Linguistic Features of Scribal Spacing

Virginia J. Cyrus

In contrast to the regularity of the letterforms in hand-produced Old English manuscripts, the scribal spacings show great variation in both size and positioning. Examination of the spacings in the Tollemache manuscript of Alfred's *Orosius* reveals extensive patternings which can be shown to correspond with such linguistic features of the text as syntactic structures as they are identified by immediate constituent analysis. Thus, the manuscript provides graphic evidence of nonsegmental features of the language which before were only indirectly inferable.

Some ancient hand-produced manuscripts may contain linguistic data which mechanically prepared language records do not. Because scholars generally have approached manuscripts prejudiced by the assumptions of modern reproductive processes, their attention has been focused primarily upon phonologically distinguishable morphemes as represented by alphabetic letterforms, or, rather imprecisely, upon words. Type-produced editions of a manuscript essentially provide only a standardized notation of the segmental text; that is to say, they reproduce a manuscript as a string of words, often with punctuation and capitalization supplied according to modern conventions. Close examination of one manuscript, the Tollemache manuscript of King Alfred's translation of Paulus Orosius's *History of the World* (British Museum Additional Manuscript 47967),[1] reveals other regular features not represented in machine reproductions. Apparently as the scribes attempted to reproduce language visually, they incorporated certain linguistic features into their transcriptions which have since been normalized or eliminated due to the mechanical demands of the printing process.

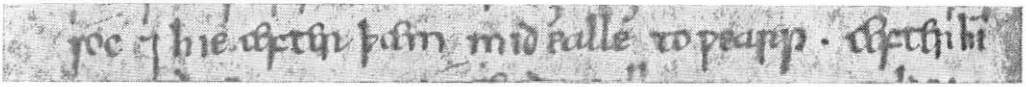
The Tollemache manuscript was selected for this study for a variety of reasons. First, because Alfred was intent on interpreting meaning rather than producing a slavish word-for-word translation,

he took great freedom with his Latin original, thereby furnishing today's scholars with a text which represents an extended composition by a native speaker of Old English. As a result the Tollemache is one of the few manuscripts upon which the grammar of Early West Saxon, the traditional norm of Old English, was constructed. Furthermore, paleographic analysis indicates that this manuscript was transcribed at the center of origin, presumably Winchester, within the quarter-century immediately following the literary activity of Alfred and his circle. Moreover, both the formation of individual symbols and the general formation of letters indicate that the manuscript was prepared by a single contemporary scribe, who, it is reasonable to assume, was a native speaker of English, and thence thoroughly familiar with the language. Because of the minimal possibility of corruption or contamination of the original text, this manuscript provides a unique record of the use of the English language 1,000 years ago.

Anyone examining the manuscript must be struck immediately by the incredible regularity of the hand-produced letterforms. Even though some pages are slightly more crowded than others, there is almost no difference in the formation of any given grapheme (loosely speaking, letterform) in its numerous occurrences on any page or series of pages. Throughout the manuscript the linear measurements of various letters in relation to each other are consistent. The linear measure of a non-terminal occurrence of the triple stroke m or the æ ligature is almost always exactly twice that of the mid-word appearance of the single stroke letters i or l. Halfway between these two extremes is the space allotted to those letters made with two vertical penstrokes—h, n, u, and r. Most graphemes made with one circular stroke or combining a straight stroke and a circular stroke—b, d, ð, f, o, p, s, þ, Þ, and the low forms of c and e—take virtually equal space along the line. The linear measurements of various occurrences of a, g, t, and x do differ slightly, but this deviation seems to be dependent on and consistent with the sequential environment in which they appear.

In contrast to the exceptional regularity of the size and shape of the letters, the linear measure of the spaces separating letters or groups of letters varies noticeably. About 8% of all spacings can be attributed to such arbitrary causes as manuscript damage, erasures,

Page 50, line 27

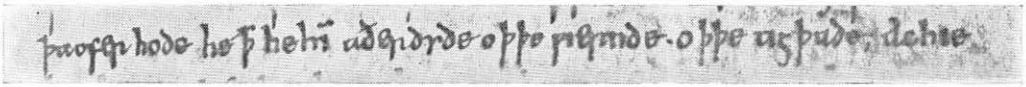


sóx ⁊ hiæfter þæm mid ealle to wearp. æfter hi(m)

Page 104, line 6

þa oþre to scipun oðflugon. ⁊ raðe þæs þe ham coman

Page 168, line 11



þa oferhode he þ(æt) he hi(m) aðer dyde oþþe wiernde. oþþe tigþade; Ac hie

Figure 1. These representative lines illustrate the uniform formation and proportionate regularity of the graphemes relative to their environments. The irregularity of the size and placement of the spacings is especially apparent when the manuscript line is compared to modern lexical spacing. From *The Tollemache Orosius, Early English Manuscripts in Facsimile*, vol. III. Reproduced by the kind permission of the publisher, Rosenkilde and Bagger.

imperfections in the vellum, interlinear corrections and the right margin line end. Even when these indeterminate spacings are disregarded, the scribal spacing still is characterized by a remarkable irregularity compared to the regularity of the letter shapes.

The scribal placement of spaces does not always comply with the modern principle of identifying each lexical unit, or "word," by spaces immediately preceding and following it; this becomes readily apparent when the manuscript is compared with Henry Sweet's edition of the text.[2] In the manuscript articles, prepositions, the negative particle, pronouns, determiners, and conjunctives frequently are not separated by space from the following "word." On the other hand, one eighth of all spaces occur within "words," although the sequence of letters between two spaces is usually at least a morpheme, a linguistically meaningful unit of speech. Also the

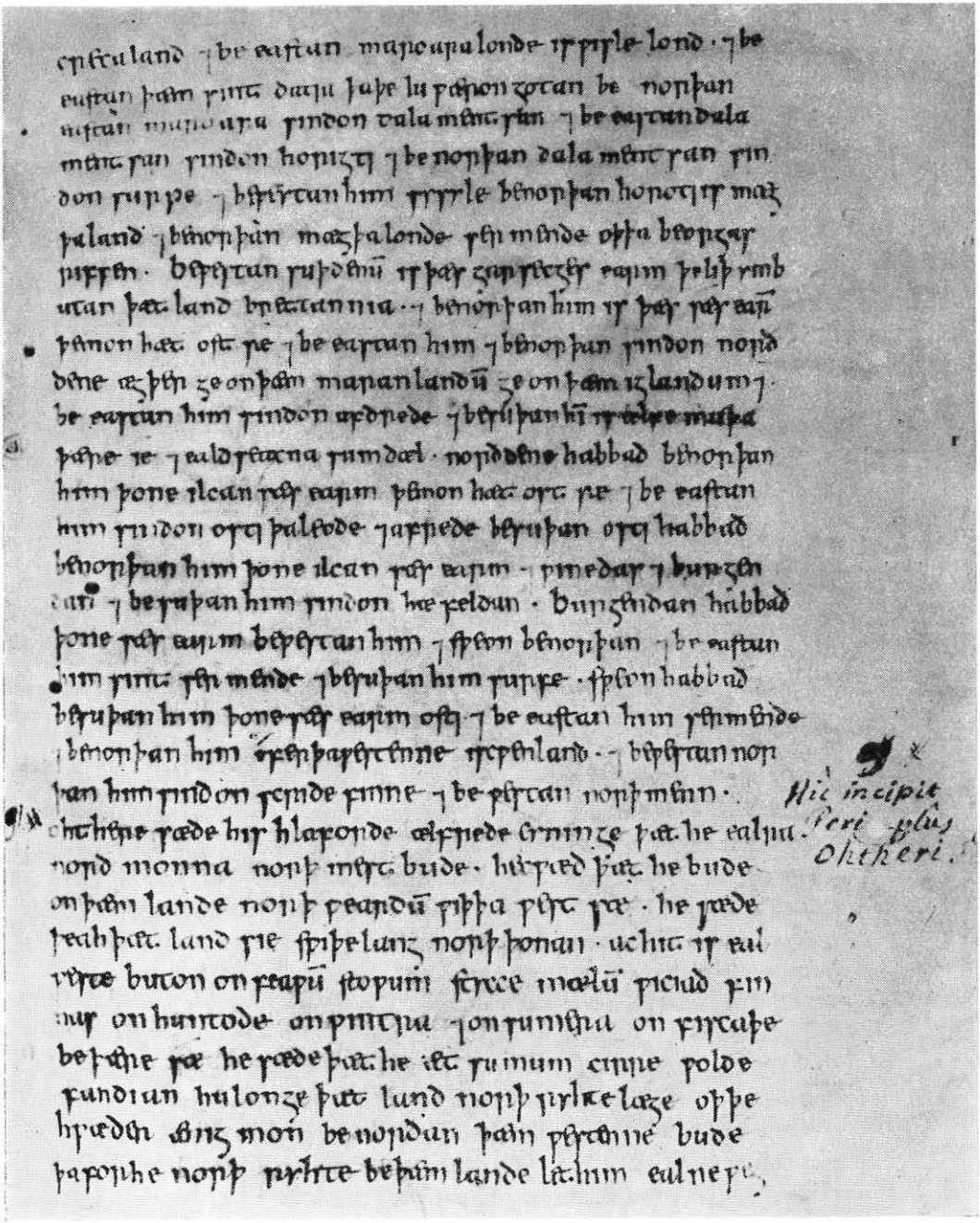
positioning of spacings shows a consistent relationship to certain classes of morphemes. A space commonly appears between the elements of a compound; prebases, postbases, and prefixes frequently are separated from their bases by spacing, while inflectional suffixes almost never are separated. Tabulations of the occurrences of different spacing widths in various morphological environments indicate that the variation of scribal spacing definitely is not random.

Scholars who work with manuscripts have noted correspondence between scribal spacing and certain linguistic signals. In his *Old English Grammar*, Alistair Campbell says, "Word-division is inconsistent in Old English manuscripts, for the accentual group obscures the word, and proclitics (e.g., articles and prepositions) are frequently joined to the following word, or divided from it by a space less than that which normally divides words. On the other hand, compounds are frequently divided into their component parts."[3] Although Campbell does not pursue this point, his identification of the relationship between the spacing patterns and "accentual groups" provides the key to the significance of the scribal spacings.

Examination of the manuscript in this light reveals that both the positioning and the variation of the width of the scribal spacings display extensive patterning which coincides with syntactic (word order) as well as morphotactic features of the text. Briefly, morphtactics describe the distribution of the basic elements of speech, both segmentals, that is recurrent and meaningful sound sequences such as bases (root words), affixes, and pre- and post-bases; and suprasegmentals, that is, prosodic features like stress, pitch, and juncture.

Even though page 13 of the manuscript (Fig. 2) does not provide the most striking example of spacial patterning, it was selected as an illustration for several reasons. It provides samples of both very simple sentence structure (11.1—21) and more complex constructions (11.22–31). Also the vocabulary on this page should seem less

Figure 2. Page 13 of *The Tollemache Orosius, Early English Manuscripts in Facsimile*, vol. III. Reproduced by the kind permission of the publisher, Rosenkilde and Bagger.



105 Cyrus : Linguistic Features of Scribal Spacing

foreign than most to readers unfamiliar with Old English. Finally, since Alfred's description of northern European geography and his account of Ohthere's voyage are supplements to the Latin original, they represent totally free composition in Old English by a native speaker.

In the following discussion of the text a more or less arbitrary code will be used to indicate the measure of the scribal spaces according to the following scale:[4]

0 not enough room for a single penstroke between letters at a morph boundary which elsewhere is marked by spacing.
1 the width of a penstroke. This is the standard size of the space between most non-contiguous letters within a single syllable.
2 the width normally allotted to a grapheme formed with a single penstroke, such as i or l, when it appears internally in a morpheme. Spaces of this size occur most frequently between words.
3–6 multiples of the above.
9 line terminal.

Of the 38 major syntactic units, or clauses, on page 13, twelve end with pointing and all but three end with a space of at least size 3. Two of the exceptions are the closely linked complex constructions "he sæde þeah þæt" (he said however that) and "he cwæþ þæt" (he said that). The third is the unaccountable size 1 space which separates "iglandum" from the "⁊" (and) followed by a point at the end of line 10.

All but six units end with a space as large as or larger than any space which occurs within them. Three of the exceptions seem to be merely the result of vagaries of human transcription: the end of line 10 mentioned before; the wide space (size 5) between "be" and "norþan" at the end of line 2; irregularities not only in spacing but even in the straightness of the last line on the page. The other three exceptions are probably more significant: in line 22 a size 4 space sets off the appositional "ælfrede cyninge" (Alfred the king); the two size 4 spaces on line 5, following "surfe," and on line 6, preceding "sermende," both seem to be related to the fact that in the compound constructions the verb "to be" is understood.

Close examination of one passage from midpage (11.16–21) will illustrate more clearly than generalizations the astonishing degree of correspondence between sequences identified by the spacings and

principle constituent boundaries independently inferable from syntactic features. Boundaries of major syntactic constituents, or sentences, and the internal relations of constructions are signalled by word order, function words, parts of speech categories and inflections. Using these syntactic signals, the boundaries of the principle constituents are marked for these lines. To facilitate modern preconceptions, spaces in this sample are positioned on the basis of lexical units or words; superscript numerals indicate the measure of the scribal spacing at the point of each syntactic cut. The first cut produces the major syntactic units or clauses:

> [5]burgendan habbað þone sæs earm be westan him & sweon be norþan[4] / & be eastan him sint sermende & be suþan him surfe.[5] / sweon habbað be suþan him þone sæs earm osti & be eastan him sermende[9] / & be norþan him ofer þa westenne is cwen land.[4] / & be westan norþan him sindon scride finne & be westan norþ menn.[9] /

Translation:

> [5]Burgendans have the sea arm to the west of them and Swedes to the north[4] / and to the east of them are Sermendes and to the south of them Surfes.[5] / Swedes have to the south of them the sea arm Osti and to the east of them Sermendes[9] / and to the north of them over the waste is Queenland.[4] / and to the west-north of them are Scride-finns and to the west North-men.[9] /

Size 5 spaces with pointing both precede the passage and separate its two main thought units, the Burgendan and the Sweon. Of the five major syntactic units (independent clauses), three end with pointing and all end with size 4 or 5 spaces or at the end of a manuscript line. Further cutting produces the following boundaries:

> [5]burgendan habbað þone sæs earm be westan him[3] / & sweon be norþan[4] / & be estan him sint sermende[3] / & be suþan him surfe,[5] / sweon habbað be suþan him þone sæs earm[2] / osti[3] / & be eastan him sermende[9] / & be norþan him[3] / ofer þa westenne[4] / is cwen land.[4] / & be westan norþan him sindon scride finne[4] / & be westan norþ menn.[9] /

These markings distinguish the major constructions and their main constituents. The correspondence between these syntactic boundaries and the sequences identified by the spacings can be seen if the

numerals which indicate the widths of all manuscript spacings are superimposed.

> ^{5}burgendan^{2} habbað^{9} þone^{2} sæs^{2} earm^{2} be^{0} westan^{2} him^{3} / &^{2} sweon^{2} be^{0} norþan^{4} / &^{2} be^{2} eastan^{9} him^{2} sint^{2} sermende^{3} / &^{1} be^{0} suþan^{2} him^{2} surfe.^{5} / sweon^{2} habbað^{9} be^{0} suþan^{2} him^{2} þone^{1} sæs^{2} earm^{2} / osti^{3} / &^{2} be^{2} eastan^{2} him^{2} sermende^{9} / &^{2} be^{0} norþan^{2} him^{3} / ofer^{1} þa^{0} westenne^{4} / is^{0} cwen^{1} land.^{4} / &^{2} be^{0} westan^{1} nor^{9} þan^{2} him^{2} sindon^{2} scride^{2} finne^{4} / &^{2} be^{2} westan^{2} norþ^{1} menn.^{9} /

All spaces size 3 or larger coincide with the elementary syntactic cuts.

Rather than making further cuts, it is easier and just as revealing to consider the locations of the least admissible cuts in terms of immediate constituent analysis, the purpose of which is to give the maximum number of actually occurring morphemes, words, or phrases at each step. Such cuts as "& sweon be / norþan" (and Swedes to / the north) or "ofer þa / westenne" (over the / waste) do not produce constituents consistent with Old English syntax which may be inferred independently. Minimal scribal spacings usually occur at such points where major syntactic boundaries cannot be marked: "&^{2} sweon^{2} / be^{0} norþan^{4}" and "ofer^{1} þa^{0} westenne^{4}."

Such syntactic parsing does not correspond absolutely with the scribal spacings, but it would be unrealistic to expect complete consistency of any feature in a text the length of the Orosius. Furthermore, it is likely that some deviations may be affected by different features than those which generally determine spacing. For example, some spacing variations are probably determined or influenced by such factors as the stresspeak in a phrase or phonotactics, that is patterns of distribution and clustering of phonemes (sounds). One obvious example of this is the spacial separation of the contiguous vowels in "be^{2} eastan" where the corresponding syntactic units are represented with no spacial separation: "be^{0} norþan," "be^{0} westan," and "be^{0} suþan."

However, this paper is concerned primarily with those features which control the pervasive systematic patterning of the scribal spacing rather than with minor patterns or inconsistencies. On the basis of the evidence and in view of the well established and systematic correlation between graphemes (letterforms) and phonemes (sounds)—even though all vowel phonemes are not distinguished in the spelling—it is unlikely that the size and placement

of scribal spacings are governed by segmental features of the text. Instead, the degree of correspondence between the various ranges of spacings and the boundaries independently inferable from syntactic signals shows that syntactic features are clearly correlated with the spacings. The correlation is to be found, presumably, in the relation of prosody and syntax.

In spoken language syntax and prosody are complementary aspects of sentence structure. The prosodic (suprasegmental) features of juncture, stress and pitch correlate with the syntactic structure of the construction as a whole and identify the constituents of that construction. The intonation features coordinate with juncture to mark the boundaries of a construction (phrase, clause, etc.); however, it is juncture, characterized either by a pause or by the prolongation of the phonemes of the final stressed syllable or by both, which signals the end of the construction. In the linear representation of speech in the Old English manuscript, widest spacing occurs at boundaries of principle syntactic constructions. The increased spacing sizes then seem to imply these boundaries, and therefore the junctures; narrow spaces, on the other hand, would deny such boundaries and indicate that juncture is inadmissible. Probably because the production of a written record of speech is a slower, more deliberate act than that of speech, the spacings in the manuscript identify the boundaries of virtually every contruction as it would be signalled prosodically in speech; or in prosodic terms, the spacings signal nearly every potential juncture.

Once juncture is determined, the stress of the construction (phrase or clause) may be positioned, for the strongest stress usually occurs at or near the end of the construction. Although segmental morphemes do not have stress as an integral part of their structure, words and phrases do, for no word can be pronounced except as a phrase or part of a phrase, and individual words normally preserve the identity of their stress patterns in the context of an utterance or construction. So prosodic stress patterns are determined primarily by the morphological structure of the constituent words. The unconventional size and placement of certain spacings in the Old English manuscript, like size 0 spaces preceding or following such words as particles, determiners, conjunctives, prepositions, and even some personal pronouns, suggest that the stress pattern of the larger construction

suppresses, obscures, or overrides the individual lexical unit or word. In short, the scribe seems to have followed the phrasing of normal speech, rather than word boundaries, in placing and varying the size of the spacings.

The third suprasegmental, the pitch pattern, can be inferred when the grammatical stress is positioned, for the pitchpeak usually accompanies the strongest stress of a construction. Although it can be shifted to other positions, its precise onset is ultimately determined by the stress pattern of the phrase or its constituents, and it is always bound by a terminal juncture. Therefore, the scribe who prepared this manuscript, although possibly neither intentionally or even consciously, has left a graphic record by a native speaker of the general characteristics of Old English from which the prosody can be reconstructed.

However, syntactic constituents and correlative prosody are only two linguistic features that in all probability influenced the scribal spacing. They certainly indicate that the variations in scribal spacings are meaningful, but the full extent of the meaning will be clear only after all recurrent features and structural variables on all levels from phonological features to grammatical forms and larger construction patterns have been examined in relation to the spacings.

1. *The Tollemache Orosius*, ed. Alistair Campbell, Vol. III, *Early English Manuscripts in Facsimile* (Copenhagen: Rosenkilde and Bagger, 1953).
2. Henry Sweet, ed., *King Alfred's Orosius* (London, 1883).
3. Alistair Campbell, *Old English Grammar* (Oxford: Clarendon Press, 1959), p. 14.
4. While the existence of all kinds of pointing has been taken into account, the measure of the spaces which contain pointing has not been broken into that which precedes and that which follows the pointing itself. The measure, therefore, incorporates the pointing. Also, in determining each measure, variations in letterforms such as the trailing crosshatch of the terminal e are not taken into account. Instead, the measure is taken from the "edge" of the normalized letterform.