Semantic prosody for the 21ˢᵗ Century: Are prosodies smoothed in academic contexts? A contextual prosodic theoretical perspective

Bill Louw [1], Carmela Chateau [2]

[1] University of Zimbabwe, Harare, Zimbabwe

[2] University of Burgundy, Dijon, France

Abstract

Semantic prosody was first investigated over twenty years ago. It describes the consistent aura of meaning that is created through the general tendencies of the set of collocates associated with the central node word. Sinclair proposed that the semantic prosody of a word or expression would form the outer boundary of the unit of meaning. Since then, many other names have been suggested to describe the phenomenon, including semantic preference and evaluative prosody.This paper summarises work presented in the earliest papers, and then examines new evidence of semantic prosody in specialised corpora. This indicates that semantic prosody is not monolithically positive or negative. It can sometimes be smoothed by certain impersonal scientific contexts. However the underlying polarity quickly rises to the fore in more fiercely debated work. Finally, suggestions for further studies and the development of an over-arching contextualist prosodic theory will be presented.

Keywords: semantic prosody, specialised corpora, corpus linguistics, contextualist prosodic theory

1. Introduction

1.1. The origin of the expression and the earliest studies

The phenomenon that now bears the name 'semantic prosody' was brought to light for the first time through the work of John McHardy Sinclair. The first use of the expression in writing was by Louw, in an article entitled "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosody". However, in this same text, Louw attributed the authorship of the phrase to Sinclair: «This is effectively the first computationally derived 'profile' to appear in print of the phenomenon which Sinclair has begun to refer to as a 'semantic prosody' (personal communication 1988 [1]).» (Louw, 1993: 158).

Previously, in 1987, Sinclair had published a text which, through the study of the verb 'set', highlighted a peculiarity of the phrase 'set in'. In a chapter entitled *The Nature of the Evidence*, the study of the negative prosody of the verb is summarised in two paragraphs: the first contains the sentence «The most striking feature of this phrasal verb is the nature of the subjects. In general they refer to unpleasant states of affairs.» (Sinclair, 1987: 155). The second paragraph deals with the presentation of 'set in' in the COBUILD dictionary.

---

[1] In June 2008, the Forli workshop organized by Guy Aston celebrated the 20th anniversary of the concept of semantic prosody, as named by Sinclair in 1988.

The next study on the subject, from the chronological point of view, was also by Sinclair, and was published in 1991. In "Words and Phrases", the fifth chapter of *Corpus, Concordance, Collocation*, Sinclair re-examined the expression 'set in', echoing much of the 1987 text. He also mentioned the negative elements associated with the verb 'happen'.

1.2. Defining semantic prosody

A review in a recent article presents several studies by the principal authors in this domain in the form of a table (Xiao and McEnery, 2006). Although this table is not exhaustive, it nevertheless provides a reasonable idea of the work undertaken in this area. The definition of semantic prosody in the article identifies it as 'collocational meaning', describing the habitual association between a node word and words with which it co-occurs more frequently than coincidence would allow. «The collocational meaning arising from the interaction between a given node and its typical collocates might be referred to as semantic prosody, 'a form of meaning which is established through the proximity of a consistent series of collocates'» (Louw, 2000: 57) (Xiao and McEnery, 2006: 105-6).

However, semantic prosody as described by Louw extends this definition to include a consistent *series* of collocates. For Louw, the semantic coherence of its collocates taken as a whole is what generates the semantic prosody attached to a word. This key notion is what makes semantic prosody so difficult to identify automatically. It has also given rise to many of the misunderstandings and misinterpretations that have cluttered the debate since the phenomenon was first brought to light.

1.3. The importance of authentic examples

It is worth noting here that semantic prosody cannot be detected except through the study of authentic examples. It is the combined association of different words having the same polarity, positive or negative (generally negative), which identifies the polarity preferentially associated with the node word or expression, and thus its semantic prosody. This methodology reinforces the validity of corpus linguistics as a new way of looking at language: «Corpus data provides us with incontrovertible evidence about how people use language. It allows us to examine in a split second more language than we are likely to use in a lifetime. That is all. Because it is very new, and language is very complex, it is showing us many unexpected features — unexpected because unheralded by intuition, though most of them seem, once mentioned, to have a ring of truth about them.» (Francis and Sinclair, 1994: 191) [2].

Semantic prosody is not available to *a priori* intuition, but becomes visible 'to the naked eye' only through corpus analysis. Thus the corpus may surpass the experience of the native speaker, and allow non-native speakers an equal footing with native speakers with regard to judgments of validity. «Whatever it is in the reader's prior knowledge that accounts for this, it must be verifiable as a piece of text process in our experience of the whole language.» (Louw, 1993: 162).

1.4. Semantic prosody and transitivity

The semantic prosody of a word or phrase can be very specific and precise. Of all the examples of the use of the verb 'set in' studied by Sinclair, only a specific subset reveals negative semantic prosody. Louw's analysis of 'build up' shows that the presence of a transitive or intransitive

---

[2] The article is a response to an advocate of introspection rather than the study of corpus data.

verb may affect the semantic prosody that accompanies it (1993). The article also ends with an educational dimension, proposing the expression 'without feeling' to test the reader's intuition. Only one of the examples corresponds to the expected 'intuitive' usage, where the expression is found at the end of a sentence.

1.5. Semantic prosody as a tool for ironic effect

When there is sufficient distance between the expected collocation and the combination of words proposed by the author, the result is irony. It is obvious that this discrepancy has to be easy to detect, as otherwise the effect of the intended irony would fall flat. A poet seeks to create an effect through the use of language and the tools employed for this purpose are words. The paradox is that while semantic prosody is not immediately accessible to intuition, it must be implicitly perceived by the poet, since effects in a literary text are to some extent the result of deliberate creation. This paradox is not as illogical as it first seems. There is a cline, from the naïve informant, so desirable for research in socio-linguistics, through the linguist, whose intuition has been honed by many years of linguistic analysis, to the author of a literary text, seeking to create a specific effect, and far more sensitive than the norm to effects such as semantic prosody. The naive informant would obviously not be consciously aware of the semantic prosody attached to a given word, but would still be able to perceive its effects without being able to explain them. The linguist would draw on the intuition born of experience to select elements for investigation. In corpus stylistics, the starting point is the literary text, produced by an enlightened author, which corpus analysis will illuminate further: «the application of Sinclair's work to stylistics lies in matching texts against corpora: this supplies the `bottom up' *textual evidence* for what has until now been regarded as 'top down' *prior knowledge* in the act of reading» (Louw, 1989; 1993: 161).

1.6. Irony and/or lack of sincerity

Irony is often deliberate: the author intends to create the effect produced by the reversal of semantic prosody. The analysis of the adjective 'fine' shows that it normally expresses a positive semantic prosody, as in the example 'a fine actress'. This prosody can be reversed, in which case it becomes ironic. This would be the case with the expression 'a fine mother' for example, which is probably ironic. It was used by Barack Obama to describe Sarah Palin: «Obviously, she's – a fine – mother and – up-and-coming public servant» [3]. However, sometimes the speaker is unaware of the irony. In such cases, the irony is unintentional and is directed against the speaker rather than being a tool used by the author. In cases where irony is involuntary, it could be an unconscious revelation of the real attitude of the speaker, a form of Freudian slip. Exceptions to this rule would be those cases where speakers are not using their first language, or in the case of children and younger, inexperienced speakers.

2. Semantic prosody and specialised corpora

Most of the studies investigating semantic prosody have been based on large, general-language corpora. An early study by Partington (1998) used a corpus of newspaper English, arguing that this genre encompassed most others. Gabrielatos and Baker (2008) also used a newspaper corpus for their diachronic study of attitudes to RASIM (Refugees, asylum seekers, immigrants,

---

[3] From the TV show, 60 minutes, on August 30, 2008: http://www.politico.com/blogs/bensmith/0808/Obama_on_Palin_A_fine_mother_and_an_upandcoming_public_servant.html.

and migrants) in the press. Sinclair (1987; 1991; 1996) often used the COBUILD corpus for his lexicographical studies. Louw (1993; 2000; 2007) generally uses data from the Bank of English, also at the University of Birmingham, to assist in the study of literature, a methodology also referred to as corpus stylistics. For this type of analysis, the corpus must contain evidence of all the devices used in literature. A corpus purely composed of newspaper texts would be too restricted to provide sufficient evidence of relexicalisation, as it would contain a greater proportion of delexicalised language. Several authors have used corpus evidence to enrich findings from other domains and specific genres, such as court transcripts (Cotterill, 2001), fairy tales (Levorato, 2003) and poetry (Louw, 2007).

Many criticisms have been directed at corpus linguistics, for example by Chomsky, who stated in a well-publicised interview "Corpus linguistics doesn't mean anything" (Andor, 2004: 97). More serious criticism has been addressed to the problem of representativeness. A corpus is a sample of language. Scientists often study samples when it would be impractical to collect a full set of data. Statistical analysis is often used to prove the validity of the study. This methodology has been used to good effect in corpus linguistics. Kretzchmar (2006) even goes so far as to praise: «the accomplishments of the NeoFirthians regarding collocational frequency and semantic prosody, which only good science could have achieved».

The tools used to study the corpora may also affect the results obtained. This problem can be avoided by using only one set of tools, but for practical reasons this is not always possible. The Corpus of Contemporary American English (COCA, 385 million words) is freely available online through an interface developed by Mark Davies; the British National Corpus (BNC) and the British Academic Written English corpus (BAWE) were consulted using the Corpus Architect interface developed by Adam Kilgarriff et al. (2004), which allows various types of statistical measurements to be computed. The BAWE corpus contains 6,506,995 running words or 8,336,262 tokens. The BNC contains 100 million words or 112,180,030 tokens. Token counts using the Corpus Architect interface include punctuation.

One of the questions frequently raised in corpus work is the appropriate size for a corpus. While general language corpora now contain hundreds of millions of words of text, specialised language can be studied with a much smaller dataset, chiefly because of the greater lexical density of scientific language (Gotti, 2003). The specialised corpora used in this study include subcorpora drawn from the academic writing sections of large, general-language corpora. An even more specialised subset of texts, ScEngL1, was created from the BAWE corpus, using Corpus Architect. Only texts written by English L1 authors were selected, to avoid possible bias, such as transfer of prosody from L1 to L2. It is beyond the scope of this paper to investigate such effects. As the goal here is to examine whether semantic prosody is affected by genre or register, a Physical Sciences subcorpus was created, the working hypothesis being that the hard sciences would be the least likely to be affected by the transfer of semantic prosodies which have been shown to exist in general language. This subcorpus, with just over one million tokens, is composed of texts from ten different fields. Details are given in Tab. 1.

Architecture	18 491	Mathematics	64 086
Chemistry	139 663	Meteorology	51 751
Computer Science	109 648	Other	187 458
Cybernetics & Electronics	57 534	Physics	218 989
Engineering	427 967	Planning	28 577

Table 1: BAWE SciEngL1 subcorpus (1,088,815 tokens)

This broad selection of corpora should also help to address another frequent criticism of corpus studies: the problem of reproducibility. A recent study (Louw and Chateau, submitted) found evidence of similar examples of the same semantic prosody in several different datasets. Another study (Hunston, 2007) re-evaluated semantic prosody, notably drawing upon Stubbs' previous investigation of the negative prosody associated with the verb 'cause', in comparison with a series of examples from the New Scientist. Unfortunately, Hunston's paper does not clearly indicate how the examples were selected, or whether they were merely a random sample. Scientific experiments can only be reproduced correctly if the methodology is clearly described. It is important to avoid such pitfalls wherever possible.

3. Semantic prosody in scientific texts: CAUSE and consequences

Several studies of variation in English, most notably by Biber (1988; 1998) and Swales (1990), have investigated genre or register differences. Academic writing in particular has often been the focus of research. Biber (1998) and Gledhill (2000) have shown that phraseology and other features of academic writing vary in scientific articles from one section of the text to another. It is also generally acknowledged that the context of situation in scientific texts tends to be impersonal, particularly in the methods and materials section. This could perhaps partly explain why there have been few studies of semantic prosody in specialised corpora. And yet the investigation of this phenomenon in different types of texts could provide useful information, notably for pedagogical purposes, whether for native or non-native speakers of English, with a major goal being «to help students comprehend and produce these registers appropriately» (Biber, 1998: 157). This is particularly the case for graduate studies, where most students are required to write in English, regardless of their native language. Semantic prosodies identified in a corpus of business English could be used to «give students insights into how business people use the language» (Nelson, 2006: 233). Investigation of a specialised corpus could also be useful for teachers of English for Specific Purposes if they are not experts in the particular field under study.

One factor which may have delayed investigation of semantic prosody in specialised corpora may be that many specialised contexts would seem to fall outside the bounds of the categorisation of "context of situation" as defined by Firth:

A.   The relevant features of participants: persons, personalities.
(i)   The verbal action of the participants.
(ii)   The non-verbal action of the participants.
B.   The relevant objects.
C.   The effect of the verbal action.
(Firth, 1950, reprinted 1957).

3.1. Science subcorpora and CAUSE

Hunston (2007: 253) warned that it would be preferable to specify that the negative semantic prosody associated with 'cause' is limited to cases involving "human self-interest". Although this would be perfectly coherent with the principles expressed by Firth, the evidence from the BNC tends to suggest that the negative semantic prosody of 'cause' extends even to less animate contexts. In the science subcorpus of the BNC (1,303,172 tokens or around one million words, selected using Dave Lee's classification: W_ac_nat_science, in Corpus Architect), the most frequent lexical collocate for 'cause' is 'damage'. There are in all 23 occurrences, in seven of which 'damage' is followed by CAUSE (the lemma), while in the remainder, 'damage' follows

CAUSE. The next most frequent lexical collocates are 'changes', 'change', 'DNA', 'reduction', 'differences', 'sea', 'loss', 'local', and 'major'. Data from the BAWE Science English L1 subcorpus show a similar pattern. The most frequent lexical collocates are 'error', 'problems', 'change', 'flow', 'pressure', 'damage', 'field', 'current', 'magnetic' and 'interference'.

In the 'academic' section of the COCA corpus, there are 26,461 occurrences of the lemma CAUSE. Almost half (11,419) are in the form 'cause', over half of the remainder (6,822) are in the form 'caused', with almost as many (6,211) in the form 'causes'. The final form 'causing' is three times less frequent (2,009 occurrences). The hundred most frequent collocates (by mutual information) are overwhelmingly negative. The first ten nouns are 'damage', 'death', 'disease', 'harm', 'loss', 'injury', 'pain', 'illness', 'cancer' and 'deaths', which would all come under the category of 'unpleasant outcomes'. The COCA query syntax also allows searches for context synonyms rather than precise words. The search for 'cause' in the context of synonyms of 'good' produces no results, but the search for synonyms of 'bad' within the nine-word window gives five results: 'severe', 'adverse', 'faulty', 'decomposing' and 'putrid'. On this evidence, it might be worth investigating further the counter-examples cited by Hunston (2007). Were the texts perhaps not written by native-speakers? Or was 'cause' used simply for lack of an alternative expression?

3.2. Synonyms of CAUSE: the need for change?

Although the definition for 'cause' in the corpus-based Macmillan dictionary [4] underlines the negative semantic prosody of the verb, it is difficult to find verbs or expressions with the same meaning that do not contain some element of negativity. Louw (2000: 58) showed that 'bring about' was used as an alternative to 'cause' to put a positive spin on death in the debate on euthanasia. The Macmillan definition for 'bring about' is «to make something happen, especially to cause changes in a situation», which brings us back to 'cause' and one of its most frequent collocates in the Physical Sciences subcorpora, the lemma CHANGE. It is possible that the very idea of causing things to happen is unconsciously perceived as negative, since society prefers stability to change. The word 'change' is found only 26 times in the *King James Bible*, for example, but it is used 92 times in Darwin's *On the Origin of Species* [5]. Scientists perform experiments and they cause things to happen. Sometimes their actions are rather like those of the sorcerer's apprentice, but more often their research produces positive results and everybody benefits from their discoveries and inventions.

3.2. Actively bringing about positive change

Further examination of the SciEngL2 subcorpus provides 22 occurrences of 'BRING about'. The first eight are presented in Tab. 3a. They are all active or infinitive forms, and the most frequent collocates are 'energy' and 'change'. Any negativity is given a positive spin, as in the second example, where 'limitations', which are generally negative, "will bring about the necessity for new types of computers". Although 'change' may be frightening and unwanted, 'new' is more often positive than negative. The 'problem' in the final example is linked to 'causation' that will be developed later, rather than to the previous sentence.

---

[4] "to make something happen, usually something bad" <http://www.macmillandictionary.com>.
[5] Data from Bookscraper http://labs.timesonline.co.uk/bookscraper/words/change.

photon no longer carries sufficient energy to	**bring about**	the photolytic dissociation reaction.
increase in performance, these limitations will	**bring about**	the necessity for new types of computers
this goes hand in hand with the means to	**bring about**	improvement" and "the labelling of some
application of radiation. The energy required to	**bring about**	this spin transition depends on the strength
the frequency of radiation required to	**bring about**	resonance is different to what would be
phase difference of the two beams alters	**bringing about**	the interference fringes observed.
These are interrelated i.e. a change in one	**brings about**	changes in the rest of the attributes.
Of an event, 'B', if 'A' is a reason that	**brings about**	the effect 'B'. The problem with causation

Table 3a: 'BRING about' in the SciEngL1 BAWE subcorpus

3.2. Negative effects being brought about

The remaining 14 occurrences, shown in Table 3b, all contain the passive form 'brought about'. Here the prosody tends to be more negative. The sentence containing the word 'disease' is repeated twice, and comes from two different sections of the same text, the first being the abstract. Although this is the most obvious example of negativity, there are several others which are unquestionably negative: 'setbacks', 'bottlenecks' and 'perturbation'; some are more neutral: 'changes', 'action' and 'modification'. Examination of the extended context for the third example shows that 'breakdown' or 'decomposition' is what is 'brought about'.

1 prototypes that can be seen on CAD. Changes	**brought about**	in one part mostly affect the dimensions
2 locally acting signalling hormone. Action is	**brought about**	through binding to the histamine H 2 receptor
3 occurs at the post-synaptic cleft, and is	**brought about**	by the enzyme acetylcholine esterase which
4 , and greatly depend on any modification	**brought about**	to the variables involved in the processes
5 spread of disease. Water-borne disease can be	**brought about**	either by contamination brought on by flooding
6 Climate sensitive. Water-borne disease can be	**brought about**	either by contamination brought on by flooding
7 perturbation of the photostationary state can be	**brought about**	the radical oxidising species, hydroperoxyl
8 due to the single cause that the effect is	**brought about**	. It is also important to point out that
9 seawater intruding into freshwater aquifers	**brought about**	by the rising sea level owing to the melting
10 sub-contractors, reducing resource bottlenecks	**brought about**	by this multitude of contractors and suppliers
11 Nickel(II) but not Copper(I). The specificity is	**brought about**	by the minimum energy conformations of
12 product comprises of many changes that are	**brought about**	in the raw materials and any modifications
13 Minor alterations to account for the changes	**brought about**	by the additional motor control function
14 plans - necessary to deal with setbacks	**brought about**	as a result of environmental and other

Table 3b: 'brought about' in the SciEngL1 BAWE subcorpus

3.2. Giving rise to smoother prosodic effects

The final potential synonym for 'cause' is the expression 'GIVE rise to', which is generally found in negative contexts, as is shown in the COCA corpus, with phrases such as 'false expectations', 'widespread concerns' and 'considerable controversy'. The exception to this negative semantic prosody occurs when the effect is countermanded by the positive aspect of the adjective 'new'. Then there are phrases such as 'new opportunities', 'new challenges' and 'new theories'. In the SciEngL1 BAWE subcorpus (Tab. 4), there are 30 occurrences of the expression, which generally indicate a smoothing of the negative semantic prosody, although there is nothing palpable to counterbalance the effect of the expression. In this type of scientific writing, it is more obviously the absence of the human context that removes the need for negativity. The world of hard science is an impersonal world of cause and effect without human agency. There are six clear exceptions: in examples 1, 6, 10, 18 and 26, there is negativity, most notably in the final example, where people are present and suffering. In examples 1 and 6, 'group' and

'organisation' imply human agents, and for 10 and 18, both scepticism and colour-blindness suggest human participants. Example 18 continues: "Colour blindness is a condition that affects a small percentage of the population".

1 Also, the group "made an acquisition that	**gave rise to**	negative goodwill" in March 2005 which
2 he postulated that the original equations	**gave rise to**	a universe that was expanding or contracting
3 was backed by the photoelectric effect and	**gave rise to**	the knowledge of the region of light. The
4 tested to shed light on the grey areas and	**give rise to**	better accounting standards while at same
5 However, the knowledge acquired from it may	**give rise to**	further experiments that would not have
6 unethical consultancies by an organisation may	**give rise to**	a number of problems. Two such issues are
7 of the complex was square planar, would	**give rise to**	no unpaired electron (with a d 8 config
8 destination, only fairly slow transition	**give rise to**	strong crisp peaks. Proton exchange between
9 electron transitions between orbits that	**give rise to**	the emission spectrum obtained by three
10 , but its development over the years has	**given rise to**	much scepticism as to whether it can really
11 Entry region because of turbulence, which	**gives rise to**	pressure losses. Pressure gains are apparent
12 the products of turbulent fluctuations,	**gives rise to**	additional shear stresses. These turbulent
13 Photons have almost identical energies. This	**gives rise to**	many scattered gamma-rays with energies
14 level to S . Since this is a continuum it	**gives rise to**	many different wavelengths of light, and
15 magnetic field, often defined as , which	**gives rise to**	different spin energy levels. In the absence
16 particularly those near the Fermi sphere,	**gives rise to**	the many conduction properties of metals
17 And so frequencies) can be emitted. This	**gives rise to**	the peaks in the spectral distribution.
18 short-wavelength light (blue) This	**gives rise to**	the occurrence of colour blindness.
19 the body. This rapid change in velocity	**gives rise to**	a velocity gradient perpendicular to the
20 800nm are these d-d transition. This is what	**gives rise to**	the colour of the complex. The strong adsorption
21 adsorption characteristic of the vanadyl centre	**gives rise to**	it's typical blue-green colour. All higher
22 magnetic components to the field. This	**gives rise to**	a hyperfine, 2I+1 splitting pattern. The
23 the electron spins can be parallel - this	**gives rise to**	the configuration 5s 14d 5. So now, with
24 , photolytic mechanisms operate that	**gives rise to**	a vast diversity in the chemical nature
25 The flow generated by the stream-function	**gives rise to**	a pattern of streamlines known as the '
26 impacts has caused suffering of many people,	**giving rise to**	great opposition to future projects.2
27 analogous substitutions for the other limits,	**giving rise to**	the expression baweformula Green
28 no fundamental flaw, or omission, in the model	**giving rise to**	some systematic failure in the simulation
29 divide the waves by partial reflection,	**giving rise to**	two wave fronts with the same wavelength
30 debris that fall away from the frozen core	**giving rise to**	the well known comet tails. (3) They follow

Table 4: Data for 'GIVE rise to' from the SciEngL1 BAWE subcorpus

4. Contextual prosodic theory and scientific writing

Several linguists have sought to find a new name for the phenomenon first expressed using the term 'semantic prosody'. Like many terms describing naturally occurring phenomena, it is metaphorical. Metaphors are not always easy to understand, particularly when some participants in the discussion insist on taking them literally. To avoid further fruitless discussion, although the term semantic prosody has been used throughout this text to describe the phenomenon, the far-reaching implications of semantic prosody will be referred to as contextual prosodic theory, as first suggested in Louw (2000). The implications of the theory for scientific writing in general will be based on the corpus-derived examples discussed here.

Definitions of 'cause' as a verb in corpus-based dictionaries stress its negative semantic prosody. When used as a noun, 'cause' is less monolithic in its prosody, but even so the most frequent expression in the COCA corpus is 'cause for concern'. Even in the academic section of the COCA corpus, the semantic prosody of CAUSE is overwhelmingly negative. However, the academic section of the COCA corpus contains relatively little 'hard science': science and technology is only one of nine categories, the others being mostly related to the Arts and Humanities.

In the subcorpora drawn from the BNC and the BAWE corpus, CAUSE is still frequently found in negative contexts, but when the Firthian context of situation is incomplete, negative prosody is smoothed and becomes neutral. The verb 'cause' is difficult to replace: synonyms are often not single words but multi-word expressions such as 'bring about' or 'give rise to'. Scientists generally seek to avoid verbiage and, in scientific writing, a single-word verb will often be preferred to a phrasal verb or multi-word expression. This is more particularly the case in the methods and materials section, where precision, clarity and brevity are the main aims. However, when the context of situation is close to that of normal language, the negative prosody is reactivated.

Although science aims to be objective, there are times, for example at moments of paradigm shift, when scientists are forced to take sides. Attitudinal or evaluative language comes to the fore, as does 'unscientific' vocabulary, such as verbs of belief, and even religious terminology, as in this description of a well-known paradigm shift in the Earth Sciences, by one of the key figures of the time: «This scientific revolution, like others before it, was long in the making, but it was not until the late 1960s that it achieved wide acceptance. Until that time most earth scientists regarded the theory of continental drift as heresy, but at a meeting of the world's geophysicists held in Moscow in August 1971 it became clear that suddenly a majority of earth scientists throughout the world had accepted this revised version of continental drift as the prevailing orthodoxy. The new form is called global plate tectonics» (Wilson, 1976).

Semantic prosody is generally used appropriately if unconsciously by native speakers. It is now often possible to find some information about the typical semantic prosody of a word or expression in most corpus-based dictionaries. Careful examination of the examples in such dictionaries can also provide information about semantic prosody. In the Macmillan dictionary, for example, although 'cause' as a noun is not marked as negative, the first two sentences cited are "The major cause of these accidents is drivers going too fast." and "an essay on the causes of the First World War".

Scientists as actors must perform experiments, and must therefore make things happen. This may to some extent be seen as 'playing God' but although Darwin was more partial to 'change' than the authors of the King James Bible, 'cause' is equally frequent in both texts (64 and 328 occurrences, respectively), as is 'caused' (19 and 94 occurrences). However, 'causes' occurs only seven times in the King James Bible, whereas it is used 44 times in Darwin's text, just like 'causing', which occurs 12 times in Darwin and only four times in the Bible. The more active forms are far more frequent in the scientific text, while the more passive forms are equally frequent in both.

Further investigation is required into the phenomenon of smoothing in incomplete contexts of situation, such as the impersonal forms of scientific writing. For the non-native user of English as the international language of science, it would be helpful to have more glosses on scientific prosody, in both writing manuals and in dictionaries, in order to avoid as much as possible the negative remarks of reviewers who are sensitive to incorrect prosodies, even if they are unable to explain exactly what it is they find unnatural, and perceive perhaps as unintentional irony.

References

Biber D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber D., Conrad S. and Reppen R. (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Firth J.R. (1957). *Papers in Linguistics.* London: Longman.

Francis G. and Sinclair J. (1994). 'I Bet He Drinks Carling Black Label': A Riposte to Owen on Corpus Grammar. *Applied Linguistics*, 15(2): 190-200.

Gabrielatos C. and Baker P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press 1996-2005. *Journal of English Linguistics*, 36(1): 5-38.

Gotti M. (2003). *Specialized Discourse: Linguistic Features and Changing Conventions*. Bern: Peter Lang.

Hunston S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics*, 12(2): 249-268.

Kilgarriff A., Rychly P., Smrz, P. and Tugwell D. (2004). The Sketch Engine. In Williams, G. and Vessier, S., editors, *Proceedings of Lorient 2004 (11th EURALEX International Congress)*, pp. 105-116.

Kretzschmar Jr. W.A. (2006). Book Review: Empirical Linguistics. *Journal of English Linguistics*, (34): 161-165.

Levorato A. (2003). *Language and Gender in the Fairy Tale Tradition: A Linguistic Analysis of Old and New Story Telling.* Houndsmills (UK): Palgrave Macmillan.

Louw W.E. (1989). Sub-routines in the Integration of Language and Literature. In Carter, R.A. and Walker, R., editors, *Literature and the Learner: Methodological Approaches*, British Council ELT Documents 130, pp. 47-54.

Louw W.E. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Baker, M., Francis, G. and Tognini-Bonelli, E., editors, *Text and Technology: In Honour of John Sinclair.* Amsterdam: John Benjamins, pp. 157-176.

Louw W.E. (2000). Contextual Prosodic Theory: Bringing Semantic Prosodies to Life. In Heffer, C. and Sauntson, H., editors, *Words in Context. In Honour of John Sinclair*, Birmingham: ELR.

Louw W.E. (2007). Collocation as the determinant of Verbal Art. In Miller, D. and Turci, M. editors, *Language and Verbal Art Revisited*, London: Equinox.

Louw W.E. and Chateau C. (submitted). La prosodie sémantique: miroir de la variation contextuelle, à l'écrit comme à l'oral. In *Actes du séminaire Consiglia*, Paris.

Nelson M. (2006). Semantic associations in Business English: A corpus-based analysis. *English for Specific Purposes*, 25(2): 217-34.

Nishina Y. (2007). A Corpus-Driven Approach to Genre Analysis: The Reinvestigation of Academic, Newspaper and Literary Texts. *ELR Journal*, 1 (2).

Partington A. (1998). *Patterns and Meanings: Using Corpora for English Language Research and Teaching*. Amsterdam-Philadelphia: John Benjamins.

Sinclair J.McH. (1987). *Looking Up*. London: Collins.

Sinclair J.McH. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair J.McH. (1996). The search for units of meaning. *Textus*, IX: 75-106.

Wilson J.T. (1976). *Continents Adrift and Continents Aground*. Scientific American.

Xiao Z. and McEnery A. (2006). Collocation, semantic prosody and near synonymy: a cross-linguistic perspective. *Applied Linguistics*, 27(1): 103-129.

Corpora and corpus tools

Corpus of Contemporary American English (COCA) <http://www.americancorpus.org/>.

Corpus Architect <http://ca.sketchengine.co.uk> This interface provides access to several corpora, including: British Academic Written English Corpus (BAWE) and British National Corpus (BNC).