

Lo spazio lessicale della poesia antico inglese

Giuseppe Brunetti

abbreviazioni

| | | |
|--------------|---|--------------------------|
| An | = | Andreas |
| Beo | = | Beowulf |
| ChrA | = | Christ A |
| ChrB | = | Christ B |
| ChrC | = | Christ C |
| Dream | = | Dream of the Rood |
| El | = | Elene |
| Eleg | = | Elegies |
| Ex | = | Exodus |
| GenA | = | Genesis A |
| GenB | = | Genesis B |
| GuthA | = | Guthlac A |
| GuthB | = | Guthlac B |
| Jud | = | Judith |
| Jul | = | Juliana |
| Mald | = | Battle of Maldon |
| Max | = | Maxims |
| Phoen | = | Phoenix |
| Phys | = | Physiologus |

Stile lessicale

Per stile lessicale qui si intende la distribuzione delle classi lessicali in un testo, con l'aggiunta di composti e parole poetiche (quelle usate solo in poesia). Prendendo in considerazione solo le forme, cioè il numero di occorrenze dei lemmi, ecco i profili di *Maldon*, *Phoenix* e *Beowulf*:

| | noun | verb | adj | adv | pron | prep | conj | other | pnoun | total | cpnd | poetic |
|--------------|------|------|------|------|------|------|------|-------|-------|-------|------|--------|
| Mald | 508 | 475 | 115 | 208 | 409 | 176 | 133 | 4 | 64 | 2092 | 66 | 186 |
| % | 24.3 | 22.7 | 5.5 | 9.9 | 19.5 | 8.4 | 6.4 | 0.2 | 3.1 | 100.0 | 3.1 | 8.9 |
| | | | | | | | | | | | | |
| Phoen | 1284 | 608 | 395 | 318 | 399 | 303 | 259 | 8 | 20 | 3594 | 214 | 251 |
| % | 35.7 | 16.9 | 11.0 | 8.8 | 11.1 | 8.4 | 7.2 | 0.2 | 0.6 | 100.0 | 5.9 | 7.0 |
| | | | | | | | | | | | | |
| Beo | 5231 | 3498 | 1455 | 1566 | 2649 | 1170 | 960 | 95 | 641 | 17265 | 1535 | 1816 |
| % | 30.3 | 20.3 | 8.4 | 9.1 | 15.3 | 6.9 | 5.6 | 0.5 | 3.7 | 100.0 | 8.9 | 10.5 |

Testi di così diversa lunghezza (2092, 3594 e 17265 parole) non sono direttamente comparabili attraverso il confronto delle percentuali: serve il confronto con una norma – quella del corpus a cui appartengono.

Il corpus

Sono stati fatti i conteggi di circa la metà dell'intero corpus poetico antico inglese: 19 testi per un totale di oltre 15800 versi e quasi 90 mila parole (occorrenze o forme di 9302 lemmi). Sono state accorpate le *Elegie*, le *Massime* e il *Fisiologo*, in modo che il testo più breve fosse di almeno 1000 parole. Ecco i dati aggregati:

| | noun | verb | adj | adv | pron | prep | conj | other | pnoun | total | cpnd | poetic |
|---------------|-------|-------|------|------|-------|------|------|-------|-------|-------|------|--------|
| corpus | 27118 | 17166 | 7975 | 7672 | 14644 | 7048 | 5500 | 433 | 2018 | 89474 | 6054 | 6772 |
| % | 30.3 | 19.2 | 8.9 | 8.5 | 16.4 | 7.9 | 6.1 | 0.5 | 2.3 | 100.0 | 6.8 | 7.5 |

Frequenza teorica e frequenza effettiva

Lo stile lessicale di un testo può essere allora rilevato dal confronto con i valori del corpus. La statistica adoperata è la differenza tra la frequenza teorica e quella effettiva. Il *Beowulf*, per esempio, con le sue 17265 forme è poco meno di un quinto del corpus, e la frequenza teorica dei suoi nomi è perciò di circa un quinto dei nomi del corpus: la differenza tra la frequenza effettiva e quella teorica indicherà se il testo ha un eccesso o un difetto di nomi, se è sopra, sotto o nella norma.

Scarti ridotti

Questa differenza viene data in termini di scarti ridotti (*z-scores*), che hanno il vantaggio di indicare direttamente se essa è significativa o no: lo è, se lo scarto è superiore a +2 o inferiore a -2 (cioè due deviazioni standard).

| | noun | verb | adj | adv | pron | prep | conj | other | pnoun | total | cpnd | poetic |
|-------|-------|------|------|------|------|------|------|-------|-------|-------|-------|--------|
| An | 3.2 | 0.3 | 3.6 | -0.1 | -4.4 | 2.2 | -3.7 | 0.3 | -5.9 | 9302 | 6.3 | 8.2 |
| Beo | 0.0 | 3.6 | -2.4 | 3.0 | -3.7 | -5.7 | -3.5 | 1.4 | 14.1 | 17265 | 11.9 | 15.7 |
| ChrA | -2.6 | -0.7 | -0.4 | 2.3 | 2.7 | -0.9 | 2.2 | 0.3 | -1.1 | 2487 | -3.1 | -4.7 |
| ChrB | 2.9 | -1.7 | -0.7 | -1.5 | -1.0 | 1.7 | 1.3 | -0.8 | -3.5 | 2434 | -1.7 | -2.6 |
| ChrC | -1.3 | -2.4 | 3.9 | 0.0 | 1.0 | 2.4 | 2.9 | -1.5 | -7.0 | 4708 | -5.0 | -14.8 |
| Dream | -3.1 | 1.2 | -1.7 | 0.9 | 4.8 | 1.5 | -1.2 | -1.0 | -3.4 | 1109 | -3.5 | -5.4 |
| El | 1.1 | -1.7 | -1.4 | -0.4 | -1.4 | 5.5 | -2.1 | 2.2 | 0.4 | 7291 | -0.4 | 1.2 |
| Eleg | -1.1 | 1.6 | 1.0 | 3.0 | 0.1 | -1.7 | 0.1 | -0.8 | -5.3 | 2220 | 2.0 | 0.4 |
| Ex | 8.0 | 0.1 | 4.2 | -6.0 | -7.2 | -0.7 | -3.5 | 0.7 | 0.0 | 2985 | 10.2 | 6.1 |
| GenA | 2.0 | -3.0 | -3.6 | -1.7 | -1.3 | -3.0 | 3.7 | 8.3 | 11.2 | 12608 | -1.8 | 0.0 |
| GenB | -10.3 | 2.3 | -5.3 | 3.0 | 13.2 | -0.6 | 1.5 | -1.3 | -0.8 | 4525 | -11.1 | -11.4 |
| GuthA | -2.6 | 1.8 | -4.4 | -3.2 | 6.6 | 0.1 | 3.2 | -3.9 | -2.1 | 4825 | -8.6 | -11.9 |
| GuthB | 2.0 | -1.6 | 3.9 | 0.0 | -1.8 | 1.1 | -1.3 | -1.8 | -4.6 | 3113 | 2.8 | 1.7 |
| Jud | -0.9 | -1.9 | 5.1 | 1.0 | 1.0 | 0.0 | -1.4 | -2.8 | -2.0 | 1999 | 3.1 | 6.8 |
| Jul | -2.8 | -1.1 | 1.7 | -1.4 | 4.3 | 1.2 | 1.5 | -2.3 | -2.2 | 4101 | -1.5 | -5.5 |
| Mald | -5.1 | 3.7 | -5.3 | 2.3 | 3.6 | 0.9 | 0.4 | -1.9 | 2.5 | 2092 | -6.4 | 2.2 |
| Max | 0.7 | 5.3 | 1.4 | -2.7 | -1.5 | -1.4 | -1.9 | -2.4 | -4.2 | 1866 | -5.5 | -0.6 |
| Phoe | 6.0 | -3.2 | 4.3 | 0.8 | -8.0 | 1.2 | 2.6 | -2.3 | -6.9 | 3594 | -1.9 | -1.3 |
| Phys | 0.8 | -2.4 | 1.8 | 0.0 | -0.1 | 2.2 | 0.7 | -1.2 | -3.8 | 950 | 1.2 | -0.3 |

Lo spazio lessicale

C'è un modo efficace di visualizzare l'insieme di queste differenze. Le variabili lessicali (=classi, composti e parole poetiche) possono essere pensate come coordinate di uno spazio multi-dimensionale in cui sono collocati i testi: la statistica del *multi-dimensional scaling* consente di ridurre le dimensioni in forma rappresentabile.

Due grafici complementari rappresentano le distanze dei testi fra loro e dalla norma, e le distanze delle variabili fra loro.

La norma è il quadrato nero al centro: l'ipotetico testo con scarto zero. Sono stati esclusi i nomi propri, perché legati al genere: la poesia narrativa ne ha tanti, la gnomica e l'elegiaca pochi o nessuno.

Fig. 1 Lo spazio lessicale: i testi

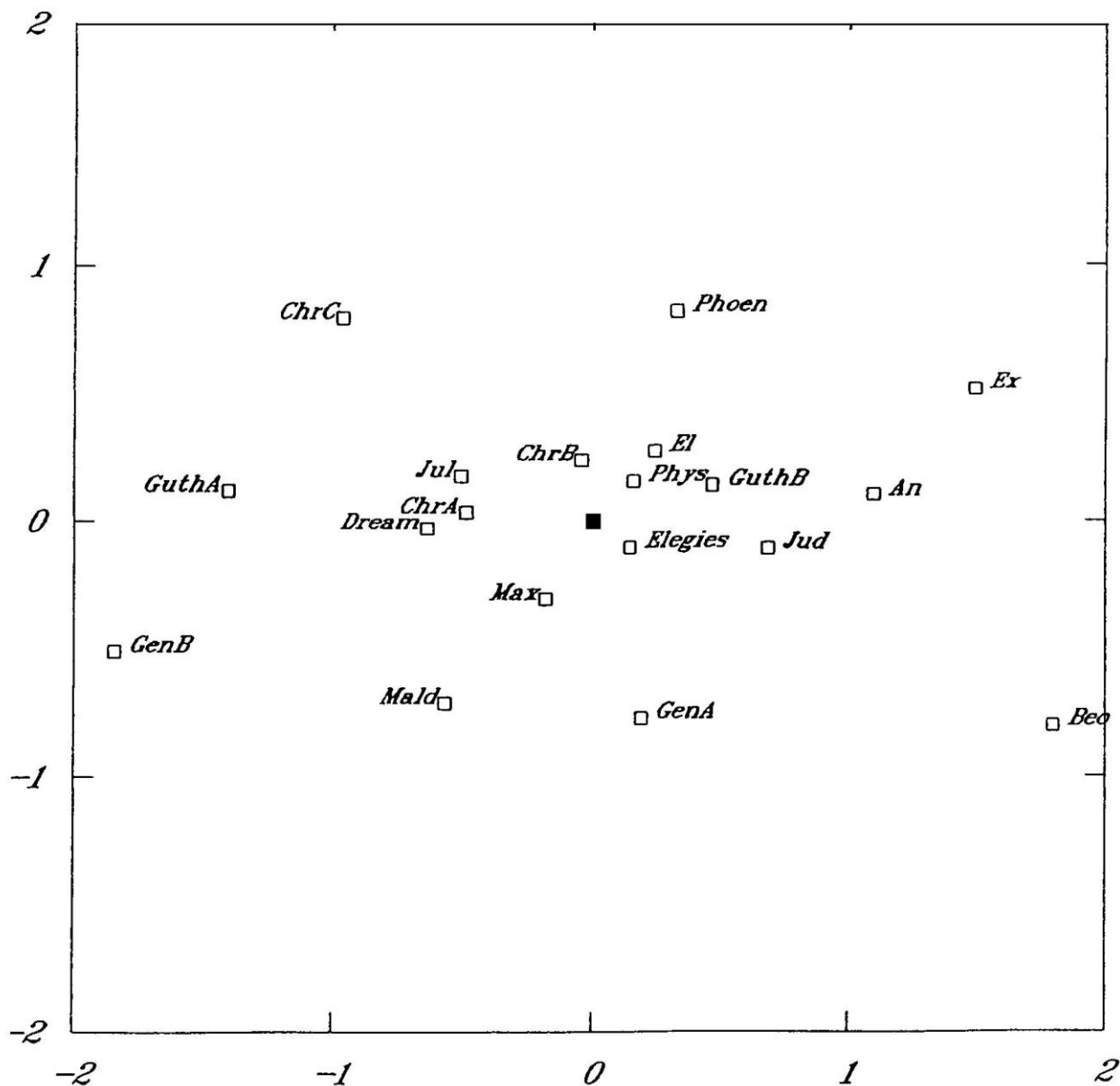
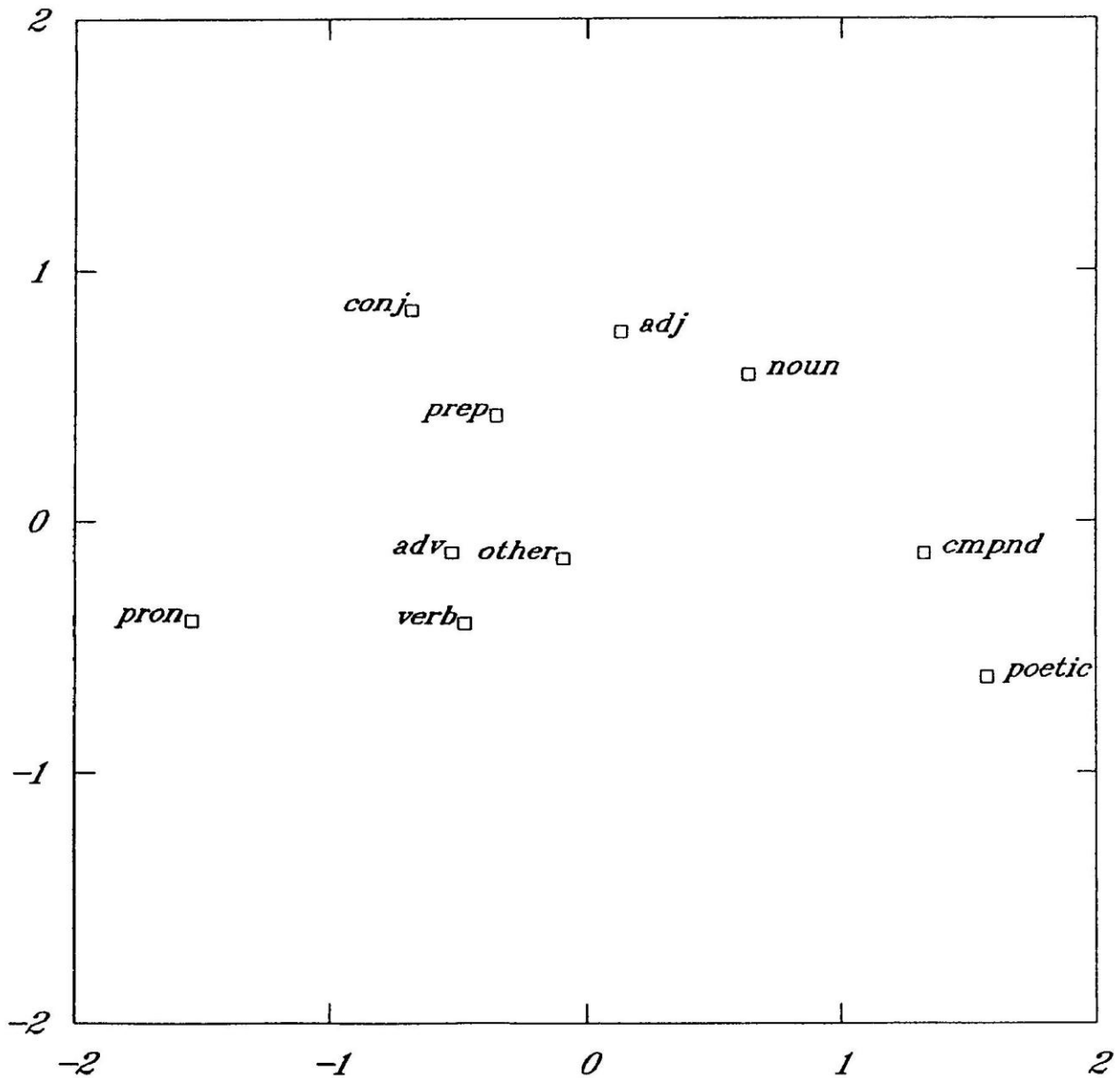


Fig. 2 Lo spazio lessicale: le variabili



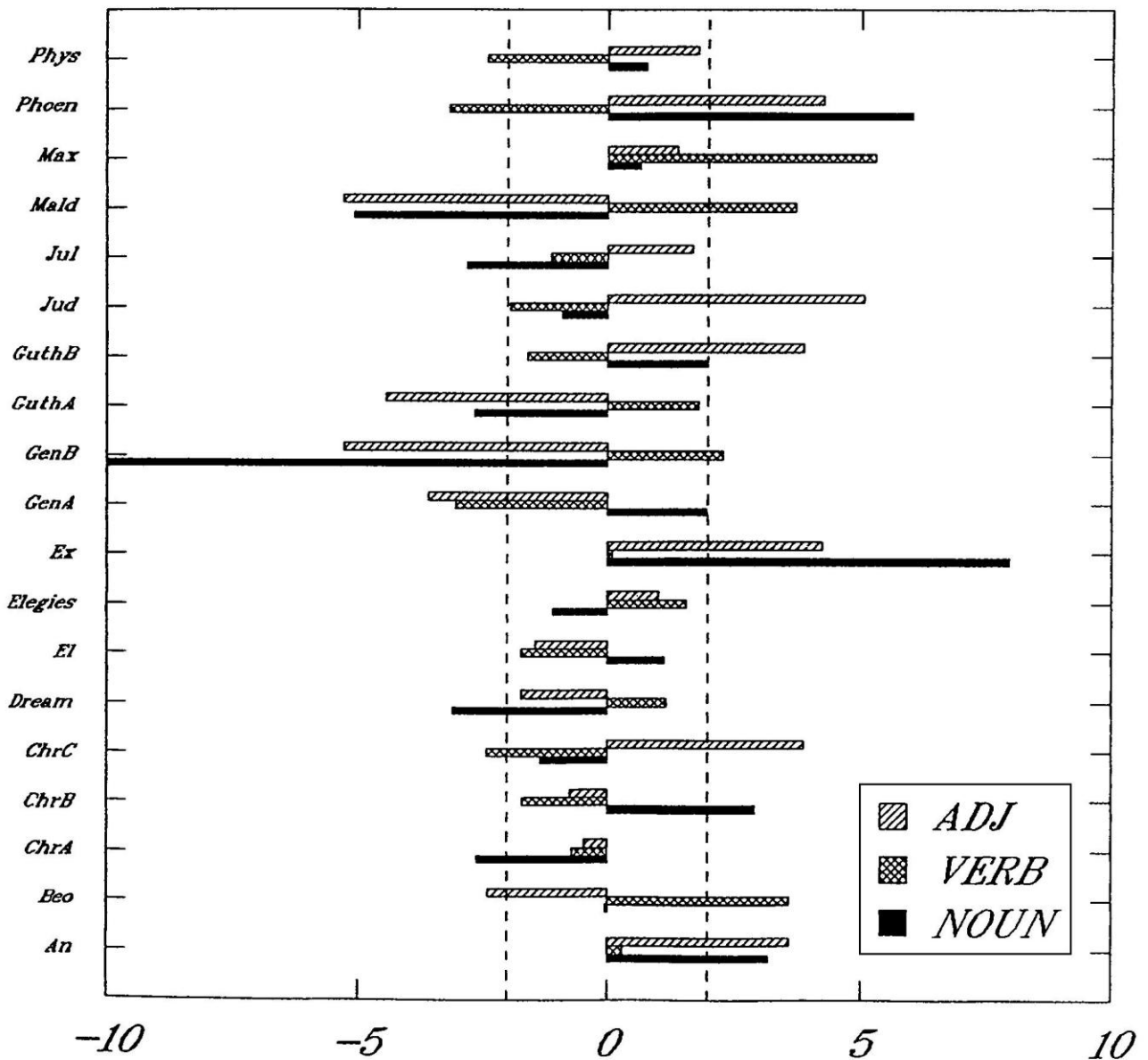
Il secondo grafico indica la forza e la direzione delle variabili nel distanziare i testi. La variazione maggiore è lungo l'asse orizzontale: verbi e parole funzionali sono tutti a sinistra; aggettivi, nomi, composti e parole poetiche tutti a destra. Quello che varia lungo quest'asse – dal minimo di *Genesis B* al massimo di *Beowulf* – è la nominalità, anche qualitativa: composti e parole poetiche sono per tre quarti nomi.

Beowulf ha una marcata scarsità di parole funzionali (è il solo testo in cui pronomi, preposizioni e congiunzioni sono tutti significativamente sotto la norma), le quali invece abbondano spropositatamente in *Genesis B*. La lontananza del *Beowulf* è data da composti e parole poetiche.

Stile nominale o verbale

I valori delle singole variabili (=classi lessicali) possono essere rappresentati come lunghezze negli istogrammi del tipo profitti e perdite: la banda tra +2 e -2 indica la norma.

Fig. 3 Nomi, verbi e aggettivi



Dal grafico di nomi, verbi e aggettivi si vede che lo stile lessicale del *Beowulf* (il secondo dal basso, *Beo*) è attorno alla norma: i suoi nomi hanno scarto zero, mentre ha un piccolo difetto significativo di aggettivi e un piccolo eccesso significativo di verbi.

Stile nominale è quello di *Exodus* (il nono dal basso, *Ex*), con il suo marcato eccesso di nomi e aggettivi; così anche *Andreas* (primo dal basso, *An*) e la *Fenice* (seconda dall'alto, *Phoen*) – quest'ultima anche con un difetto significativo di verbi.

Si può dire invece verbale lo stile della *Battaglia di Maldon* (quarto dall'alto, *Mald*), in cui verbi da una parte, e nomi e aggettivi dall'altra sono quasi proporzionalmente divaricati oltre le due soglie.

Judith ha un alto tasso di aggettivi. Colpisce la forte denominalizzazione di *Genesis B*.

Elegie ed *Elena* sono interamente nella norma.

Lessico caratteristico

I verbi modali sono un esempio di lessico caratteristico: indicano la modalizzazione dell'azione – il potere, il dovere, il voler fare. *Maldon* si distingue per l'uso di **willan**: dei guerrieri che fuggono dalla battaglia si dice, non che lo fecero per paura o sbandamento, ma che fuggirono perché “non vollero esservi” (v. 185). Altri quattro testi hanno scarti significativi: le *Elegie* per **magan**, *Genesis B* per **sculan** e **magan**, *Guthlac A* per **willan** e **motan**, e le *Massime* per la nota funzione gnomica di **sculan**.

Fig. 4 Verbi modali

