Semantic preference and semantic prosody re-examined*

MONIKA BEDNAREK

Abstract

In this paper I want to re-examine the key corpus-linguistic notion of semantic preference. This is defined here as the collocation of a lexical item with items from a specific (more or less general) semantic subset. The article aims to throw some light on the term semantic preference*, and to examine in more detail some aspects of semantic preference that are frequently neglected in research. It also discusses how semantic preference interacts with syntax and meaning, and what happens when semantic preferences are not 'realized' in context. Finally, it seeks to illuminate the distinction between semantic preference and semantic prosody, and points to future research in this area.*

Keywords: semantic preference; semantic prosody; collocation; connotation; irony; creativity; attitudinal/evaluative meaning; collocational clash.

1. Defining semantic preference

The concept of semantic preference and the related notion of semantic prosody are currently very important topics in corpus linguistics (some recent references are Partington 2004; Whitsitt 2005; O'Halloran 2007). However, these two terms (and what they designate) have occasionally been confused; they have sometimes been employed for the same phenomenon, and at other times for different but closely related phenomena. This paper provides a discussion of some relevant issues, and tries to define the two terms in a more precise way. I shall start with gradually building up a definition of the term *semantic preference* and examining it with respect to a number of aspects such as syntax, meaning and context. I then discuss the difference between semantic preference and prosody, and identify areas of research that are delineated for future study.

In order to shed some light on this issue, I will start this article by avoiding the terms *semantic preference/prosody* (though they will be introduced later on), and will instead focus on one of the phenomena that they describe. Thus, both of these terms have been used to refer to the fact that some lexical items predominantly co-occur with what can be called 'negative' ('bad', 'unpleasant') and 'positive' ('good', 'pleasant') collocates. To avoid using the terms *preference/prosody*, let us call this phenomenon simply NEG/POS collocation at this stage. Compare some examples mentioned in the literature (Bublitz 1998; Sinclair 1990; Louw 1993; Stubbs 1995):

Lexical items investigated	Collocation identified as
completely, utterly, totally	NEG collocation
perfectly	POS collocation
break out, outbreak	NEG collocation
bent on, symptomatic of	NEG collocation
cause	NEG collocation
provide	POS collocation

Typical examples (from the BNC baby, a 4-million word corpus comprising samples of about 1 million each from academic discourse, conversation, news, and fiction; http://www.natcorp.ox.ac.uk/babyinfo.html) are:

(1) Geoffrey Bailey, head of Southern Region Testing Team: **utterly** and **completely** <u>failed to raise standards or curb excessive overtime</u>. [NEG]

(2) I think you've got a **totally** <u>illiterate</u> person, no doubt [NEG]

(3) He was **perfectly** <u>amiable</u> to her on the few occasions when they did meet; sometimes she even felt that he liked her. [POS]

      (see also appendix 1 for further examples)

This phenomenon is related to the habitual collocation of lexical items with linguistic expressions that belong to certain semantic fields. For now, let us call this SEM collocation. Some examples (Stubbs 2001: 65; Partington 2004) are:

Lexical items investigated	SEM collocation identified as
large	Quantities and sizes
sheer	Magnitude, weight, volume Force, strength, energy Persistence Strong emotion
entirely	Absence Change of state (In)dependency

As becomes evident, these lexical items "co-occur typically with other words that belong to a particular semantic set" (Hunston 1995: 137), collocating with "items which share a semantic feature, for example that they are all about, say, sport or suffering" (Sinclair 2004a: 142). Again, some examples demonstrate this:

(4) The house seemed very quiet and she suddenly realised she was **entirely** alone. [absence]

(5) Then he leaned towards her and took both her hands in his and began to study them as if they were **entirely** new creations, the like of which he had never seen before. [change of state]

(6) He coughed, then went on, They won't make you **entirely** independent, but they will ease any worry about the maintenance of the house and estate. [independency]

     (see appendix 2 for more examples)

Thus, POS/NEG collocation and SEM collocation are arguably very similar phenomena: the only difference is that POS/NEG is a more general semantic group than the groups usually listed in SEM collocation studies: "It is a question of how open-ended the list of collocates is: it might be possible to list all words in English for quantities and sizes, but not for 'unpleasant things'" (Stubbs 2001: 66). In fact, sometimes the two phenomena, POS/NEG collocation and SEM collocation, occur together, when lexical items exhibit both a preference for POS/NEG collocates and for collocates from other semantic groups (cf. Partington 2004: 140 on *happen*, *occur*, *set in*, and *take place. Happen* and *set in* have also been discussed by Sinclair in earlier studies – see Partington 2004 for the relevant references).

   The two types of collocation are hence very similar, differing only in degrees of 'generality', and frequently occur together. I would therefore suggest that we treat them as two types of the same *collocational* phenomenon. Since this phenomenon has now been described in some detail, it is useful to introduce a cover term for it, rather than continuing to avoid the usage of any linguistic term as I have done so far. Therefore, I shall from now on use the term *semantic preference* as a cover term for this kind of collocation (including *both* POS/NEG and SEM collocation), though the terms *POS/NEG collocation* and *SEM collocation* will be employed whenever there is a need to distinguish between these two types of semantic preference. As Partington puts it, there is a view that POS/NEG collocation

   is a sub-category, or a special case, of semantic preference, to be reserved for instances where an item shows a preference to co-occur with items that can be described as bad, unfavourable or unpleasant, or as good, favourable or pleasant. This description of the relation between the two phenomena as set and sub-set, probably works as a rule of thumb in most cases (Partington 2004: 149).

To recapitulate, *semantic preference* in this paper refers to a particular type of collocational phenomenon, which is (relatively objectively) observable by looking at corpus evidence. However, the labels that are assigned to the semantic subsets to which collocates belong (e.g. 'force/strength/energy', 'strong emotion' etc) are necessarily the analyst's – it is s/he who decides how to interpret, categorize, and classify the collocates semantically. The *identification* of POS/NEG collocation is particularly problematic: in order to do so, it is necessary for the analyst to determine whether the collocates of a lexical item are positive or negative, which is difficult to establish objectively. Admittedly, some words are explicitly negative and easily recognizable, but others are not. Thus, Stubbs argues that many collocations for *intellectuals* have negative connotations because "collocates such as the following would be interpreted negatively in many circles: activists, contempt, dissident, hippie, ideology, leftist, leftwing, liberal, students, young" (Stubbs 1996: 188).

This quote shows the problem: it is true that these collocates would be evaluated negatively in many circles, but they might also be evaluated positively in others. Simply counting these collocates as negative disregards the preconceptions that readers bring to the text; some readers may in fact regard *leftwing*, *liberal*, *activist* and *student* as more or less positive. Similarly, dissidents, as pointed out by Lee and Solomon are often admirable in the context of countries that are not democracies, but undesirable in the context of the US (Lee and Solomon 1991: 264). Thus the percentage of collocates identified as negative may vary from researcher to researcher: where one sees a POS/NEG collocation the other may not. In some cases the evaluative meaning of lexical items might be dependent on the attitude/ideology of the whole text – but this is sometimes not taken into account when doing large-scale corpus research. In contrast, collocations involving neither POS/NEG collocation nor SEM collocation are much more objectively analyzable, because the results need not be further interpreted (being concerned with word *forms* rather than *categories* of words).

2.   Semantic preference in interaction

Having now established a relatively clear definition of semantic preference, we can move on to examine semantic preference in interaction with speakers, genres, syntax and meaning. The aspects of semantic preference outlined in the following have come to light in large-scale corpus analyses in recent years, even though they are often neglected in discussions of the phenomenon.[1] Firstly, some more general remarks.

Semantic preference is probably shared among speakers of a given speech community (Partington 2004: 152). In this respect, semantic preference has recently been considered in terms of "priming" (Hoey 2003; Partington

2004), that is, "as the word is learnt through encounters with it in speech and writing, it is loaded with the cumulative effects of those encounters such that it is part of our knowledge of the word that it co-occurs with other words" (Hoey 2003). However, there are different degrees, or strengths of semantic preference (Louw 1993: 164; Bublitz 1995: 14; Partington 2004: 144). If lexical items have a strong semantic preference, their occurrence sets up specific expectations regarding the semantics of their collocates – what Bublitz (1995) calls "syntagmatic predictability" (1995: 1). If these expectations are fulfilled, the result is either (intra-sentential) cohesion (Stubbs 1995: 16), or, in the case of intensifiers, an increase in the intensity of (e.g.) negativity involved, as in the utterance ***Arch-plotter** Francis Maude openly took $50 bets that IDS is doomed as leader* (THE SUN).

Furthermore, semantic preference is probably context-, genre- and domain-dependent (cf. also Partington 2004: 153; and Hoey 2003 on priming). For instance, anecdotal evidence (broadcasts on MTV) suggests that *totally* in AE young adult language does not collocate with 'negative' lexical items. There might be a problem, then, in using general language corpora to research semantic preference: for instance, an analysis of concordance lines for *responsibility for* in the Bank of English shows a negative semantic preference (it co-occurs with *recent bombings*, *Sunday's suicide bombing*, *one of the murders*, *some of the explosions*, *the killing*, *these murders*, *the Holocaust*, *the kidnapping*, *the worst act of terrorism*) while an analysis of the same lexical expression in a business English corpus does not (here *responsibility for* collocates with *budgets*, *a major functional area*, *the product*, *a new stream of scheme*, *outcomes*, *decisions*, *aspects*, *the technical aspects*, *scale-up and operation*, *outcomes*) (Walker 2004). Analogous to O'Halloran (2007) we might call this a *register* or *genre preference* – a semantic preference that is dependent on register or genre.

Secondly, semantic preference interacts with syntactic patterning (colligation), word class and meaning: Semantic preference applies to patterns (or colligations) as well as to individual words. For example, the patterns <*it* v-link ADJ *of* n *to*-inf> and <*there* v-link *something* ADJ *about* n> have a semantic preference for evaluative adjectives (Hunston and Francis 2000: 105).[2] And Partington (2004: 145) notes that in the pattern <*the sheer* (noun phrase) *of* (noun phrase)> the semantic preference is for 'magnitude/weight/volume' or 'force/strength/energy', whereas in the pattern <preposition of means/manner (e.g. *through, out of, by, because of, by virtue of*) + *sheer*> the preference is for 'persistence'.

Related lexical items that belong to different word classes may also have distinct semantic preferences: O'Halloran (2007) points out that in newspaper texts the verb *erupt* (in the word forms *erupted/erupts*) has a semantic preference for human phenomena whereas the noun *eruption* (in the word forms *eruption/eruptions*) has a preference for geological phe-

nomena. The problem is that nouns such as *eruption* are usually simply 'run-on' entries of the related verbs in dictionaries. This means the word is "not itself defined, but is printed, along with its word class and possibly an example, usually in a bold typeface, at the end of the entry for the word it is derived from" (Kilgarriff 1997: 153). For example, the *Oxford Advanced Learner's Dictionary* (OALD) only gives the following entry for *eruption*:

> **eruption** noun [C, U]: *a major volcanic eruption; an eruption of violent protest; skin rashes and eruptions*

Since access to knowledge about semantic preferences is considered important for language learners (Partington 1998: 72), and since such preferences may differ across word classes, this is rather problematic. In such cases helpful information for the learner about the preferences of related items belonging to different word classes is missing.

Finally, Bublitz (1995: 12) points out that different meanings have "different collocational patterns", and shows how the claim that verbs such as *happen* (for extended analyses see Bublitz 1995; Partington 2004) have a particular preference "cannot be supported as it stands, but has to be modified taking different meanings of the verb into account" (Bublitz 1995: 14). If a lexical item has a semantic preference with respect to one of its meanings, and if it is then used in a different meaning, it is likely that the semantic preference is not 'activated' by the reader/hearer, and that they do not expect the semantic preference to be 'fulfilled'. Related to this is the fact that semantic preference interacts with the difference between more literal and more metaphorical meanings. Let us illustrate this with a brief analysis of *short-sighted* in the BNC:[3] in its literal sense ('If you are **short-sighted**, you cannot see things properly when they are far away, because there is something wrong with your eyes'; definition according to *Collins COBUILD English Dictionary for Advanced Learners* (*COBUILD*)) lexical collocates (disregarding grammatical words such *as*, *and*, *this*, *as* etc) are:

> eyes, people, window-cleaners, scribbler, extremely, steel-rimmed, flat-footed, dozy, patients, bulldog, infirm, disciple, amos, piggy, relentlessly, prematurely, beetles, blinking, asked, burglar, bulky, mole, incompetent, exceedingly, peering, blinked, bald, betrayed, mansell, disclosed, congregation, bat, vanished, man, lined, scarcely, throwing, neighbour, severely, uncertain, trace, glasses (all occurring at least once and with a log-likelihood value of between 11,6 and 37,8; ordered according to log-likelihood).

Generalizing slightly, 'literal' *short-sighted* hence occurs with:

– intensifying lexis (*relentlessly, exceedingly, extremely, severely, scarcely, short-sighted as a mole/as a bat*);

– lexical items from the semantic subset '(physiological) weakness/age/ infirmity' (*flat-footed, patients, infirm, bald, uncertain, dozy, bulky, incompetent*);
– lexis from the semantic field 'seeing' (*eyes, steel-rimmed, glasses, blinked, blinking, peering*);
– human phenomena from the more 'private' domain (*people, window-cleaners, man, neighbour, congregation,* names such as *Amos*, *Piggy* (from *Lord of the Flies*), *Mansell*);
– animals (*bulldog*, *beetles*).

In its metaphorical sense ('If someone is **short-sighted** about something, or if their ideas are **short-sighted**, they do not make proper or careful judgements about the future'; COBUILD) in contrast lexical collocates are:

uninformed, selfish, cripplingly, politicians, rottenness, dealing, profit, attitude, debunking, widely-held, opposition, non-productive, serpell, cutback, narrow-minded, redolent, humourless, servile, privatizing, uncovering, thwart, empiricism, sinhalese, stupidly, speculators, regrettable, snapping, view, characterize, gallop, trifle, simplistic, capitalists, rave, fashions, stubborn, irresponsible (all occurring at least once and with a log-likelihood value of between 14, 47 and 37, 5, ordered according to log-likelihood).

'Metaphorical' *short-sighted* thus occurs mainly with:

– negative adjectives (*uninformed, selfish, narrow-minded, humourless, simplistic, servile, non-productive, regrettable, stubborn, irresponsible*), adverbs (*cripplingly, stupidly*) and nouns (*rottenness, capitalists*);
– human phenomena from the more 'public' domain (*politicians, opposition, speculators*);
– lexical items from the financial domain (*dealing*, *profit*, *cutback*, *privatizing*)
– nouns denoting 'opinion' (*attitude, view*) or 'fashion' (*rave*, *fashions*)

In other words, literal and metaphorical uses of lexical items may result in different semantic preferences. Purely automatic corpus analysis does not capture such differences in meaning, because it

defines a word as any string of characters separated by blanks or punctuation. [...] [T]hat makes counting easy and has the advantage that everyone knows where they stand and will arrive at the same numbers. The disadvantage, of course, is that it doesn't tell the truth. [...] Any step towards the truth (as linguists strive to define it) tends to be a step away from anything that is computationally straightforward (Kilgarriff 1997: 144).

Lexicographic works based on such automatic analysis mistakenly often (though not always) make no meaning distinctions when listing collocates. For example, *Collins COBUILD English Collocations on CD-ROM* (*COBUILD COLL*) does not distinguish between *guilty* as in 'If you feel **guilty**,

you feel unhappy because you think that you have done something wrong or have failed to do something which you should have done' and *guilty* defined as 'If someone is **guilty** of a crime or offence, they have committed that crime or offence' (COBUILD) and lists collocates for *guilty* indiscriminately:

> *found – pleaded – feel – charges – plead – felt – murder – verdict – court – plea – feeling – because – also – pleading – charge – innocent – yesterday – manslaughter – counts – jury*

It is questionable whether all language learners that use such a resource are advanced enough to be able to deal with this sort of collocate listing.

3.   "For God's sake, don't be ironical" (Brontë 1981: 246) – semantic preference and context

One of the most interesting (and most often addressed) questions is, what happens when a lexical item co-occurs in a text ('intratextually') with a lexical item that is not part of the semantic subset to which its collocates usually ('intertextually') belong? An example would be the occurrence of *totally* in a text with a neutral or positive collocate rather than a negative one. This is related to the phenomenon of collocational "clash" (Louw 1993: 157), "dislocation" (Partington 1995: 34) or "deviance" (Krishnamurthy 1995: 13), i.e. the usage of a collocation that is non-conformant. Depending on the strength of the semantic preference (and perhaps other factors), this may result in:

1) irony or the expression of a 'hidden' negative attitude on the part of **the speaker/writer** (Louw 1993; similarly Bublitz 1995, 2002).[4] For instance, in the following text the collocational clash (*symptomatic of* usually co-occurs with negative collocates) shows that "the speaker believes that the University of Zimbabwe badly needs assistance from Britain" (Louw 1993: 170):

> I mean, it's symptomatic of the University of Zimbabwe which has such a high reputation that there are fifteen links between departments in the university here and equivalent departments in all sorts of institutions, universities, polytechnics in Britain. (Louw 1993: 169)

This effect is by now well-known, and many other examples have been identified. Sinclair (2004a) describes this phenomenon as "semantic reversal" (Sinclair 2004a: 134), elaborating in relation to *borders on*:

> Whatever follows this phrase indicates the limits of normality by specifying a mental state that lies just outside normality. When the adjective is *obsessional*, the feature of abnormality is already present in the meaning of the word, and the co-ordinated choice will be felt to be coherent; in the case of *holy* [as in the example *The ambience borders on the holy*], the required feature has to be added by reversal. And if there were

an instance of "borders on the normal", this would be interpreted as fully ironical, suggesting that the normal is unexpected (Sinclair 2004a: 135).

The phenomenon of evaluative or ironic meaning resulting from a collocational clash is thus not restricted to POS/NEG collocation (for Sinclair, *border on* has a preference for collocates that share the semantic feature of (admittedly negative) 'unusuality' or 'unexpectedness' rather than simply for POS/NEG collocates). Another example is: *The game against Turkey will be difficult enough*, *particularly given the precarious security situation. As it is*, *England have **managed to** derail their pre-match build-up themselves*. (Financial Times). Usually people *manage to* do something that is both difficult and a desirable goal. The proposition *derail their pre-match build-up themselves* does not belong in this semantic subset, and the result of the collocational clash seems to be an expression of ironic disapproval on the part of the writer.

However, it is important to point out that such evaluative meanings are not produced automatically whenever collocational clashes occur, but often lie at the intersection of intra- and intertextuality. Only if lexical items exhibit a very strong preference for e.g. negative/positive collocates can collocational clashes result in pragmatic meanings such as irony or disapproval (and even then this may not always be the case; see below).

2) irony or the expression of evaluative meaning on the part of **someone other than the writer.** This is how we can interpret Hunston's (2002) example:

> In Britain, despite a bad choice of school – 'my parents didn't know the system' – and a careers teacher who tried to **condemn her** to a life of nursery nursing, Elaine got together enough qualifications to go to teacher training college. (Hunston 2002: 48, boldface mine)

where "*a life of nursery nursing* may not seem to be a terrible fate to many people, but the phrase *condemn her to* [meaning 'make something bad happen'] indicates that Elaine would see it in this way. (Hunston 2002: 47). This is an example of how collocational clashes can contribute to the adoption of a perspective that is internal to the text world.

In fiction, this might be quite a frequent phenomenon, and a collocational clash may also result in the evaluation of a character, as in:

> Yesterday, Nazma – whose brother-in-law had turned up unexpectedly – had popped around to borrow a pinch of saffron. She rolled into the sitting room holding up her hands. "Can't stay, she said, can't stay," ready to fend off any **onslaught of** hospitality. But she stayed long enough to drop a few hints about Razia "Do you know? The woman *smokes*!" (Ali 2003: 243, boldface mine, italics in the original)

Here the context as well as the collocational clash involved with *on-slaught of* (which, according to an analysis in the BNC has a preference for NEG collocation[5]) results in a negative evaluation of a participant's character, Nazma. It seems that, from the focalizer's (rather than the writer's) point of view, hospitality is construed by Nazma as an attack rather than a positive happening. But Nazma's attitude is construed by the focalizer, rather than by Nazma herself, on the basis of Nazma's (bodily and linguistic) behaviour (*She rolled into the sitting room holding up her hands. "Can't stay, she said, can't stay,"*), and the negative evaluation might be strengthened by the contrast with *But she stayed long enough to drop a few hints about Razia* (a reference to Nazma's bitching). Similarly, compare the following example from the *Guardian* subcorpus of the *Bank of English*, which is part of a review of Channel 4's *Psychos*:

> I certainly didn't believe in Dr Kelly when he took Dr Nash to see a woman who'd just come in with slashed wrists: 'We treat for you, darling. You're going to love this. Haven't seen one of these in ages. (Janet Brown the unhappy woman), *enjoy*.' Shocking, sure. But real? Surely not.

Again, the effect of using the verb *enjoy* with reference to concepts which are not something you usually enjoy, is some sort of sarcasm, cynicism or irony (Bednarek and Bublitz 2007). The irony here does not derive directly from the writer of this review, but rather from the writer(s) of the TV show *Psychos*, who draw(s) on it to characterize Dr Kelly through his way of speaking (Dr Kelly's unusual or even deviant use of *enjoy*). Clearly, the way in which writers and narrators use collocational clashes to create irony with respect to the construal of characters and their point of view seems to be worthy of further investigation.

3) other creative language usage, such as poetic or humorous language use (Bublitz 1995; Partington 1995; Krishnamurty 1995). As Hoey (2003) points out,

> [c]reativity involves a selective overriding of the primings. Thus if, for example, a word is positively primed for one particular collocation, positively primed also for one particular colligation, and negatively primed for another particular colligation, the creative writer may utilise both the positive primings but override the negative, or s/he may override the collocational priming but conform to the colligational primings.

4) a different, non-ironic, non-evaluative, non-'creative' text meaning (of course, all text meaning is creative in a sense). In the following example *totally* occurs in the context of a positive collocate without expressing irony or the writer's "real attitude" (Louw 1993: 157). Neither is it particularly poetic or 'creative':

As energetic and diverse as the American life it describes, Walt Whitman's *Leaves of Grass* has been loved by generations for its celebration of a brash young nation and one man's exuberant spirit. [...] Though the sights and sounds of these poems are those of the nineteenth century, Whitman's freewheeling, open-ended style and earthy voice are **totally** modern. By turns joyous and haunting, sensual and spiritual, his poetry celebrates with fierce originality nature, the body, and the contradictions of the human soul. As fresh today as it was when written over a century ago, *Leaves of Grass* remains one of the most ambitious and exciting works of American literature ever published. http://www.microsoft.com/reader/promotions/free_shop.asp (accessed August 30, 2003, emphasis mine)

Collocational clash can hence result in a number of different effects, by no means simply irony or evaluation (which are most often mentioned in research). Though it "can be exploited by speakers for particular effects" (Partington 2004: 147), it does not always result in creative or evaluative meanings (in cases where the collocational preference is not 'activated' by the genre, text type, context or the meaning of the lexical item). As Halliday argues in connection with a probabilistic grammar, "a speaker, or writer, may always 'choose' to subvert the probabilities of the system" (Halliday 2005: 92).

4.   Semantic preference and semantic prosody

I have argued above that POS/NEG collocation is a type of semantic preference and that semantic preference simply involves collocation with semantic subsets of different degrees of generality. It is nevertheless necessary to return specifically to POS/NEG collocation, in order to discuss a further complex and controversial issue.

It is sometimes suggested by researchers that the result of POS/NEG collocation is that the given expression is "imbued by its collocates" (Louw 1993: 157) with "a consistent aura of meaning" (Louw 1993: 157); in other words, the co-occurrence of an item with emotive collocates has been seen as resulting in a "colouring" of the item itself (Louw 1993: 159).[6] In the strongest case of such coloring, it is proposed that items with a preference for NEG collocation can acquire negative connotations themselves, i.e. evaluative meaning (Stubbs 1995; Bublitz 1995; Hunston 1995; Hunston 2002: 142; Partington 1998: 75, 2004).[7] It is then claimed that "the 'neutral' words themselves carry 'implied meanings' and 'connotations'" (Hunston 1995: 137). Stubbs (1995) proposes that this is a diachronic process (similarly Channell 2000 for *roam the streets*):

CAUSE is near the stage where the word itself, out of context, has negative connotations. [...] The selection restrictions on CAUSE are not (yet?) categorical: it is not (yet?) ungrammatical to collocate CAUSE with explicitly positive words. But it is easy to see how an increase in frequency of use can tip the balance and change the system. (Stubbs 1995: 16).

The problem lies with the evidence for this claim. If connotations associated with POS/NEG collocation are said to lie *outside* native speaker intuition (Louw 1993; Stubbs 1995; Channell 2000; but see Whitsitt 2005: 293–296), on what grounds can we confirm this hypothesis? Channell (2000) seems to suggest that corpus data can help to trace the semantic change: as soon as an expression is only used in "negative contexts", the expression can be said to have "settle[d] in to having only a negative meaning" (Channell 2000: 53). Stubbs (1995) speaks of categorical selection restrictions and the stage at which it becomes "ungrammatical" to collocate an expression with non-negative lexical items as providing evidence for such connotative meaning. However, POS/NEG collocation may be an indication but not the only grounds on which such a hypothesis can be made. As a counter-example, take the verb *tackle*. According to COBUILD this predominantly collocates with 'negative' items such as a 'difficult problem or task', but it certainly does not have negative connotations (if anything, it has positive connotations, in that it is good to tackle problems). Even if such verbs occur only in negative contexts and even if their collocation with non-negative expressions is ungrammatical, they do not acquire negative connotations. Whitsitt (2005) also notes that "verbs like *alleviate, heal, relieve, soothe*, etc [. . .] all habitually appear in the company of clearly unpleasant words, yet it seems clear that a word like *alleviate*, to take one example, certainly does not come to have an unpleasant meaning because of that company" (Whitsitt 2005: 296–297). Partington (2004) hides a similar observation in his footnotes:

> Simply being primed to appear in the environments of collocates of a certain evaluative sense, good or bad, is not a sufficient condition for an item to acquire the same sense. If the relationship between the item and its collocates is one of opposition or detraction, then the combination does not acquire the evaluative sense of the collocates. For example, relief [sic] and ease [sic], which are inherently favourably evaluative, collocate very frequently with unfavourable words like *debt*, *pain*, *poverty*, *suffering* and so on. The overall meaning of "the relief of suffering", "easing the pain" remains, of course, highly favourable.

In my view, this should be stated very strongly: the co-occurrence of an item with negative collocates (i.e. NEG collocation) does not automatically indicate the presence of a negative connotation. Another example of this is *brook* which has a negative semantic preference but a favourable prosody (Partington 2004: 151), expressing a justified "intolerance of intrusive behaviour by another" (Sinclair 2004b: 37).

It is thus extremely important for corpus linguists to distinguish between collocational patterning (semantic preference) – whether this relates to positive/negative lexical items or to items from more specific semantic subsets – and the connotations of a lexical item (including those of 'individual' words and extended units of meaning). Sinclair suggests the term *semantic prosody* for the latter:[8]

> The semantic prosody of an item is the reason why it is chosen, over and above the seman-
> tic preferences that also characterize it. It is not subject to any conventions of linguistic
> realization, and so is subject to enormous variation, making it difficult for a human or a
> computer to find it reliably. It is a subtle element of attitudinal, often pragmatic meaning
> and there is often no word in the language that can be used as a descriptive label for it
> (Sinclair 2004a: 144–145).

For example, employing the verb *budge*, "the user wishes to express or report
frustration (or a similar emotion) at the refusal or inability of some obstacle
to move, despite pressure being applied" (Sinclair 2004a: 145). *All* of this is
its semantic prosody (not simply the expression of emotion, but the expres-
sion of *a particular emotion at the refusal of some obstacle to move despite
pressure being applied*). The term *semantic prosody* in this sense refers to the
complex attitudinal and/or functional meaning of lexical items: "It [semantic
prosody] expresses something close to the 'function' of the item – it shows
how the rest of the item is to be interpreted functionally" (Sinclair 2004b: 34).
Semantic prosody refers to attitudinal meanings of all kinds, not just positive
and negative evaluative meanings (as becomes apparent by Sinclair's vari-
ous examples of semantic prosody, e.g. a prosody of 'informal invitation'),
and is, despite its name, "on the pragmatic side of the semantics/pragmatics
continuum" (Sinclair 2004b: 34).

The crucial problem for linguistics is how semantic prosody can be de-
termined: as Sinclair states, it is extremely difficult to observe and describe.
And sometimes "the prosody is, paradoxically, not necessarily realized at all"
(Sinclair 2004a: 145). An argument which is very frequently made in corpus
linguistics (concerning positive/negative semantic prosody) goes like this:
'lexical item x occurs with negative items (i.e. it has a negative semantic pref-
erence); ergo, it has evaluative meaning (i.e. a negative semantic prosody)'.
The argument seems to be that semantic preferences more or less automati-
cally "build up" or "form" (Partington 2004: 150, 151) prosodies. As I have
suggested above, this argument does not always hold, and the two different
phenomena (collocational and evaluative meaning, i.e. semantic preference
and semantic prosody) should, in a careful linguistic analysis, be described
individually and separately. Whereas the observation of semantic preference
is *relatively* objective and straight-forward (see above), the analysis of se-
mantic prosody is much more subjective and problematic (even though, as
suggested in section 1, semantic preference studies are not unproblematic
either). There are great "limitations on how reliably evaluative meaning can
be identified" (Stubbs 2001: 107):

> the amount of inference necessary to identify [. . .] evidence [of a semantic prosody] is
> often great; moreover in several instances there is no actual piece of language that can be
> quoted, but a more general appeal is made to **experience** (Sinclair 2004a: 146, emphasis
> mine).

Substitute *intuition* for *experience* in this citation, and the problem of semantic prosody for corpus linguistics becomes clearer: the analysis of semantic prosody is in fact partly intuitive since it cannot be 'objectively' derived from corpus data and requires a lot of inference on the part of the analyst. The paradox mentioned by Sinclair (that prosody need not be realized linguistically) thus remains problematic; the only answer he offers is that "if we make a strong hypothesis we may establish a search for it [the prosody] that will have a greater chance of success than if we were less than certain of its crucial role" (Sinclair 2004a: 145) – an answer that is not wholly satisfying.

Whereas these issues concern the reliable (and objective) identification of semantic prosody in corpus linguistics, a further problem is that the complex nature of semantic prosody is often disregarded. Arguably, it is of great importance for future research to examine semantic prosody in interaction in more detail – in the way this was outlined for semantic preference above: what is clear so far is that there is "no reason to believe that the prosody cannot be neutralized or reduced substantially in impact by other choices in the vicinity" (Sinclair 2004b: 44), i.e. that prosody, like preference is context-dependent. Additionally, "it is highly likely that the quality and strength of the prosody of a good many items will differ from genre to genre or from domain to domain" (Partington 2004: 153). Thus, O'Halloran (2007) has shown that "some prosodies have probabilistic relationships to register", i.e. vary according to register. For example, *erupted* has a positive prosody in sports reporting but a negative prosody in hard news reporting. He calls this kind of prosody *register prosody*, and also notes that there might be a prosody that is genre-specific (*genre prosody*), e.g. relevant to genres such as news discourse or romantic fiction in general. Clearly, semantic prosody also depends on meaning: returning to the analysis of *short-sighted* above, it seems obvious that 'literal' *short-sighted* can be used more or less neutrally with lexis from the semantic field of 'seeing' or to describe people. Even though it does seem to carry associations that short-sightedness is a physiological disadvantage, it can be used neutrally, and certainly does not necessarily imply that the person who is *short-sighted* is evaluated negatively. In contrast, 'metaphorical' *short-sighted* evaluates the person who is *short-sighted* or who acts in such a way negatively, and indicates disapproval of views and opinions or of human participants in the public domain.

To sum up, it appears to me crucial to uphold the distinction between semantic preference and semantic prosody as proposed by Sinclair (2004a, b).[9] If we want to keep the terms *semantic prosody* and *semantic preference*, we should reserve the term *semantic preference* for collocations of lexical items with (more or less specific) semantic subsets, and use the term *semantic prosody* for connotations of all kinds. *Semantic prosody* then refers to POS/NEG connotation as well as more complex attitudinal connotations, affecting both single words and larger units of meaning such as phrases, i.e.

concerns both 'traditional' connotation (said to relate to single words) and 'prosodic' connotation (connotation that is "distributed prosodically across a textual sequence" (Stubbs 2001: 202)).

Even though most discussions of semantic prosody (apart from those by Sinclair) have predominantly focused on positive and negative attitudinal meanings, it should be kept in mind that there are many semantic prosodies that do not relate to '(un)pleasantness', and that evaluation is much more multi-faceted (Bednarek 2006). But if one would want to identify in particular POS/NEG collocation, one could use the term *positive/negative semantic preference*; and if one would want to identify in particular POS/NEG connotation, one could talk about *positive/negative semantic prosodies*. It would then become possible to argue that verbs such as *soothe, ease* etc have a negative semantic preference but probably a positive semantic prosody.

5.   Conclusion

In this paper I have argued firstly that the collocation of a lexical item with positive/negative lexical items should be regarded as a special kind of semantic preference, and that semantic preference involves semantic sub-groups of different levels of generality.

Secondly, I have pointed out that semantic preference interacts with syntactic patterning, word class, meaning, genre, domain and context, and that collocational clash can have a number of different discourse effects (by no means limited to evaluative meaning).

Thirdly, I have proposed that it is crucial to uphold the difference between semantic preference and semantic prosody, and furthermore, that both phenomena involve more than just positivity and negativity.

Finally, it would seem to me that a number of interesting questions remain yet to be addressed in more detail concerning both semantic preference and semantic prosody:

– Are there additional semantic groups that form collocational subsets apart from those identified so far, and which of these are evaluative/attitudinal? For instance, are there semantic preferences for evaluative dimensions such as '(un)importance' or '(un)expectedness' (Bednarek 2006)?
– Does it affect semantic preference if the collocate precedes or follows the node? (For example, does this affect the strength of the semantic preference?)
– Is semantic preference really a question of semantics (cf. the discussion in Whitsitt 2005: 198), i.e. is it "part of the lexical meaning of a word" (Bublitz 1995: 14)? And what about semantic prosody (Stubbs 2001: 65–66; Partington 2004: 132): is it one of the "inherent components" (Sinclair 2004b: 35) of lexical items?

– Which words and which word classes exhibit the feature of semantic preference and semantic prosody? What factors decide this?
– Which semantic preferences and prosodies are context-, genre- and domain-dependent, and which are not? Should dictionary entries be modified accordingly? For example, the different semantic preferences of nouns and verbs should be listed in the dictionary; collocational preferences for different meanings of lexical items needs to be considered, and information about *genre* or *register preferences* and *prosodies* should be included.

To conclude, it must be stated that both semantic preference and semantic prosody are by no means fully explored, and that corpus linguistics should continue to tackle the two concepts in more detail, taking into account their complex nature more thoroughly than this has so far been the case.

Appendix 1: Examples for POS/NEG collocation

– The askaris open fire and it drops. Now the Masai are gathering, arguing and shouting. A fight **breaks out** and one man is struck on the head by a stick. [NEG]
– Rising unemployment turned a lingering and growing malaise into a crisis. As in the United States, so in Britain it was an **outbreak** of violence that dramatically brought the plight of the inner cities to public notice. [NEG]
– Indeed, towns were as anomalous as the independent lordships the count of Champagne was **bent on** integrating into his castellan network. [NEG]
– An extreme case, perhaps, but Mr Cawley believes it is nonetheless **symptomatic of** the financial malaise affecting all authorities. [NEG]
– Well Bob will do me a swap if it's gonna come to a crisis. I think I'll ring Madge up. Cos it will **cause** a crisis Oh yeah. with Jim. [NEG]
– Air show: The famous Red Arrows are to **provide** a grand finish to the Southport Air Show; organisers say [POS]

Appendix 2: Examples for SEM collocation

– The raider, who masked his face with a tartan scarf, forced Mr Bridgeman to open his safe and escaped with a **large** amount of money. [quantity]
– Further, **large**-scale studies are needed to confirm the encouraging results obtained so far. [size]
– The includers are not always merely disinterested advocates of a philosophy, but people who have become so through experience, through the **sheer** amount of time they spend with members of the excluded social group. [magnitude, weight, volume]

– Overwhelmed by the **sheer** size and <u>strength</u> of the surging crowd, Clasper's supporters suddenly got the message and broke rank, dispersing into the milling throng. [force, strength, energy]
– In the pictures of Belsen, the **sheer** <u>anguish</u> on the faces of the Tommies contrasts starkly with the indifference of the SS recorded in their cheery souvenir albums. [strong emotion]
– But Ian keeps us all going by his **sheer** <u>determination.</u> [persistence]

Bionote

Monika Bednarek is a Chancellor's Post-doctoral Research Fellow at the University of Technology, Sydney. She is the author of *Evaluation in Media Discourse* (2006, Continuum) and *Emotion Talk across Corpora* (2008, Palgrave Macmillan) and is currently co-editing a volume on multi-modality (forthcoming, Continuum).

Notes

* I would like to thank two anonymous reviewers for their constructive criticism of an earlier version of this paper, and am also grateful to Kieran O'Halloran for his helpful comments, and for bringing the word *short-sighted* to my attention.
1. In what follows I shall use the term *semantic preference* even when referring to findings of researchers who use other terms (*semantic prosody/association* etc), because otherwise the discussion would become too confusing.
2. It is a matter of discussion whether the feature [evaluative] is a semantic or pragmatic feature of such adjectives. Perhaps there might be a reason for distinguishing between *semantic* and *pragmatic preference*.
3. For this analysis I used BNCweb (http://escorp.unizh.ch/), an online concordancer for the BNC. Excluding identical occurrences there were 55 instances of 'literal' and 71 instances of 'metaphorical' *short-sighted*. For the collocation analysis at first 'literal' and then 'metaphorical' instances were manually deleted, and collocations for each sense were calculated by BNCweb (window span +/- 3, using log-likelihood, BNC as basis, frequency in BNC at least 5). One of the anonymous reviewers has suggested that a corpus analysis of *short-sighted* has already been undertaken previously, but unfortunately I was not able to locate the source of this previous analysis.
4. Irony may also be the result of other kinds of priming. For example, *presumably* is primed to occur in contexts in which speakers/writers are not certain about their conclusions concerning a given state of affairs. When *presumably* is used with reference to a state of affairs that is obviously true and whose truth cannot be doubted, the result is irony:

> In the context of the war on terror we are constantly being told that we must sacrifice our civil rights and liberties for the greater good. The latest voice to join this defeatist chorus is that of Dame Eliza Manningham-Buller, head of the counter-espionage agency MI5. In a speech that has, significantly, been given prominence on the MI5 website, she acknowledged that human rights had been hard fought for. But, she went on, "the world has changed and there needs to be a debate on whether some erosion of what we all value may be necessary to improve the chances of our citizens not being blown apart as they go about their daily lives". This, **presumably**, is a reference to the

risk from terrorists, as opposed to armed police. (*The Guardian Weekly*, Sept 16–22 2005, emphasis mine; the ironic reference is to the shooting of a suspected terrorist by British police who turned out to be completely uninvolved in terrorism.)

5. I examined the first 30 concordance lines of 87 occurrences of *onslaught of* in the BNC. Of these, the majority (24) occurred clearly with negative right-hand collocates (e.g. *alcoholism*, *propaganda*, *stress and tension*, *behavioural difficulties*, *chaos*, *cold weather*, *criticism*, *death imagery*, *dotty hangers-on*) or in an otherwise negative context (an onslaught is something that one must *defend* oneself *against*, *protect from*, *withstand*, *tackle*, *survive*, *be equipped to meet* etc).

6. For a very recent critique of Louw's concept of semantic prosody and his notion of 'coloring' and 'imbuing' see Whitsitt (2005). While a critique of Whitsitt is beyond the scope of this paper, two points need to be made. Firstly, there is a disagreement between the arguments made in this paper and those in Whitsitt, in that I argue for an overlap between semantic prosody and connotation (see especially section 4). Secondly, it must be noted that he interprets Louw as defining semantic prosody essentially as semantic transfer: "what is central to Louw's [. . .] definition [. . .] can be put as follows: if several different words all sharing the same semantic trait are frequently used with another word, meaning will be passed, over time, from that group of words to the other word" (Whitsitt 2005: 284), and that his critique thus focuses mainly on this aspect (disregarding largely Sinclair's concept of semantic prosody). Whether this captures the essence of Louw's notion of semantic prosody is questionable; in any case, the notion of semantic transfer will not be discussed here.

7. The term *connotation* itself is problematic because it is used differently in linguistics, and sometimes not explicitly defined (cf. also Partington 2004: 154). Thus, it is employed in the sense of *intension* in opposition to *denotation/extension* especially by philosophers. Hence John Stuart Mill, who seems to have coined the term, says that the word *white* denotes all white things and connotes the semantic feature of whiteness (Lyons 1977: 175f). For Leech (1974), connotative meaning consists of the non-criterial properties of referents. Nowadays, it is often used as a cover term for stylistic, regional or emotional meaning components (cf. e.g. Bußmann 1996: 96). Space does not permit further discussion of the ways in which *connotation* can and has been defined. Suffice it to say that in this paper *connotation* refers to all attitudinal (but **not** simply pos/neg) components of a given unit of meaning (word, phrase etc), which may or may not relate to the properties associated with its referent, and which may or may not be accessible to native speaker intuition.

8. Both semantic preference and semantic prosody are part of Sinclair's (2004a, b) model of the lexical item or unit of meaning. This extended linguistic unit is made up of the following optional and obligatory components: semantic prosody (obligatory) and preference (optional), colligation (typical co-occurrence of grammatical phenomena; optional), collocation (typical co-occurrence of words; optional), and the invariable core (obligatory). The semantic prosody is the functional component, and that which must be chosen first by speakers: they select a semantic prosody which is applied to a specific semantic area (semantic preference); this preference in turn controls colligational and collocational patterns of the core of the lexical item (Sinclair 2004b: 34). Hence, the phrasing that is used by speakers "must be seen as the result of a single choice" [of a particular prosody] and "all subsequent choices within the lexical item relate back to the prosody" (Sinclair 2004b: 34). For example, the collocation *true feelings* is the core of a compound lexical item which has the following inherent components:
   – a semantic prosody of reluctance/inability
   – a semantic preference of expression (and a strong colligation of a verb with the semantic preference)
   – a colligating possessive adjective
   – the core
     (Sinclair 2004b: 35–36)

Collocation, colligation, semantic preference, and semantic prosody "are related to each other in increasing abstraction" (Sinclair 2004a: 142; Stubbs 2001: 88). With the examples that Sinclair (2004a, b) gives, the core of the lexical item is lexical (*naked eye*, *true feelings*, *brook*, *budge*), and it is this lexical core that appears to have a semantic preference. A question that remains to be discussed is how this notion of an extended lexical item can be combined with pattern grammar (Hunston and Francis 2000): what is the core in phrases such as <*it* v-link ADJ *of* n *to*-inf> or <*there* v-link *something* ADJ *about* n> where the core seems grammatical rather than lexical? As Sinclair proposes, it is "likely that this lexical item [the model outlined above] is only one of several. One possibility is a type of item based on a grammatical core rather than a lexical one" (Sinclair 2004b: 39; see also Sinclair and Renouf 1991 on collocational frameworks). In such cases (where a colligation carries a semantic preference), it is not simply prosody but also preference that extends across stretches of text, and the claim that semantic preference is "a narrower phenomenon [...] than prosody which can affect wider stretches of text" (Partington 2004: 151) does not hold.

9. It is still important today to emphasize Sinclair's (2004a, b) distinction between preference and prosody (originally made in 1996, 1998), since the central problem of linguistic research on these phenomena seems to lie in the fact that the term *semantic prosody* has been used to refer both to 1) POS/NEG collocation, 2) POS/NEG connotation (usually *on the basis of* an observed POS/NEG collocation) and 3) connotations of all kinds. Similarly, *semantic preference* has been used to refer to 1) POS/NEG collocation, 2) POS/NEG connotation, and 3) SEM collocation (cf. Louw 1993; Stubbs 1995; Bublitz 1995, 1998; Hoey 2000; Hunston 2002; Partington 2004 on semantic prosody; Channell 2000 on *evaluative polarity*; and Hoey 2003 on *semantic association*. Whitsitt 2005 is the most recent overview of some of the different ways *semantic prosody* is defined in linguistics.) As said, *semantic prosody* is very frequently used to refer to lexical items that have a preference for POS/NEG collocates, and are simultaneously (automatically) said to express evaluative meaning. Strictly speaking, the difference between collocation and connotation is not upheld in such cases, and sometimes results in the assumption of relations between the two that appear to be too automatic and unproblematic.

References

Ali, Monica
   2003   *Brick Lane*. New York: Scribner.
Bednarek, Monika
   2006   *Evaluation in Media Discourse. Analysis of a Newspaper Corpus*. London/New York: Continuum.
Bednarek, Monika and Wolfram Bublitz
   2007   *Enjoy!* – The (phraseological) culture of having fun. In Skandera, P. (ed.), *Phraseology and Culture in English*. Berlin/New York: Mouton de Gruyter, 109–135.
Brontë, Charlotte
   1981   *Jane Eyre*. New York: Bantam Books.
Bublitz, Wolfram
   1995   Semantic prosody and cohesive company: somewhat predictable. *General and Theoretical Papers* 347. Duisburg: L.A.U.D. (Linguistic Agency University of Duisburg), 1–23.
   1998   I entirely dot dot dot: copying semantic features in collocations with up-scaling intensifiers. In Schulze, R. (ed.), *Making Meaningful Choices in* English. Tübingen: Narr, 11–32.

2002   Anmerkungen zur emotiven Prosodie des Textes. In Bublitz, W, M. von Roncador and H. Vater (eds.), *Philologie, Typologie und Sprachstruktur. Festschrift für Winfried Boeder zum 65. Geburtstag*. Frankfurt: Peter Lang, 25–40.

Bußmann, Hadumod
1996   *Routledge Dictionary of Language and Linguistics*. London: Routledge.

Channell, Joanna
2000   Corpus-based analysis of evaluative lexis. In Hunston, S. and G. Thompson (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 38–55.

COBUILD = *Collins COBUILD English Dictionary for Advanced Learners*. (2001). CD-ROM, London: HarperCollins. [Lingea Lexicon 2002, version 4.05; includes 5 million word corpus]

COBUILD COLL = *Collins COBUILD English Collocations on CD-ROM: A Comprehensive Database of Common Word Patterns from The Bank of English*. (1995). CD-ROM, London: HarperCollins.

Halliday, M.A.K.
2005   *Computational and Quantitative Studies* (Volume 6 in the *Collected Works of M.A.K. Halliday*, ed. by Jonathan J. Webster). London/New York: Continuum.

Hoey, Michael
2000   A world beyond collocation: new perspectives on vocabulary teaching. In Lewis, M., *Teaching Collocation*. Hove: LTP, 224–245.
2003   Lexical priming and the properties of text, available to download at: http://www.monabaker.com/tsresources/LexicalPrimingandthePropertiesofText.htm

Hunston, Susan
1995   A corpus study of some English verbs of attribution. *Functions of Language* 2, 133–158.
2002   *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Hunston, Susan and Gill Francis
2000   *Pattern Grammar*. Amsterdam/Philadelphia: John Benjamins.

Kilgarriff, Adam
1997   Putting frequencies in the dictionary. *International Journal of Lexicography* 10, 135–155.

Krishnamurthy, Ramesh
1995   The macrocosm and the microcosm: the corpus and the text. In Payne, J. (ed.), *Linguistic Approaches to Literature*. University of Birmingham: ELR, 1–16.

Lee, Martin A. and Norman Solomon
1991   *Unreliable Sources: A Guide to Detecting Bias in the News Media*. New York: Lyle Stuart.

Leech, Geoffrey
1974   *Semantics*. Harmondsworth: Penguin Books Ltd.

Louw, Bill
1993   Irony in the text or insincerity in the writer? In Baker, M., G. Francis and E. Tognini-Bonelli (eds.), *Text and Technology. In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 157–176.

Lyons, John
1977   *Semantics*. Vol. 1,2. Cambridge: Cambridge University Press.

OALD = *Oxford Advanced Learner's Dictionary*
[7]2005   Oxford: Oxford University Press.

O'Halloran, Kieran A.
2007   Critical discourse analysis and the corpus-informed interpretation of metaphor at the register level. *Applied Linguistics* 28, 1–24.

Partington, Alan
   1995   Kicking the habit: the exploitation of collocation in literature and humour. In Payne, J., *Linguistic Approaches to Literature*. University of Birmingham: English Language Research, 25–44.
   1998   *Patterns and Meaning. Using Corpora for English Language Research and Teaching*. Amsterdam/Philadelphia: Benjamins.
   2004   "Utterly content in each other's company". Semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9, 131–156.
Sinclair, John. McH.
   1990   *Collins Cobuild English Grammar*. London: HarperCollins.
   2004a   The lexical item. In Sinclair, J. McH. (ed.), *Trust the Text. Language, Corpus and Discourse*. London/New York: Routledge, 131–148.
   2004b   The search for units of meaning. In Sinclair, J. McH. (ed.), *Trust the Text. Language, Corpus and Discourse*. London/New York: Routledge, 24–48.
Sinclair, John. McH. and Antoinette Renouf
   1991   Collocational frameworks in English. In Aijmer, K. and B. Altenberg (eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*. London: Longman, 128–144.
Stubbs, Michael
   1995   Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language* 2, 23–55.
   1996   *Text and Corpus Analysis*. Oxford: Blackwell.
   2001   Texts, corpora, and problems of interpretation: a response to Widdowson. *Applied Linguistics* 22, 149–172.
Walker, Crayton
   2004   Factors which influence the collocational behaviour of business English nouns and verbs. Paper given at the annual meeting of the British Association of Applied Linguistics (BAAL) at King's College, London, September 9–11, 2004.
Whitsitt, Sam
   2005   A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics* 10, 283–305.
   n. d.   http://www.natcorp.ox.ac.uk/babyinfo.html (information about BNC baby), last accessed 18 January 2006.