

# Corpus-based cognitive semantics: Extended units of meaning and their implications for Translation Studies

**Karen Korning Zethsen**

Aarhus School of Business, University of Aarhus

*Traditional lexical semantics focuses on the meaning of individual lexemes. Firth (1957) brought our attention to collocations and the fact that meaning is not isolated in the lexeme. In 1996 Sinclair argued for the existence of extended units of meaning which, as the expression indicates, go beyond the lexeme. In recent years Stubbs (2001b), and other corpus linguists have convincingly shown that meaning is a phraseological phenomenon to a high degree. Corpus searches allow us to study lexemes in their immediate context, study their most frequent collocates and thus help us reveal their semantic preferences (Sinclair 1987, 1996) and semantic prosodies (evaluation) (Louw 1993). Some of the findings confirm intuitions, whereas some make us aware of connotations which we have never before consciously known the existence of. In this article, I shall argue for the application of corpus-based cognitive semantics as a tool for researchers within translation studies (TS) who are particularly interested in revealing evaluative aspects of the units of meaning of source texts and their translations. What may formerly have been described as something intangible like an 'atmosphere', now becomes tangible because of the patterns emerging from large numbers of examples. I shall provide empirical examples in various languages of such evaluative patterns which are of course not automatically generated but come about as the result of computer-generated concordance lines and thorough manual analysis.*

## **1. Introduction: TS and corpus linguistics**

Electronic corpora have been used within Translation Studies since the early 1990s and so-called corpus-based Translation Studies is a research area which has really gained momentum during the past decade or so. In a well-know article from 1993, Baker predicts the usefulness of corpus-based research within TS (incidentally in the same anthology that includes Louw's seminal article on semantic prosody) and recent years have seen a profusion of publications on the subject (Laviosa 2002; Olohan 2004; Baker 1999, 2004). Empirical studies within corpus-based TS initially focused on universal features of translation investigating e.g. the hypotheses of simplification and explicitation (Laviosa 2002: 58). Mainstream corpus-based translation research now focuses on the nature of translated language (a corpus like the Translational English Corpus (TEC) has for instance been used to study the distinctive nature of translated text, the style of individual

translators and the impact of individual source languages on the patterning of English) and on studying the differences and similarities between translated and non-translated text (see Laviosa 2002 and Munday 2008 for an overview).

Much research has been carried out within TS using parallel (originals and their translations (Baker 1995)) or comparable (original text and translated text within the same language (Baker 1995)) corpora to study the language of translation and compare it with non-translated language, and this particular strand of TS has developed into a paradigm in its own right with close links to descriptive TS. However, perhaps precisely because corpus-based TS has become so established it is important that we do not neglect the fact that corpus linguistics in general has much to offer the TS scholar, the translation trainer and ultimately the practitioner. In fact, Baker (1999: 282) speculates on the reasons for the failure of corpus linguistics to make more of an impact on TS so far and one of the main reasons she suggests is “the negative image of mainstream linguistics that developed within translation studies during the 80s and 90s, following several decades of simplistic linguistic theorising of translation”. Malmkjær (1998: 534-535) makes more or less the same point and concludes that linguistics and TS have something to learn from each other as both disciplines have language and linguistic activity at their centre. TS scholars can make good use of results and methods from corpus linguistics, which has been around since the early 1970s, and especially from cognitive lexical semantics where exciting progress has been made concerning the unit of meaning and evaluation in language. In this article I intend to explain the theory behind the revolutionary findings and discuss their potential within TS. The first subject I shall turn my attention to is the unit of meaning.

## **2. The unit of meaning (the unit of translation)**

The approach to semantics on which this article is based takes as a starting point the fact that meaning is established in an interpretive process performed by language users. Part of the input to this process is linguistic expressions. The information provided by these expressions is then combined with information from the context in which they occur, triggering different interpretations. As we shall see below, some words nearly always appear in the same context. The meaning of the context therefore contributes to the meaning of the word; that is, certain words presuppose a certain context to such an extent that this context can be said to form part of the lexical meaning of the word. It follows from this that individual word meaning cannot be considered a sound concept within semantic analysis.<sup>1</sup> This is particularly interesting to TS. For, if the unit of meaning has to be extended, so has the unit of translation. Though translators today are often taught to take a textual approach to translation it is difficult, for all practical purposes, not to focus on the word as the unit of translation (and the word is also the most

common dictionary entry). However, if corpus linguistics argues convincingly for a more phraseological approach to meaning – beyond what is traditionally known as idioms and collocations – then this supports the importance of larger units of translation and should be reflected in translation theory as well as in practice.

Lexical semantics within the structuralist tradition entails independent word meaning which ideally can be finitely described by means of componential analysis, and denotation, not connotation, is considered of main interest – a rather static view which does of course not provide a satisfactory description of meaning in language use.<sup>2</sup> Many factors which are significant for the meaning of a word are idiosyncratic and strongly context-dependent. The structuralist tradition has given us many useful findings, but the limitations are evident once we move into the area of language use. If we take a Roschian view of meaning and apply the theory of prototypology,<sup>3</sup> context and encyclopaedic knowledge is taken into account. Instead of merely defining meaning as a question of sense relations within the *language system*, the theory of prototypology considers meaning as a *mental phenomenon* which in addition to inherent lexical meaning helps us account for and describe evaluative meaning which is not necessarily inherent in the lexeme.<sup>4</sup> For practical purposes, we can still work at the level of semes, but instead of attempting an exhaustive analysis of a lexeme, we should aim at a description of prototypical features, inherent or contextual. However, even if we reject the theory of meaning which believes in a finite description of the vocabulary, if we accept the existence of prototypes, if we include inherent as well as contextual aspects of meaning, the structuralist approach still implies that meaning is more or less isolated in the lexeme.

As early as 1934, Porzig made the very interesting observation that certain words co-occur. He pointed out the existence of essential meaning relations such as lick/tongue, blond/hair and bark/dog (see Lyons 1977: 261). Essential meaning relations are what Firth for the first time in 1957 calls “collocations” (“...I propose to bring forward as a technical term, meaning by ‘collocation’, 1957: 194). As is often pointed out (e.g. Lyons 1977: 612), Firth does not define collocability as precisely as one may wish, but one thing is clear: Firth (1957: 196) rejects the Saussurean dualistic notion of signification:

Meaning by collocation is an abstraction at the syntagmatic level and is not directly concerned with the conceptual or idea approach to the meaning of words. One of the meanings of night is its collocability with dark, and of dark, of course, collocation with night.

Firth was not always clear in his writings on collocation, so exactly what kind of importance he attached to the notion of collocation is a difficult question to answer. However, there is no doubt that Firth considered the tendency of lexemes to co-occur in texts an important part of their meaning.

Firth died before the age of computers and corpus linguistics and therefore the pervasive nature of collocation in language in general was not observable to him – as we shall see below, the extent of collocation even defies native-speaker introspection. Since the advent of computers and the subsequent development of corpus linguistics, important developments have been made within lexical semantics. Instead of mainly relying on introspection and a few illustrative examples the semanticist of today is in many cases able to obtain corpus evidence. One of the important consequences of these developments is the discovery that lexical meaning is not so much a question of meaning isolated in the lexeme, but rather in so-called *extended units of meaning* – a term introduced by Sinclair in 1996.

A student of Firth's, John Sinclair has been a central figure within corpus linguistics for decades. Amongst other things, he has studied collocational patterns and the importance of these patterns for the concept of meaning. Louw (1993: 161) writes: "Sinclair's stated position has long been that the pursuit of independent word meaning has been as illusory as it has been sustained". In an article from 1996, "The search for units of meaning", Sinclair develops a model which convincingly argues for the existence, or rather salience, of extended units of meaning (or compound lexical items as he also calls them). Sinclair puts forward the hypothesis that units of meaning are 'largely phrasal', that only a few words are selected independently of other words. His model consists of "four types of co-occurrence relations in extended lexico-semantic units" (Stubbs 2001b: 64 and see also Stubbs 2001a: 449), these four relations being collocation, colligation, semantic preference and semantic prosody (some of the examples below have been taken from Stubbs (2001b: 64ff), who explains Sinclair's model well):

*collocation:*

A frequent co-occurrence of word forms (physical evidence). Directly observable in textual data, e.g.: 'rancid butter'; 'thunderous applause'; 'sustainable development'.

*colligation:*

The co-occurrence of grammatical choices. Colligation is one step more abstract<sup>5</sup> than collocation as it is the outcome of long sequences of analysis (structural evidence, e.g.: 'cases' frequently co-occurs with the grammatical category of quantifier "in **some** cases", "in **many** cases").

*semantic preference:*

A lexical set of frequently occurring collocates, which share a semantic feature, i.e. they belong to the same lexical field. It is an abstract set which is not directly observable, but the preferred lexis can be listed: the adjective 'large' is often followed by words from a group which could be called "quantities and sizes" such as 'number',

‘scale’, ‘part’, ‘amounts’. The verb ‘commit’ is always followed by a noun phrase belonging to a semantic field which could be called “crimes and/or behaviour which is socially disapproved of” such as ‘murder’, ‘adultery’ or ‘sin’ (incidentally, this preference does also express a semantic prosody, which is the subject of the final, and evaluative, co-occurrence relation)

*semantic prosody:*<sup>6</sup>

Word forms which have a tendency to be (or in some cases which are always) followed by words with certain connotations, basically positive or negative (see below for further elaboration), e.g.: the verb ‘cause’ is almost always followed by something negative such as ‘problems’, ‘serious illness’, ‘death’ or ‘damage’. The verb ‘provide’ is mostly followed by positive things such as ‘service’ or ‘support’. By choosing a word form which in itself does not carry negative connotations but which has a negative semantic prosody, the entire extended unit of meaning becomes attitudinal. According to Sinclair (1996: 87-88) a semantic prosody (or “discourse prosody” as Stubbs calls it) “shows how the rest of the item is to be interpreted functionally. Without it, the string of words just “means” – it is not put to use in a viable communication”.

Sinclair (1996: 94) concludes:

So strong are the co-occurrence tendencies of words [collocation], word classes [colligation], meanings [semantic preference] and attitudes [semantic prosody] that we must widen our horizons and expect the units of meaning to be much more extensive and varied than is seen in a single word.

When Sinclair talks about these extended units of meaning he does not mean the completely fixed expressions which we normally understand by collocations or idioms, but rather a fixed system, framework or matrix in which there is room for variation. In the case of colligation the framework may dictate that there has to be a preposition, but it can be one of many, with semantic preference a word with a certain sense may be required, but there may be many words to choose from, with semantic prosody there will be something attitudinal, but this may take many forms.

Stubbs (2001b: 63) concludes that Sinclair’s model contains two closely related key ideas:

- Meaning is typically dispersed over several word-forms which habitually co-occur in text.
- These co-occurring word-forms ‘share’ semantic features.

In this article I shall follow Sinclair and Stubbs and maintain the position that meaning, and in particular evaluative meaning, cannot be limited to the lexeme. Meaning is rather a phrasal phenomenon and it makes more sense to work on the basis of extended units of meaning. Within a model of extended units of meaning, it is at the level of semantic prosody that we find evaluation.

### 3. Evaluative aspects of texts: Semantic prosody

Within a translational context, it is of course important (to the translator and the TS scholar alike) to understand the evaluation, or more specifically the implied attitudinal meaning (Hunston 2007), of a source text as well as possible. Irrespective of the role the source text is to play in a translation, the translator will not be able to decide whether to keep, change, adapt or omit evaluative aspects (in accordance with the *skopos* of the target text) if (s)he is not aware of their existence in the first place as may be the case in connection with subtle attitudinal meaning. And likewise the translator needs to be aware of the more subtle attitudinal features of the target language. The concept of semantic prosody shows us how pervasive evaluation is and certainly that evaluation in text is much more widespread than traditionally assumed.

As early as in 1966, Sinclair noted that the word and the lexical item would not always coincide. In 1987 Sinclair found computationally derived evidence for the existence of basically ‘good/positive’ or ‘bad/negative’ semantic profiles<sup>7</sup> or in other words, whether a word form is likely to be followed by something basically positive or negative.<sup>8</sup> Sinclair found, for instance, that the expression ‘set in’ has a negative subject in the majority of cases, such as ‘rot’, ‘decay’, ‘despair’ and ‘bitterness’, and ‘set in’ is thus described as having a bad semantic profile (see Sinclair 1987: 155-56). That is, the most frequent collocates of ‘set in’ gradually<sup>9</sup> colour the expression itself so that taught by experience we come to expect something negative as a kind of default value when ‘set in’ is uttered. In this way ‘set in’ cannot be seen in isolation – it cannot be semantically accounted for without including the influence of its most frequent co-texts.<sup>10</sup> This phenomenon was later named semantic prosody<sup>11</sup> (Louw 1993: 157), when Louw wrote his much cited article “Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies” in 1993. Louw’s article was directly based on Sinclair’s 1987 work and he defines semantic prosody as “A consistent aura of meaning with which a form is imbued by its collocates” (Louw 1993: 157). Partington (1998: 68) defines the phenomenon more precisely further emphasising the phrasal element: “Semantic prosody refers to the spreading of connotational colouring beyond single word boundaries.”<sup>12</sup>

Louw (1993) has carried out several corpus analyses to corroborate Sinclair's evidence of the existence of semantic prosody. A well-known example from this article is Louw's analysis of 'utterly' which he finds to have an overwhelmingly bad prosody with typical sentences such as "The farmers were utterly against the union" and "In my experience it gets utterly confused" (1993: 160). According to Louw (1993: 157) the phenomenon is largely inaccessible to human intuition and it cannot be retrieved reliably through introspection: "Semantic prosodies [...] are essentially a phenomenon that has been only revealed computationally, and whose extent and development can only be properly traced by computational methods". Olohan (2004: 82) makes the same point.

The important discovery of the existence of semantic prosodies means that we cannot reveal connotative meaning in a text by simply looking at individual words. We must take into account the wider semantic/collocational patterns which these words form part of in order to reach the evaluations which are likely to be triggered in a reader's mind and for this we need computers and corpus studies. Semantic prosody is not a static phenomenon: it develops constantly (which is also why it is impossible to reach a finite description of the vocabulary) and may be difficult to pin down entirely, but it must be considered an indispensable tool for eliciting speaker attitude and making qualified guesses at likely hearer interpretation. The pioneering work by scholars such as Sinclair, Louw and Stubbs points in the direction of a phraseological approach to meaning and in particular to, often subtle, evaluation or speaker/hearer attitude – an area highly relevant for TS.

#### **4. Empirical examples of semantic prosody and the usefulness of the concept as a tool for analysis within TS**

In order to carry out an analysis of semantic prosody it is necessary to have a representative corpus or rather a corpus which is as representative as possible considering the dynamic nature of language. A corpus may be biased in time, in genres, in medium, etc. (for a more detailed discussion see, e.g. Stubbs, 2001b: 223–224), but Stubbs claims that "for many of the more frequent features of language, relatively modest corpora provide adequate evidence" (Stubbs, 2001b: 224). Once the corpus is in place it is possible to carry out a search by means of the head word under investigation and concordance lines are obtained for the words or expressions in question (Key Words in Context). In most cases, the immediate context allows one to interpret each concordance line and to establish a semantic profile of the head word and the extended unit of meaning of which it forms part. In this way electronic searches facilitate the investigation of language patterns which cannot be explored reliably or at a large enough scale manually or by way of introspection.

#### 4.1. Empirical examples

Most empirical studies of semantic prosody (and there are quite a number by now) have been carried out on the English language including those by Stubbs (1995 and 2001b), who provides a classic example with the verb 'cause' where he points out that the traditional definition 'make something happen' should really be 'make something bad happen' as corpus searches show that 'cause' has an overwhelmingly negative prosody (confirmed by Xiao & McEnery 2006 and Dam-Jensen & Zethsen 2008). Other classic examples of words with strong negative prosodies are, as mentioned above, 'set in' (Sinclair 1987: 155-56) and 'utterly' (Louw 1993: 160). Dam-Jensen & Zethsen (2007) have studied 'lead to' as have Xiao & Mcenery (2006), Tognini-Bonelli (2001) has looked at 'largely' and 'broadly' and Channell (2000) has among other lexical items studied 'par for the course' to mention a few.

Little work has been done though on languages other than English, as pointed out by Xiao & McEnery (2006: 103), Munday (forthcoming) and Berber Sardinha (2000). Examples include Partington (1998) and Tognini-Bonelli (2001) on Italian, Berber Sardinha (2000) on Portuguese, Tao (2003) and Xiao & McEnery (2006) on Chinese, and Dam-Jensen & Zethsen (2007) on Danish. That is, up till now, only an insignificant number of non-English studies of semantic prosody have been carried out and very few indeed are contrastive and compare the semantic prosody of equivalent lexical units in different languages. I know of no other studies than the following which were all based on comparable monolingual corpora: Partington (1998) (English and Italian), Berber Sardinha (2000) (English and Portuguese), Xiao & McEnery (2006) (English and Chinese), Dam-Jensen & Zethsen (2008) (English and Danish), and Munday (forthcoming) (English and Spanish).

Partington (1998) provides a good example with the Italian/English pair of seemingly similar words 'impressionante'/'impressive'. The English adjective 'impressive' typically collocates with words such as 'achievement', 'talent' and 'dignity' and can therefore be said to have a positive profile whereas its Italian look-alike 'impresionante' often collocates with neutral or negative terms such as 'series of price rises' and 'assassination attempts' giving it a somewhat unfavourable bias. A search in the Danish Korpus 2000 (accessed June 2005, 778 examples found) shows that the Danish corresponding adjective 'imponerende' like 'impressive' has an overwhelmingly positive semantic profile the most frequent collocates being 'samling' ['collection'], 'resultater' ['results'], 'karriere' ['career'], 'evne' ['ability'/'talent'], 'præstation' ['achievement'] and 'opbud' ['turn-out'] (Zethsen 2006: 16)

Berber Sardinha (1999, referred to in his 2000 article) offers interesting analyses of translational equivalents to words whose semantic profiles have been established in English by other scholars. Among others he analyses 'causar'/'cause' and 'acontecer'/'happen' and Berber Sardinha (2000)

offers a contrastive analysis of ‘cometer’/‘commit’ again among other pairs of translational equivalents. The author concludes that ‘cause’ and ‘causar’ share a negative semantic prosody, ‘acontecer’ and ‘happen’ have different semantic prosodies and that the semantic prosodies of ‘commit’ and ‘cometer’ are similar.

Xiao & McEnery (2006) have carried out a contrastive analysis of ‘cause’ and its Chinese equivalent and their results echo those of Berber Sardinha (2000) and Tognini-Bonelli (2001). Other words such as ‘result’/‘outcome’ and ‘price(s)’ are analysed and compared with their Chinese equivalents. On the basis of their thorough analyses Xiao and McEnery (2006: 124-125) conclude that semantic prosody and semantic preference are as observable in Chinese as they are in English and that the collocational behaviour and semantic prosodies of near synonyms are quite similar in the two languages, even though they are distinctly unrelated. However, they do point out that just as different languages can have different ranges of near synonyms, near synonyms and their close translation equivalents in different languages may also demonstrate, to some extent, different collocational behaviour and semantic prosody.

Dam-Jensen & Zethsen (2007) study the Danish verbs ‘medføre’ and ‘forårsage’ [‘lead to’ and ‘cause’] and Dam-Jensen & Zethsen (2008) compare them to the English verbs ‘lead to’/‘cause’. Even though Danish dictionaries in no way indicate any negative associations in connection with the verb ‘forårsage’ [‘cause’], analysis shows an extremely negative semantic profile, which fits in well with Stubbs’s (1995) analysis of the English verb ‘cause’. The analysis of the synonym ‘medføre’ [‘lead to’], on the other hand, shows that there are clear negative evaluation tendencies, but they are not so overwhelming that they prevent the verb from being used neutrally or even positively, and in most cases the negative implications of ‘medføre’ are not as devastating as are those of ‘forårsage’. In Dam-Jensen & Zethsen (2008), contrastive analyses of ‘forårsage’/‘medføre’ and their English equivalents ‘cause’/‘lead to’ show similar results.

Munday (forthcoming) analyses ‘loom large’ and one of its Spanish dictionary correspondents, ‘cernerse’ and demonstrates that both expressions have generally negative semantic prosodies, but also that there are some notable differences particularly as regards the most common collocates and the syntactic structures in which the two verb phrases are used.

#### **4.2. The potential of the concept of semantic prosody within TS and related areas**

In the above paragraphs, I have discussed the theoretical implications of extended units of meaning versus the traditional focal point – the lexeme. Following Sinclair (1996) I have argued in favour of the more recent tendency which points towards a phraseological approach to meaning, which again entails a more phraseological approach to translation and relevant translation theory. Interpreting the intended meaning of the ST and

of the linguistic choices of the TT is important within TS and practical translation alike, and a phraseological approach means that the concept of semantic prosody becomes of central interest for the interpretation of implied attitudinal meaning and for increased awareness of its salience. In the following I will discuss the potential of semantic prosody analyses within TS and related areas.

The attitude and the intended effect of a text is often not expressed in so many words, but emerges from more subtle connotative meaning. The concept of semantic prosody creates awareness about the fact that attitudinal meaning is frequently part of words, phrases or stretches of text which are conventionally considered to be neutral. The well-developed corpus-based method for revealing these attitudinal markers provides access to an area of research which has only become possible since the advent of electronic corpora. As research on semantic prosody has shown, the connotations acquired in repeated contexts are far from always recorded in monolingual dictionaries and hardly ever in bilingual ones which means that the translator or the translation scholar often do not have access to information on the semantic profiles of words. Time is of the essence to the professional translator and it goes without saying that it is rarely possible to carry out individual queries and analyses (though in some cases it may be worth the effort in connection with keywords or cultural words (see Fairclough 2000; Partington 1998: 8, 76 and Laviosa 2002: 85-86). In other words, translators will undoubtedly welcome lexicographical developments within the field of semantic prosody (Berber Sardinha 2000: 1). Semantic prosody is a huge challenge to lexicographers, as the prosodies rarely apply to all occurrences of a word or phrase and are highly context-dependent just to mention a few of the obstacles.

Semantic prosody is bound with time to influence our perception of the concept of equivalence. A likely hypothesis is that the traditional problem of 'false friends' within translation is much more pervasive than assumed up till now. Presumably equivalent words may have developed differently in two languages and have in time been influenced by the company they have kept and thereby developed different prosodies. Partington (1998: 77) indicates that his own research shows that "look-alike words from two related languages can have very different semantic prosodies" and he concludes: "The pitfalls for translators unaware of such prosodic differences are evident" (1998: 78). Zethsen (2004) shows that Latin-based words of the same roots have often developed very differently in different languages and may therefore end up with very different meanings. Both Partington and Zethsen are concerned with the well-known phenomenon of 'false friends' which concordance lines are well suited to reveal. What is particularly interesting is the fairly recently discovered fact that some words which are not look-alike words (i.e. non-cognate words) in two languages and which, on the strength of dictionary information, can be considered close translational equivalents may in fact have different semantic prosodies, and as noted by Munday (forthcoming: 5), may be said to be a subtler

variation on the old concept of false friends. According to Munday (forthcoming: 6), the translator may in some cases not be intuitively aware of a prosody or may, influenced by source text lexis and structure, inadvertently choose an equivalent which has a different prosody from the original. This could result in a collocational clash potentially altering or blurring the meaning or perhaps with an unintended comic or ironic effect.

Xiao & McEnery (2006) demonstrate the importance of corpus-based contrastive work on semantic prosody to language learning in order for L2 learners not to use words and expressions at odds with their semantic prosodies. It goes without saying that such contrastive studies are equally important to students of translation and TS (Munday forthcoming: 6; Berber Sardinha 2000: 16; Zethsen 2006). In connection with translator training contrastive semantic prosody analyses may also be useful for translation assessment. Kenny (1998) provides a useful illustrative example with the English word 'giro' and an instance of its translation into the German 'Scheckheft'.

On the basis of intuition or specific instances of translation, TS scholars may generate hypotheses about subtle attitudinal features of particular words and phrases in the source as well as the target language. These hypotheses may now be testable by means of comparable, parallel or monolingual corpora. It may also prove useful to test results from monolingual corpus studies (which are far more common than contrastive ones) on comparable or parallel corpora – in this way results from general corpus linguistics may provide input for studies relevant for TS and the applied areas of translator training and translator practise.

## 5. Conclusion

In this article I have argued for the application of corpus-based cognitive semantics as a tool for researchers within Translation Studies who are particularly interested in revealing evaluative aspects of the units of meaning of source texts and their translations. Within a translational context, it is important to be aware of the concept of extended lexical units, on which semantic prosody builds, because it shifts undue focus on individual lexemes as the unit of translation to larger (and more meaningful) units of meaning. The phenomenon of semantic prosody is highly relevant to Translation Studies and allows the translation scholar access to more hidden layers of meaning. If corpora comply with the requirements of sufficient volume and representativity, they facilitate the investigation of language patterns which cannot be explored manually or by way of introspection. Numerous other findings from corpus-based semantics are relevant to Translation Studies and the discipline will benefit from keeping track of major developments within semantics in general. Like all other methods, corpus analysis can of course be criticized, e.g. for providing too little context, for producing only positive data, or for studying only performance and not competence (see Stubbs 2001b: 221–226;

Partington 1998: 144–148). Needless to say corpus searches cannot stand alone – the data must be interpreted and a corpus search is therefore merely a tool (but a very useful one) that complements traditional linguistic analysis.

### Bibliography

- Baker, Mona (1993). "Corpus Linguistics and Translation Studies: Implications and Applications". M. Baker, G. Francis & E. Tognini-Bonelli (eds): *Text and Technology*. Amsterdam: John Benjamins, 233-250.
- Baker, Mona (1995). "Corpora in Translation Studies: An overview and some suggestions for future research". *Target* 7(2), 223-243.
- Baker, Mona (1999). "The Role of Corpora in Investigating the Linguistic Behaviour of Professional Translators". *International Journal of Corpus Linguistics* 4(2), 281-298.
- Baker, Mona (2004). "A Corpus-based View of Similarity and Difference in Translation". *International Journal of Corpus Linguistics* 9(2), 167-193.
- Berber Sardinha, Tony (2000). "Semantic prosodies in English and Portuguese: A contrastive study". *Cuadernos de Filología Inglesa* 9(1), 93-110, available online at [http://www2.lael.pucsp.br/~tony/temp/publications/2000murcia\\_prosodies.pdf](http://www2.lael.pucsp.br/~tony/temp/publications/2000murcia_prosodies.pdf)
- Channell, Joanna (2000). Corpus-based analysis of evaluative lexis. S. Hunston & G. Thompson (eds) (2000). *Evaluation in Text - Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 39-55.
- Dam-Jensen, Helle & Karen Korning Zethsen (2007). "Pragmatic patterns and the lexical system – a reassessment of evaluation in language". *Journal of Pragmatics* 39(9), 1608-1623.
- Dam-Jensen, Helle & Karen Korning Zethsen (2008). "Translator awareness of semantic prosodies". *Target* 20(2), 203-221.
- Laviosa, Sara (2002). *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam/New York: Rodopi.
- Fairclough, Norman (2000). *New Labour, New Language*. London: Routledge.
- Firth, J. R. (1957). *Papers in Linguistics 1934-1951*. London: Oxford University Press.
- Hunston, Susan (2007). "Semantic Prosody Revisited". *International Journal of Corpus Linguistics*, 12(2), 249-268.
- Kenny, Dorothy (1998). "Creatures of Habit? What Translators Usually Do with Words. *Meta* 43(4), 515-523.
- Louw, Bill (1993). "Irony in the text or insincerity in the writer? - the diagnostic potential of semantic prosodies". M. Baker, G. Francis & E. Tognini-Bonelli (eds): *Text and Technology*. Amsterdam: John Benjamins, 157-176.
- Lyons, John (1977). *Semantics*. Cambridge: Cambridge University Press.
- Malmkjær, Kirsten (1998). "Love thy Neighbour: Will Parallel Corpora Endear Linguists to Translators?". *Meta* 43(4), 534-541.
- Munday, Jeremy (2008). *Introducing Translation Studies*. 2nd ed. London/New York: Routledge.
- Munday, Jeremy (forthcoming) "Looming large: a cross-linguistic analysis of semantic prosodies in comparable reference corpora". A. Kruger (ed.). *Corpus-based Translation Studies: Research and Applications*. Manchester: St Jerome.
- Olohan, Maeve (2004). *Introducing Corpora*. London/New York: Routledge.
- Partington, Alan (1998). *Patterns and Meanings. Studies in Corpus Linguistics*. V. 2. Amsterdam: John Benjamins.
- Partington, Alan (2004). "Utterly content in each other's company: Semantic prosody and semantic preference". *International Journal of Corpus Linguistics* 9(1), 131-156.
- Porzig, Walter (1934). "Wesenhafte Bedeutungsbeziehungen". *Beiträge zur deutschen Sprache und Literatur* 58, 70-97.
- Rosch, Eleanor (1973). "On the internal structure of perceptual and semantic categories". T.E. Moore (ed.) (1973). *Cognitive Development and the Acquisition of Language*. New York: Academic Press, 114-144.
- Sinclair, John (1966). "Beginning the study of lexis". C.E. Bazell, J.C. Catford, M.A.K. Halliday & R.H. Robins (eds). *In Memory of J.R. Firth*. London: Longman, 410-30.

- Sinclair, John (ed.) (1987). *Looking Up. An Account of the COBUILD Project in Lexical Computing*. London/Glasgow: Collins.
- Sinclair, John (1996). "The search for units of meaning". *Textus: English Studies in Italy* 9, 75-106.
- Stubbs, Michael (1995). "Collocations and semantic profiles: on the cause of the trouble with quantitative studies". *Functions of Language* 2(1), 23-55.
- Stubbs, Michael (2001a). "On inference theories and code theories: corpus evidence for semantic schemas". *Text* 21(3), 436-465.
- Stubbs, Michael (2001b). *Words and Phrases. Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Tao, Hongyin (2003). "Toward an Emergent View of Lexical Semantics". *Language and Linguistics*, 4(4), 837-856.
- Tognini-Bonelli, Elena (2001). *Corpus Linguistics at Work. Studies in Corpus Linguistics*. V. 6. Amsterdam: John Benjamins.
- Whitsitt, Sam (2005). "A Critique of the Concept of Semantic Prosody". *International Journal of Corpus Linguistics* 10(3), 283-308.
- Xiao, Richard & Tony McEnery (2006). "Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective". *Applied Linguistics* 27(1), 103-129.
- Zethsen, Karen Korning (2004). "Latin-based terms: True or false friends?". *Target* 16(1), 125-142.
- Zethsen, Karen Korning (2006). "Semantic Prosody: Creating Awareness about a Versatile Tool". *Tidsskrift for Sprogforskning*, 4(1), 275-294.

## Corpora

British National Corpus <http://sara.natcorp.ox.ac.uk/lookup.html> (consulted 11.03.08).  
 Korpus 2000 <http://korpus.dsl.dk/korpus2000/indfang.php> (consulted 11.03.08).

<sup>1</sup> Furthermore, this approach to meaning of course challenges the strict semantics/pragmatics boundary, as context is typically seen as belonging to the interface between semantics and pragmatics (see Dam-Jensen & Zethsen 2007).

<sup>2</sup> Though componential analysis and the notion of semantic features, or semes, as tools for the description of meaning (instead of as a general theory of meaning) are very useful.

<sup>3</sup> According to Rosch (1973), human beings categorise by means of prototypes, i.e. many categories are mentally represented by means of schemata of their most characteristic members. Other members constitute borderline cases and are peripheral in nature. Put in another way our linguistic categories have a hard core and blurred or fuzzy edges.

<sup>4</sup> For an in-depth discussion of evaluative meaning see Dam-Jensen & Zethsen (2007).

<sup>5</sup> In Stubbs (2001a) 'colligation' is listed before 'semantic preference' and vice versa in Stubbs 2001b.

<sup>6</sup> The distinction between semantic preference and semantic prosody is not entirely clear-cut and the problem is linked to the semantics/pragmatics question – for a detailed discussion see Partington (2004).

<sup>7</sup> Louw (1993) also uses the expression 'prosodic profile', but interchangeably with 'semantic profile'. It seems that both expressions stand for the results of a corpus search. I.e. whether a lexeme or a text sequence has a good or a bad profile.

<sup>8</sup> That is basically negative/positive, but a more refined categorisation of semantic profiles can of course be made with headings such as 'difficulty', 'reluctance', 'uncertainty' or 'desirable', 'necessary'.

---

<sup>9</sup> “Prosodies are undoubtedly the product of a long period of refinement through historical change [...]” (Louw 1993: 164). Consequently, there must be strong and less strong prosodies as well as prosodies under development.

<sup>10</sup> Also, if we take two consecutive words in a text – which would not normally be described as a collocation – instead of just one lexeme we may find semantic prosodies as well (see Louw 1993 and his example with ‘days are’).

<sup>11</sup> According to Louw the term ‘semantic prosody’ was first coined by Sinclair in 1988 (personal communication between Louw and Sinclair), but Tognini-Bonelli (2001) attributes the term to Louw himself. Anyway, Louw (1993) was the first time the term was seen in print.

<sup>12</sup> According to Whitsitt (2005) Partington’s definition is very widespread, but Whitsitt problematises the fact that various definitions are in use each emphasising different, and according to Whitsitt incompatible, aspects of the phenomenon. Whitsitt is one of the few scholars who severely criticises, in fact rejects, the very concept of semantic prosody. It is not within the scope of this article to discuss Whitsitt’s very detailed and philosophical critique, but I would like to refer particularly interested readers to Whitsitt’s work as expressing not the mainstream view, but as discussing aspects which deserve to be considered. These include a comparison of the various definitions, whether semantic prosody (if it exists) is semantic or pragmatic in nature and the role of intuition.