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 evidence 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 (kā wolde wē ās ge-restan, for example); here is evidence about the hierarchic clustering of particle + prefixed verb (e.g., on be-cwâne). 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
- hyphen numeral hyphen
- — numeral —

represents separation of words
represents morphemic separation within words
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 syllabic 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 Ælfric’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.

---

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., grammari]. The craft is the beginning and foundation of all book-like arts. Grammaticus is one who knows the art of grammar entirely. And that craft has thirty divisions. The first division is vox (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.
If the ‘-etic’ numeral transcription is interpreted in ‘-emic’ terms, the general consistency of the spacing variations becomes clear.

This is the kind of regularity of spacing patterns to be found also in the prose ‘Alexander’s Letter to Aristotle.’ The features of spacing variations in this text fall somewhere between the pedantic style of Ælfric’s text illustrated earlier and the poetic style of the Beowulf text which now follows it in the Nowell Codex. Its syntax is principally paratactic, rather than didactic or poetic. Here is a small excerpt of edited copy.

fol. 120r 35 ðæ 0 hē 3 ðæ 3 þæ 1 wæter 3 më 1 tō 1 brōhte 4 swă 1 ic 1 ër 1 sêge 1 ðæ 1 hē 1 ic 1 më 1 wæorod 3 7 0 calle 3 ðæ 3 dugupe 4 tō-0-somme 3 7 1 hit 1 ðæ 1 be-0-foran 2 heora 2 calra 3 on-0-syn 3 nōser 3 ñ-geat 4 þyl 1 læs 1 ic 0 drunce 7 0 þone 2 mînne 4 þegn 3 þyrste 2 7 0 mînne 3 hre 7 8 ealne 3 þe 1 mid 1 më 1 wæs 4 Ond ...

[When he brought the water to me, as I said earlier, then I ordered together my troop and all my dughth, and before the view of all of them poured in out, lest I should drink and that one my thane and my army and all who were with me should thirst. And...]

There are many instances of the graphotactic consistency in syntactic phrases that repeat. For example:

109v.11 un-1-æ-0-rimed-1-lican
109r.12 un-1-æ-0-rimed-1-lican
110v.1 3 7 1 wunideñon 3
110v.4 4 ge-1-wundoñdan 3

Yet another illustration is drawn from the portion of Beowulf which was copied by the scribe of ‘Alexander’s Letter.’ Here are three small samples to show the general consistency of the graphotactics of the text: verse lines which are widely scattered have similar spacing patterns when the text repeats either exactly on or formulaic variation.

0706 35 hē 3 ne 1 mōste 3 ðæ 1 meto 3 nolde 6
0967 ic 4 ne 2 mihte 3 ðæ 1 meto 4 nolde 6

0114 lange 3 þrage 3 hē 7 him 3 ðæes 4 ëan 2 for-1-gael 1
1584 lāo-2-licu 1 lāc 4 hē 2 him 3 ðæes 4 ëan 3 for-3-gael 4

0452 On-1-send 5 hige-1-læce 5 gif 1 mec 4 hild 1 nime 7
1481 hond-2-ge-2-sellum 4 gif 1 mec 3 hild 3 nime 7
0447 ðæo-0-dröere 4 fâne 5 gif 2 mec 1 ðæ 3 nime 7
1491 ðom 3 ge-1-wyrce 3 ðæe 5 mec 3 ðæ 4 nime 6
0441 dryhtnes 4 ðome 6 sē 1 þe 5 hine 1 ðæa 4 nime 6

Fig. 2 Graphotactics of a portion of the Parker Chronicle
The extreme case of parallels occurs in the occasional ditography.

118r.1 4 ḥē 2 utan 2 wre(policy).
118r.2 4 ḥē 2 utan 2 wre(policy).
118r.14-16
3 7 1 ic 1 swiðe 3 wundrāde 3 þā 0 ge(-sālīgne)sse 4 þāre 2 eordan 1
7 0 ic 1 swiðe 3 wundrāde 4 þā 0 ge(-sālīgne)sse 4 þāre 2 eordan 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/stevick/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 syllabification for even such words as ferse-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 þet 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 pronounal, as in Se þe, but considerably more when it initiates a clause relative to a NP containing a noun head and a determinate form, as in tō þēm mere þe wē þi ge-wi−wod hæfledon.

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?

þā wolde wē | ús ge-restan.
þā mynton wē | ús ge-restan.

Or:
þā wolde | wē ús | ge-restan.
þā mynton | wē ús | 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 dā 9 wæs 4 sēo 3 priddæ 4 tid 3 þære 4 nihtæ 5 þa 3 woldæ 9 wē 1 ūs 3 ge-0-restan. 5 þa 2 cwōmān þær 3 nēdæran 3 eft 9
125r.16-17 5 þa 1 hit 2 wæs 1 sēo 1 ðifte 3 tid 3 þære 4 nihte 9 þa 2 myntōn 3 wē 1 ūs 3 ge-0-restan. 5 ac 1 þa 2 cwōmān 9 þær 3 hwīte 5 lēon 3

(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. myntōn.)

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-cūman-) 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 OND 2 Stha 7 mid 2 mēra 2 werode 3 on-1-sunde 3 in 1 patriacen 3 þ 2 lond 9 wē 0 be-1-cwōmān 3 mid 2 golde 4
111v.6 dā 2 be-1-cwōmān 3 wē 3 on 1 þa 3 lond-3-ge-2-māro 7 me 0 do 3 7 0 persa 6
111v þa 1 wē 2 tō 2 þām 1 londe 7 0 tō 1 þære 5 stōwe 5 be-1-cwōmān 4
114r.5 dā 1 be-1-cwōmān 3 wē 0 syðpan 3 tō 1 þām 3 wudum 4 in’dic 3 3 On 2 wē 4 þa 1 eft 3 in 0 fasia’cen 2 þet 111v[5] lond 4 be-1-cwōmān 3
116r 5 þa 6 usic 2 þær 2 on 1 be-1-cwōmān 6
116v.7 þa 2 fyr 3 þe 1 ūs 2 þær 2 in 0 þām 6 wicum 3 on 1 be-2-cwōmān 2
117v frīneð 4 hwæt 1 gödes 3 olde 3 yldes 3 him 1 be-1-cuμan 3 sceal 5
118r.3 dā 9 be-2-cwōm 2 ic 3 on 1 caspiam 3 þ 1 lond 3
118v.3 þa 9 laes 1 wē 1 on 0 dā 3 be-1-cwōmān 4

119r.3 þ 6 mē 2 þa 2 earfeðu 1 be-1-cwōmān 4
121v.5 5 þa 2 uisc 2 donne 3 sem’ninga 1 hwælc 3 earfeðo 4 on 1 be-2-cwōmān 5
122r.4 dā 3 be-2-cwōm 1 sum 7 on-0-gris-2-lic 7 wisse 4 on 2 hîe 3
123r 7 þa 0 wē 1 ge-3-nāg 8 rāde 3 tō 0 þām 1 be-1-cwōmān 1
124r.4 giff 2 ūs 1 on 0 nihtæ 3 un-0-cūdes 3 hwæt 3 on 1 be-2-cwōmān 5
125r.3 þāra 1 pinca 3 þe 2 ūs 1 on 1 be-1-cwōmān 3 swā 3
moniga 3 ge-0-swenc’nissa 3 7 1 earfeðo 5
125v.2 7 ge-0-swenc’nissum 2 þe 2 ūs 1 on 1 be-3-cwōm 1=1
128v be-2-cūman 3 in 0 mace’donian 4 tō 1 olmīkhraðe 5 minere 4 meder 3 7 5 mínnum 2 ge-0-swæstrum 5
128v þa 0 ic 4 eft 3 cwic 3 ne 5 móste 3 in 0 mínne 3 čēpel 4 be-0-cūman 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 linear 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, interlinear 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.

Fulk, R. D.
1998 “The role of syllable structure in Old English quantitative sound changes”, North-Western European Language Evolution 33: 3-35.

Kieman, Kevin (ed.)

Larrett Keefer, Sarah – Katherine O’Brien O’Keefe (eds.)

Malone, Kemp (ed.)

O’Brien O’Keefe, Katherine

Orchard, Andy

Rypins, Stanley (ed.)

Saenger, Paul

Stevick, Robert D.

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