

“Smudges on the canvas”?

A corpus stylistics approach to *Macbeth*

Sonia Zyngier Universidade Federal do Rio de Janeiro

I Introduction

In this paper I intend to demonstrate how corpus linguistics may be an interesting method for stylistic analysis. The title comes from L.C. Knights, who, in 1933, argued that “The main difference between good and bad critics is that a good critic points away from the work in question; he introduces extraneous elements into his appreciation – smudges the canvas with his own paint”(1933:33). Although he does say we have to add elements which are not in the text, paradoxically he also holds that “we have to allow full weight to each word, exploring its ‘tentacular roots’, and to determine how it controls and is controlled by the rhythmic movement of the passage in which it occurs” (p. 31). Sixty-three years later, Sinclair confirmed that “the starting point of the description of meaning in language is the word” (1996:75), which may mean that although influenced by a late romantic trend in literary criticism, Knights was in fact offering a vision of a discursal approach to texts, one in which we should look at collocational patterns to arrive at more solid interpretations. And he added: “... **precision** and **particularity** are exactly what is lacking in the greater part of Shakespeare criticism, criticism that deals with *Hamlet* or *Othello* in terms of abstractions that have nothing to do with the unique arrangement of words that constitutes these plays”(p. 32). What I want to claim here is that with the development of corpus linguistics we now have access to tools and methodologies which may provide more substantiation to stylistic investigation.

II Some definitions and assumptions

I start with a problem: how can we apply corpus analysis to literary texts? If, as Sinclair notes, “one of the principal uses of a corpus is to identify what is central and typical in language” (Sinclair 1991:17) you would probably agree with me that corpus linguistics

does not apply to the study of the language of literary texts precisely because writers tend to avoid typicality. Literary texts definitely cannot be considered “sample texts”. Each text is a unique instance.

But 6 decades after the Formalist paradigm, it would be hard to sustain that a text is literary because its language is unique and creative for at least two basic reasons. First, Gibbs (1994) and other scholars like Carter at the last PALA conference have demonstrated widely how everyday use of language can also be highly innovative. Gibbs holds that the so-called literary language (tropes, metaphors, metonymy, etc.) actually “constitute basic schemes by which people conceptualize their experience and the external world”(p. 1). Second, literary value depends on many different conditions established by a given community at a certain time. So what I believe corpus linguistics can do is to point out the environment of certain lexical items as used by an author to allow some statements about possible interpretations which are demonstrable. This specific use may also be set against contemporary texts so as to verify their degree of creativity. This is easier with 20th century literature, where access to all sorts of written materials is available. Although I had access to the 17th century Helsinki Corpus, the texts on which that corpus is built do not give us enough coverage of language usage of the time. This may pose a problem to interpretation. In the case of non-contemporary texts, when the written mode was not as widespread as today and when there was no means of recording spoken language, it is never enough to remind ourselves that we are limiting the investigation to a corpus which contains centrally the writer’s productions and peripherally that of his contemporaries.

Another possibility is to build a corpus of the writer’s readings. Louw¹ suggests that where we know what someone read, we know that it is part of their prior knowledge and intertextuality. The point would then be to build a corpus of an individual’s intertextuality. However, the difficulty in relation to early writers remains. The debate

¹ Personal communication, June 1998.

over which texts Shakespeare must have read has produced volumes of scholarly work and inconclusive statements.

Notwithstanding these limitations, I still think that non-contemporary literary texts can be studied against some sort of norm. These may be a collection of other texts from the same period or region, or the collected works of the same author. In the case of Elizabethan/Jacobean English, it is hard for me to talk about *corpora*, as I did not have access to a large collection of texts in electronic form. My corpus consisted of Shakespeare's Complete Works in the Oxford Electronic Texts, the 17th c Helsinki Corpus, and about 162,000 words from Holinshed's *Chronicles*², which I scanned.

Even looking at features of a text from a small collection of works, I would argue that we may still borrow the methodology and tools from corpus linguistics to look at literary texts in ways which have not been available so far. This may not only provide more substantiation to intuitive interpretations but may also help contradict interpretations which have been taken for granted. The main question then was to find out what properties of an observation distinguished it from the rest as having interpretative importance. I started both from overall frequency of occurrence, which could help establish relevance of the items to be studied and an observation of the collocational patterns of usage.

I assumed that adopting and adapting the general principles of corpus analysis, we could develop a bottom-up strategy, looking at instances of language first in order to arrive at generalizations about the significance of certain patterns and then use a top-down approach for other points. This method has strong implications for the teaching of literature. Instead of providing students with interpretations, they can find out for

² Due to the difficulty in obtaining a complete version of Holinshed's *Chronicles* in electronic form, I scanned *The Words of Holinshed as read by Shakespeare*, edited by A. Nicoll & J. Nicoll, London: J.M. Dent & Sons Ltd., 1843. This is an abridged version of the original, but it is presented in a machine-readable modern font. Although not ideal, the scanning provided me with a corpus of 161,755 words. I would like to thank the Postgraduate Programme in Psychology at the

themselves how language is used, what patterns are formed and how significant that may be for interpreting a literary text. As Miall (1990:7-8) notes, “the computer ... tends to foreground the act of inquiry, to require an active grasp of the implications and limitations of particular research procedures... The flexibility of the computer as a tool, together with the fact that attending to the screen shifts attention away from the teacher, encourages students to engage more actively in the debate that is central to the Humanities”.

The research method used in this paper is both quantitative and qualitative, since I both produce statistical accounts (the frequency lists) and from the list pick out some items to assess their behaviour within their context in the corpus. I believe this new approach may bring new kinds of evidence that may help us validate and privilege certain interpretations and perhaps even arrive at interpretations never offered before. Again I quote Miall (op.cit:11), who says that “the computer... offers to alter our perspective, shifting is from those angles of view where we or our students have perhaps grown too accustomed to standard responses and familiar modes of study”.

III The account

A Getting started: frequency lists

My first step was then to use the Shakespeare Corpus at the University of Birmingham to produce a frequency list of all the word-forms both in *Macbeth* (MAC) and in the Shakespeare Tragedy Corpus (STC)³. According to Sinclair (1996:80-81), “... the use of numerical methods is normally only the first stage of a linguistic investigation... [where] ... the focus is on repeated events rather than single occurrences”. Repeated events provide an interpretative framework by means of which we can investigate single

Universidade Estadual do Rio de Janeiro, particularly Rodolfo Ribas, who helped me with the scanning.

³ I am indebted to Dr. Geoff Barnbrook who sent me these initial lists via e-mail and with whom I had a pleasant and enlightening conversation during a two-day stay in Birmingham in July 1997.

occurrences. The intention was to verify which lexical item we would consider worth of investigation. Although this decision was ultimately arbitrary it did not invalidate the general point. In order to establish interpretative significance, the investigation followed a pragmatic orientation, one in which priority was given to the effect the text has on the reader, a view in line with reader-response theories. These theories assume the existence of structures which may not be immediately apparent to the reader but are accountable for the literary effect. It is possible that this effect may be achieved either by frequency of the use of certain lexical items, by the unusualness of single occurrences, or even by the absence of some of the items, of what has not been selected.

Starting very informally from a point I had heard before that Macbeth was a man of “shoulds” and “woulds” and that the play was future-oriented, I decided to look at the verb-forms in the play.

The apparatus was not too sophisticated: a 486 PC-compatible home computer and MicroConcord as the software. Geoff Barnbrook from the University of Birmingham was very kind to send me the frequency list of both the play and the entire Tragedy Corpus. From that, I produced the frequency list of **verbs** comparing the most frequent occurrences:

	MAC (17,839 tks)		STC (101,352 tks)	
Is	190	1.06%	1,072	1.05%
Be	139	0.77%	730	0.72%
Have	122	0.68%	672	0.66%
Do	84	0.47%	572	0.56%
Enter	76	0.42%	295	0.29%
Will	75	0.42%	513	0.50%
Are	74	0.41%	420	0.41%
Shall	68	0.38%	373	0.36%
Come	67	0.37%	328	0.32%
Hath	52	0.29%	231	0.22%
Was	48	0.26%	252	0.24%
Would	47	0.26%	261	0.25%

Table 1: Frequency list of verbs

Setting a comparison between the frequency list of verbs in the entire Shakespeare corpus (hence, STC) and in *Macbeth* (hence, MAC) did not lead me too far in terms of whether this play would differ significantly from the other tragedies in terms of verb choices. Verbs like **will** or **shall** that could be relevant in terms of the question of time were not characteristically more frequent in *Macbeth* than in the Shakespeare Tragedy Corpus. Of course one could argue that time is not exclusively conveyed by these verbs. There may be other markers of time in the play which one could investigate. What this search did was to rule out the possibility of it being conveyed by **will** or **shall**. What was interesting was to check why **enter** was more frequent in *Macbeth* than in the other tragedies. Such a frequency, especially in the stage directions, indicates that this is a dynamic play with a lot of movement on the part of the actors, especially Macbeth, who does most of the enterings. The following table shows the most frequent verbs in MAC as opposed to the most frequent ones in the STC:

	MAC	STC
1	Is	Is
2	Be	Be
3	Have	Have
4	Do	Do
5	Enter	Will
6	Will	Are
7	Shall	Shall
8	Come	Come
9	Hath	Enter
10	Was	Let
11	Would	Would
12	Make	Am

Table 2: Comparative list of order of verb frequency

Not quite knowing what to do with the verb list, I decided to look at a frequency list of **nouns**:

MAC (17,830 tokens)			STC (101,352 tokens)	
time	45	0.25%	146	0.14%
king	36	0.20%	197	0.19%
lord	35	0.19%	448	0.44%
man	30	0.16%	189	0.18%
sir	29	0.16%	259	0.25%
sleep	26	0.14%	54	0.05%
blood	26	0.14%	68	0.06%
son	26	0.14%	73	0.07%
Thane	25	0.14%	*not included in the 400 most frequent occurrences	
Heart	22	0.12%	122	0.12%

Table 3: Frequency list of nouns

The order of frequency in which these words appear is the following:

	MAC	STC
1	Time	Lord
2	King	Sir
3	Lord	Love
4	Man	King
5	Sir	Man
6	Sleep	Time
7	Blood	Father
8	Son	Heart
9	Thane	Night
10	Heart	Heaven

Table 4: Comparative list of order of noun frequency

Checking the most frequent words in MAC against those in the STC, I picked out four. Because **sleep, blood, son, thane** are not the most frequent words in the STC, they could be telling us something particular about *Macbeth* (“lord” and “thane” would seem to be selecting from the same lexical set and would still be less than the normal percentage of

usage in the corpus). Would one of the keys to interpretation lie on these word and their patterning?⁴

B Selecting from the list

The next step was to produce a **concordancing** of these words, or “a collection of the occurrences of a word-form, each in its own textual environment” (Sinclair, 1991:32) to check whether the patterns they helped produce would be worth pursuing.

• **Thane**

The concordancing showed that **thane** collocates mostly with the titles of the characters referred to in the play (Thane of Cawdor; Thane of Fife; Thane of Glamis; etc.). In 8 cases, they are part of noun phrases of honorific titles, in 6 cases they are placed in a positive semantic prosody and only in 3 cases they have negative connotations. There were 8 cases which I considered general and non-specific. If we add the 8 instances of honorific titles and the 6 cases where **thane** has positive connotations, we can see how we are made to look through Macbeth’s perspective, the honorific title thus seen in a positive light as the object of his desire.

➤ **Son**

A concordancing of this word shows that in 15 cases, it occurs in the speech directions for MacDuff’s son. As they are not in the spoken text and if we remove the other stage directions, we are left with only 9 occurrences to check, 5 of which have the meaning of progeny. Not a very exciting finding in itself, but relevant when put together with the treatment of the word **woman** (see below). It is interesting to notice that many of the

⁴ I am grateful to Mike Hoey who read an early draft of this paper and has added that “one might also draw conclusions about the absence of words from the MAC list – perhaps more interesting than what is present. It seems to me rather telling that Macbeth does not have frequent use of love or heart – after all other tragedies might be expected to feature such words – Hamlet has its Ophelia relationship. Othello centres around the Desdemona relationship, Coriolanus focuses on the hero’s relationship with his mother amongst others. But Macbeth is literally heartless. Likewise regicide and patricide are near allied, especially in Macbeth where Duncan is portrayed

characters have sons in the play (Duncan has two, Banquo, MacDuff, Siward have one each). There has even been speculation in traditional literary criticism about possible children Macbeth may have had, but as this speculation is not confirmed by the language of the play, I decided to disregard it. The child-theme has already been given much emphasis, especially by Wilson Knight (1968:163-164). Therefore, I decided to concentrate on the words **blood** and **sleep** to check how far this would lead me.

C Narrowing down: a concordancing of two items

- **blood**

Macbeth is definitely the bloodiest of the tragedies. The following table shows the frequency of this word as compared to *Hamlet* (HAM), *Othello* (OTH) and *King Lear* (LRQ), the Helsinki Corpus of English of compiled documents during Shakespeare's lifetime (HCE2) and my Holinshed corpus (HOL). If compared to the whole tragedy corpus, **blood** is 2.08 times more frequent in *Macbeth*:

MAC	HAM	OTH	LRQ	HCE2	HOL
0.14%	0.05%	0.04%	0.04%	0.01%	0.01%

Table 5: Occurrence of "blood"

In terms of colligation, there is little to say about the patterning of the occurrences. Shakespeare seems to prefer **blood** in rhematic positions, generally preceded by a preposition (*with* *, *in* *). There are very few occurrences in thematic position. However, semantic preference is more telling. Blood is used in relation to a baboon, a sow, and a bat to add the exotic aspect of witchcraft but this is not what distinguishes its use in this play from the other tragedies. Whereas in the other plays, the meanings are mostly metaphorical, a concordancing of **blood** in *Macbeth* indicates that it deviates from

as fatherly. And of course heaven is hardly appropriate in a play where hell on earth is portrayed!" (personal communication).

Shakespeare's use of the word as there are 15 literal uses and 11 metaphorical ones, thus adding to the visual impact and the violence of the play, as the following table indicates:

	METAPHORICAL	LITERAL	TOTAL
MAC	10	15	25
HOL	16	8	24
HCE	16	6	22*
HML	17	2	19
LRQ	12	4	16
OTH	10	2	12

*apparently most of these instances come from religious documents.

TABLE 6: Meaning of “blood”

With regard to collocates, in MAC, most of them occur before the core. There is a semantic preference for profusion, especially in N-3 and N-2 positions (all, much, fountain, pour). For instance, we can contrast “gouts of blood” (MAC) to “that drop of blood” (HAM). In HOL, 5 instances out of the 8 literal ones have the idea of profusion (“cakes of blood”, “beraied with blood”, “great effusion of blood”, “blood ran about”) but it still does not compare to preference of the literal use of the word in MAC. In addition, the association of “blood” and “gold” in “His silver skin lac'd with his golden blood”(II.iii.119) and “gild the faces of the grooms” in MAC produce a semantic preference for richness, for something treasured, which does not occur in HCE, LEAR, HAM, OTH or HOL.

- **sleep**

The frequency table of occurrences of this word in different plays is also interesting:

MAC	HAM	OTH	LRQ	HCE	HOL
0.14%	0.03%	0.01%	0.02%	0.0%	0.004%

Table 7: Frequency of “sleep”

Again **sleep** is 2.8 times more frequent in *Macbeth* (0.14%) than it is in the tragedy corpus (0.05%). In addition, it is more common as a noun in *Macbeth* than in the other plays (MAC: 16 nouns/9 verbs; HAM: 4 nouns/ 8 verbs; LRQ: 4 nouns/4 verbs; OTH: 3 nouns/3 verbs)⁵. How relevant is this finding? This evidence may give a different dimension to the sleepwalking scene, to the dreamworld and ambiguity in *Macbeth*. “Sleep” is a concept, not a process. It is “the season of all natures”. There are two different kinds of sleep in the language of the play: the ones with negative semantic preference, like “Curtained sleep”, “thralls of sleep”, “swinish sleep”, “equivocates him in a sleep”, the one which makes one guard cry “Murder” in his dream, and the ones set in a positive semantic preference, such as “innocent sleep”, “sleep in spite of thunder”, “the benefit of sleep”, the one which makes the other guard “laugh in his sleep”. Here again “foul is fair and fair is foul” is reinstalled. Sleep can be either good or bad and adds to the uncertain atmosphere of the play and supports the theme of equivocation. This evidence goes against Wilson Knight’s statement that “ sleep [is]... the gentle nurse of life ...‘sleep’ is twined with ‘feasting’. Both are creative, restorative, forces of nature. So Macbeth and his Queen are reft of both during the play’s action” (1968:148). The sleepwalking scene then is a further evidence of the breakdown of a strong bond between the two main protagonists. While one cannot sleep anymore, the other one falls into a continuous slumber. The couple lose common ground, once shared through letters and dialogues, and inhabit different worlds.

⁵ There was only one instance in HCE and 8 in HOL (3 nouns and 5 verbs).

Producing this list also worked as a check on other traditional statements about the play. For instance, Wilson Knight (1968: 139) claims that “throughout the main action of *Macbeth* we are confronted by fear. The word occurs ubiquitously. Fear is at the heart of this play”. A frequency list shows that **fear** occurs 35 times (17 as verbs; 16 as nouns and 2 as objects of prepositions). If we look at their collocations, we notice that there is a semantic preference for actually denying **fear** rather than reinforcing it in 10 verb cases and in 7 of the nouns. So, in 50% of the cases, **fear** is denied or questioned (e.g. “Fear not”, “What need I fear?” “Hang those that talk of fear”, “nor shake with fear”). He also holds that **honour** “occurs throughout, strongly emphasized”. **Honour** actually occurs 11 times.

D Using a top-down approach

Instead of starting from a frequency list of all word-forms in the play, I decided to look at Lady Macbeth’s role as a woman in the play. What drew my attention was that in reading the play, I noticed that she always spurred Macbeth by questioning his masculinity. There are 34 instances where the word **man** is used. A concordancing of this lexical item shows that the most frequent collocates are **a** (14) and **of** (13), which indicates a tendency to the generic reference.. In 10 instances, the verb **be** collocates with **man** and in 4, comparatives are used (“like”, “as”, “more”). An examination of the semantic area of collocates will indicate that in 12 instances, **man** is related to *power* and in 6 to *temporality*. In the play, the use of the word **man** points at the Renaissance focus on man as the centre of reference, as a paradigm, although subject to temporality. In order to spur her husband into action, Lady Macbeth attacks his manhood. The word she uses most is *great* (Louw 1991: 172) and it generally acquires an ironic overtone. This finding contradicts Wilson Knight’s (1968:141) suggestion that “Lady Macbeth wins largely by appealing to Macbeth’s ‘valour’” .

On the other hand, **woman** is taken for her capacity to reproduce. And this is typical of the play and of Holinshed’s representation of women. There are 14 occurrences of **woman** in MAC and the following list shows the most frequent collocations:

MAC		HOL	
Right: a = 14 Of = 11	Left: of = 13 Born = 11	Right: a = 10 of = 8	Left: any = 5 Borne = 3 poore = 3

Table 8: Collocates of “woman”.

What is interesting here is the adjective *poore* in HOL. Women are presented as destitute, lacking, in contrast to man, whose collocates to the left are **everie** (16), **noble** (9), **great** (8), **anie** (7). Here, the adjectives *noble* and *great* stand out in opposition to *poore*, as *any* contrasts with *everie*. According to Quirk et alii (1980:365), *any* may express indefinite amount, whereas *every* is “unambivalently singular”. Man is singled out as unique. In addition, there are two instances in *Macbeth* of **woman** with parts of body: **woman’s breasts**; **woman’s ears**. The following table illustrates the semantic prosody of **woman** in the other three tragedies:

	Negative	Positive	Neutral	Total
HAM	5	0	3	8
OTH	5	7	4	16
LR(Folio)*	3	0	5	8

*Instead of the Quarto Edition, as in the other tables, I used the Folio edition as it had one more instance.

Table 9: Semantic prosody of “woman”

Would the semantic prosody of **woman** in *Macbeth* indicate that through Lady Macbeth’s deeds Shakespeare would be condemning women who refused their role as procreators? Would this rejection be a hideous crime? Would her refusal turn Lady Macbeth into the fourth witch? This seems to be what Freud suggests when he says that: “I believe Lady Macbeth’s *illness*, the transformation of her callousness, by which she is convinced of her *impotence* against the decrees of nature, and at the same time reminded that it is through her own fault if her crime has been robbed of the better part of its fruits”(p. 135) (my

italics). This statement seems to be contradictory. If she's impotent against the decrees of nature, if, as the corpus indicates, she is socially represented as a poor element of procreation, why won't she bear children? I believe that, on the contrary, she gets what she sets out to get. What she cannot measure are the consequences of her act. Her short-sightedness does not allow her to anticipate the effect of the murder on her husband and on herself. She does not realize that it has come between herself and her partner, who begins to avoid her. They move into different worlds and this is what makes her break. Her suicide is the result of the loneliness she experiences in her state of isolation. Instead of a deranged sleepwalker, Lady Macbeth is linguistically represented as an early example of the "madwoman in the attic"⁶.

If we set this finding historically, could we see this preoccupation in terms of the instability created by the fact that Queen Elizabeth had no children? Freud also notes this when he comments that the plot "offered remarkable analogies to the actual situation. The 'virginal' Elizabeth, of whom it was rumoured that she had never been capable of child-bearing and who had once described herself as 'a barren stock', in an anguished outcry at the news of James's birth, was obliged by this very childlessness of hers to make the Scottish king her successor... The accession of James I was like a demonstration of the curse of unfruitfulness and the blessings of continuous generation" (p. 133).

IV Conclusion

In this paper I have tried to verify how the rapidly developing area of corpus linguistics can help the analysis of literary texts. My strategy was to run a canonical text with a long history of critical literature through a software which I hoped would organize the language of the text so that a fresh look could be offered into the way the language choices are patterned. The main objective was to check whether some of the traditional readings could find linguistic evidence to support their arguments and whether new readings could be produced.

⁶ Gilbert, S & Gubar, S. *The Madwoman in the Attic: The woman writer and the nineteenth-century*

To conclude, I am aware that this analysis is tentative and far from being comprehensive or thorough. It should be regarded as a preliminary exercise which has to be contextualized and historicized. One way of doing it may be to build a corpus of the history of the text's reception and check the data against the language of the text. There is also the problem of setting criteria to interpret the data. This is an area which, to my knowledge, has not been dealt with thoroughly before. My provisional conclusion is that among its benefits, corpus linguistics allows scholars to ask new questions which may not correspond to the traditional notions of reading (and interpreting) texts. It offers a high degree of verifiability. It provides high speed access and analysis of large amounts of data. One may have access to intertextual levels of analysis, as comparisons between texts are made much easier. Instead of looking at the discourse of one text, it allows the polyphony of discourses, a dialogue between texts and it can be an invaluable method for the teaching of literature. So, if we cannot avoid smudging the canvas, if objectivity is hard to sustain in a postmodern world, corpus analysis may at least help us make our own smudges more accountable.

References:

Criticism on *Macbeth*:

- Freud, S. 1968. "Some character-types met with in psycho-analytical work". In *Shakespeare. Macbeth. A Casebook*. J. Wain (ed.), London: Macmillan.
- Knight, G.W. 1968. "The Milk of Concord: Life-themes in *Macbeth*". In *Shakespeare. Macbeth. A Casebook*. J. Wain (ed.), London: Macmillan.
- Knights, L.C. 1933. *How many children had Lady Macbeth? An essay in the theory and practice of Shakespeare Criticism*. Cambridge: The Minority Press.

On Corpus Linguistics, Stylistics and Grammar:

- Barnbrook, G. 1996. *Language and Computers*. Edinburgh: Edinburgh University Press.
- Carter, R. A. 1997. *Investigating Literary Discourse*
- Gibbs, R. W. 1994. *The Poetics of Mind*. Cambridge: CUP.
- Louw, B. 1997. "The role of corpora in critical literary appreciation" in *Teaching and Language Corpora*, ed. A. Wichmann et alii. London: Longman, pp. 240-251.

literary imagination. New Haven: Yale University Press, 1980?

- Louw, B. 1993. "Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies". In *Text and Technology*, ed. M. Baker, G. Francis & E. Tognini-Bonelli. Amsterdam & Philadelphia: John Benjamins. Pp. 157-176.
- Louw, B. 1991. "Classroom concordancing of delexical forms and the case for integrating language and literature" in *Classroom Concordancing*, ed. T. Johns & P. King, *ELR Journal*, 4, 151-178.
- Quirk, R. et alii 1980. *A Grammar of Contemporary English*. London: Longman.
- Sinclair, J. McH. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. McH. 1996. "The search for units of meaning" in *Textus IX*, no. 1, January-June, pp. 75-106
- Sinclair, J. McH. 1997. "Lexical Item". Unpublished monograph.
- Sinclair, J. McH (ed.) 1987. *Looking Up*. London & Glasgow: Collins ELT

Electronic Texts:

- The Helsinki Historical English Corpus (ICAME CD-Rom)
- Shakespeare, William, *The Complete Works*. Oxford Electronic Texts
- Scott, M. & Johns, T. Microconcord. Oxford University Press, 1993

Suggestions for Pedagogical Applications

Here are suggestions for work to be carried out with students before they produce their critical appreciation of the text:

1. Observe the occurrences of fear in *Macbeth*.

- a. Classify these occurrences (nouns, verbs, etc.) and check which group is more frequent.
- b. Observe the environment of *fear* in each instance. Use **d** or **c** to indicate whether fear is denied or confirmed. Use **x** for the ones which are hard to classify.
- c. Compare your classification in **a** and **b** with that of your partner.
- d. Make a generalizing statement about Shakespeare's use of *fear* in *Macbeth*.

2. Running the electronic texts you have of a playwright's production through Microconcord, prepare two frequency lists of verbs:

- a. in a particular play
- b. in the writer's whole production.

Once the lists are ready, check whether any differences stand out and whether you think they are significant

3. Run a list of nouns according to the same procedure above.

4. **Select some of the items** you consider significant and produce a concordancing to check the textual environment of the word selected. Observe whether it has a negative or positive semantic prosody and check the implications of this finding against your interpretation of the work.
5. **Look into the syntactical patterns** (colligation) in which the items you selected above appear. If you find any pattern(s) you consider significant, describe it and point out its relevance.
6. **Make a list of the collocates** of one item and check their position in relation to the core word. The chart below is an example of some of the collocates of "blood" in *Macbeth* Modify the chart below according to the core word you choose and observe whether any pattern(s) stand(s) out. If you see one, describe it and comment on its relevance.

N-4	N-3	N-2	N-1	CORE	N+1	N+2	N+3
too	Much	Charged	with	blood	of	Thine	
	All	Badged	with	blood	,	So	were
the	Sleepy	Grooms	with	blood	.		
The	secret'st	man	of	blood	.		
Those	Clamorous	Harbingers	of	Blood	and	Death	.
and	Dudgeon	Gouts	of	Blood	,		
,	A	sip	of	Blood	,		
the	Smell	of	the	Blood	still	.	
the	Fountain	of	your	Blood	Is	Stopped	,

7. Make similar charts for other lexical items you consider relevant.
8. Choose a statement on the work in question by a literary critic and decide whether your findings support or contradict the statement.