

**Prefazione agli *Occasional Papers del CeSLiC* – ISSN 1973-221X**

***Quaderni del CeSLiC***

**General Editor – Donna R. Miller**

**Comitato Scientifico:**

**Susanna Bonaldi, Louann Haarman, Donna R. Miller, Paola Nobili, Eva-Maria Thüne**

Mi fa sempre grandissima soddisfazione proporre un nuovo contributo alla serie degli *Occasional Papers*, una collana all'interno dei *Quaderni del Centro di Studi Linguistico-Culturali (CeSLiC)*, un centro di ricerca del quale sono responsabile scientifico e che svolge ricerche nell'ambito del Dipartimento di Lingue e Letterature Straniere e Moderne dell'*Alma Mater Studiorum* – Università di Bologna.

Gli *Occasional Papers* finora pubblicati sono:

**(2005) Fusari, Sabrina, Il direct mail per le organizzazioni nonprofit: analisi retorica interculturale italiano-inglese**

[http://amsacta.cib.unibo.it/archive/00000953/01/CESLIC\\_OP1.pdf](http://amsacta.cib.unibo.it/archive/00000953/01/CESLIC_OP1.pdf)

**(2005) Louw, Bill, Dressing up waiver: a stochastic collocational reading of 'the truth and reconciliation' commission (TRC)**

<http://amsacta.cib.unibo.it/archive/00001142/>

**(2005) Nobili, Paola, 'Saper vivere' con gli altri**

<http://amsacta.cib.unibo.it/archive/00001148/>

**(2006) Witalisz, Alicja, English Linguistic Influence on Polish and other Slavonic Languages**

<http://amsacta.cib.unibo.it/archive/00000918/>

**(2006) Larisa Poutsileva, Raccontare il mondo in lingue diverse: Sara' lo stesso mondo?**

<http://amsacta.cib.unibo.it/archive/00002289/>

**(2007) Mette Rudvin, Stereotypes of 'primitivism' and 'modernity' in immigrant-related discourse in the Italian media**

<http://amsacta.cib.unibo.it/archive/00002293/>

**(2007) Ana Pano, Anglicismos en el lenguaje de la informática en español. El "misterioso mundo del tecnicismo" a través de foros y glosarios en línea.**

<http://amsacta.cib.unibo.it/archive/00002370/>

(2007) Sabrina Fusari, **Idioletti e dialetti nel doppiaggio italiano de I Simpson.**

<http://amsacta.cib.unibo.it/archive/00002182/>

(2008) Alida Maria Silletti, **La traduzione del futuro verbale in ottica contrastiva italiana-francese-inglese**

<http://amsacta.cib.unibo.it/archive/00002366/>

A questi papers si aggiungono le altre pubblicazioni del CeSLiC:

1) la serie di manuali dei *Quaderni del CeSLiC: Functional Grammar Studies for Non-Native Speakers of English*

[http://www2.lingue.unibo.it/ceslic/e\\_libri\\_1\\_func\\_grammar.htm](http://www2.lingue.unibo.it/ceslic/e_libri_1_func_grammar.htm)

che già vanta quattro volumi pubblicati; e

2) gli Atti dei Convegni patrocinati dal centro:

- a cura di D. Londei, D.R. Miller, P. Puccini, Gli atti delle giornate di studio del CeSLiC del 17-18 GIUGNO 2005:  
“**Insegnare le lingue/culture oggi: Il contributo dell’interdisciplinarietà**”,  
<http://amsacta.cib.unibo.it/archive/00002055>

ora disponibile anche in versione cartacea: Londei D., Miller D.R., Puccini P.(a cura di), 2006, **Insegnare le lingue/culture oggi: Il contributo dell’interdisciplinarietà**, Quaderni del CeSLiC, Bologna, Edizioni Asterisco.

Sono felice ora di poter proporre un nuovo e originale contributo su un tema corrente di grande attenzione scientifica: il saggio di **Jane Helen Johnson**, sull’applicazione dei metodi della linguistica dei corpora alla traduzione. La Johnson è docente a contratto presso la Facoltà di Lingue e Letterature Straniere dell’Università di Bologna, dove insegna Grammatica Funzionale nel Corso di LT, e Linguistica Inglese (Linguistica dei Corpora) nel Corso di LM, ‘Lingua società e comunicazione’. Presso la Scuola Superiore di Lingue Moderne per Interpreti e Traduttori a Forlì, insegna Traduzione dall’italiano in inglese. Da molti anni è anche traduttrice professionista.

Il titolo del suo saggio è:

### **CORPUS STYLISTICS AND TRANSLATION**

In questo lavoro, la studiosa fornisce un panorama generale sui metodi della Linguistica dei Corpora, applicati anche allo studio della stilistica letteraria, con l’obiettivo di verificare la loro utilità anche nella traduzione letteraria.

A questo scopo, fa riferimento a numerose ricerche già svolte nell'ambito di stilometria e stilistica dei corpora, dagli studi di Stubbs (2005) sul lessico 'vago' di *Heart of Darkness* di Conrad a quello di Semino e Short (2004), che utilizzano un approccio 'corpus-based' per valutare l'applicabilità del modello di rappresentazione del discorso e del pensiero (Leech e Short, 1981), concepito con riferimento al genere 'narrativa', ad altri generi di discorso scritto.

In modo particolare, evidenzia lo studio di Mahlberg (2007), che utilizza un approccio di stilistica dei corpora per esaminare le funzioni testuali 'locali' – come determinati elementi stilistici nella narrativa di Charles Dickens, soprattutto in *Bleak House* – partendo dall'alta frequenza di certi 'clusters' di parole. Anche a questo studio si ispira l'autrice per la sua ricerca sulla narrativa di Grazia Deledda, dove l'importanza di figure retoriche come la similitudine viene evidenziata tramite un approccio 'corpus-based', come documentato dalle 'concordances' qui presentate.

Riprendendo un altro studio di Mahlberg, e facendo riferimento al proprio lavoro su Deledda e le traduzioni delle sue opere, Johnson propone infine di verificare i risvolti pratici della stilistica dei corpora, specie come metodo per cimentarsi con la traduzione letteraria, utile soprattutto per gli studenti di traduzione ma anche per traduttori affermati e non, i quali cercano uno strumento concreto per prepararsi ad affrontare un testo letterario.

**Concetti chiave:**

linguistica dei corpora; stilistica dei corpora; traduzione; genere letterario

Donna R. Miller

Bologna, li 14 luglio, 2008

# CORPUS STYLISTICS AND TRANSLATION

Jane Johnson  
University of Bologna

## 1 INTRODUCTION

The purpose of this article, originally delivered as a talk for postgraduate students studying Translation Sciences, is to see how the emerging field of corpus stylistics may be brought together with the established one of translation to provide another useful tool for translators.

I will first of all summarise the benefits of using computers to study language and provide an overview of the corpus linguistics tools typically used in linguistic analysis.

I shall then move on to focus on the literary text, illustrating how corpus linguistics tools have been used to examine features of this genre. I shall provide a definition of corpus stylistics and give examples of the type of features that can be highlighted using corpus stylistics, making reference to existing research.

With reference to one particular study of the work of Charles Dickens which applies corpus stylistics techniques to both source text and target text, I shall describe my own research into the work of Grazia Deledda along similar lines, concluding with some implications for translation.

## 2 USING COMPUTERS TO STUDY LANGUAGE

There are many advantages in using computers to study language. First of all, the sheer speed and ease of processing a text by computer instead of manually is something we all perhaps take for granted nowadays. Coupled with this is reliability: a computer is not likely to miss anything out, or count things twice. Human bias is also something which can be avoided using a machine to process the information. Finally, the texts and the results themselves can be stored in an easily retrievable electronic form, making it possible for an analysis to be verified at any moment by others.

Many linguists, particularly those working with language in literature, tend to rely on their intuition in analysing texts. However it has been suggested (e.g. Hunston, 2002) that intuition is particularly unreliable where judgements about collocations, frequency, semantic prosody and pragmatic meaning, and details of phraseology are concerned. Corpora can

provide concrete and quantifiable evidence of these features, as well as a lasting record for corroboration and further analysis by other researchers.

Different techniques in corpus linguistics have been used for linguistic analysis.

The **analysis of collocation** has involved the examination of frequency lists and lists of concordances. Frequency lists may be in the form of lists of the individual words (types) and their frequency and relative frequency in a text or texts as well as lists of multiword items or clusters. These may also be processed to give information about the order of occurrence or the whereabouts in a text of particular items (by means of the Plot option in WordSmith Tools, for example), the collocates occurring with the items themselves, and the 'keyness' of the frequent words in the corpus in relation to a reference corpus.

**Corpus annotation** has also been used to investigate a particular linguistic feature and involves labelling each occurrence of this feature in a text so that it may be automatically extracted and quantified.

### 3 CORPUS LINGUISTICS AND LITERARY TEXTS

As regards the selection of features for analysis in a literary text, Leech and Short note that "[t]he reliance on 'intuitive observation' in selection means that a heavy burden falls on the reader's training in the art of alert reading, and of responding to linguistic and other cues" (1981: 70).

We may add that the burden may be alleviated to a certain extent by the use of Corpus Linguistics techniques.

Concordancing tools may be used to highlight structural boundaries in a text, as Tribble and Jones (1990) demonstrate in their examination of Katherine Mansfield's short story *The Fly*.

Queries involving character names may be used to 'Plot' the appearance of the characters through the text, while a closer examination of the collocates associated with the names themselves can help to provide more information of how the character is presented in the story, or what features are associated with him/her, as Adolphs demonstrates with her example from Tolstoy's *Anna Karenina* reproduced in Figure 1. Evidence from the concordance lines shows how she is portrayed as woman of passion and fleshly desires, emphasised by descriptions of what she wears, and how she blushes and conveys her emotions.

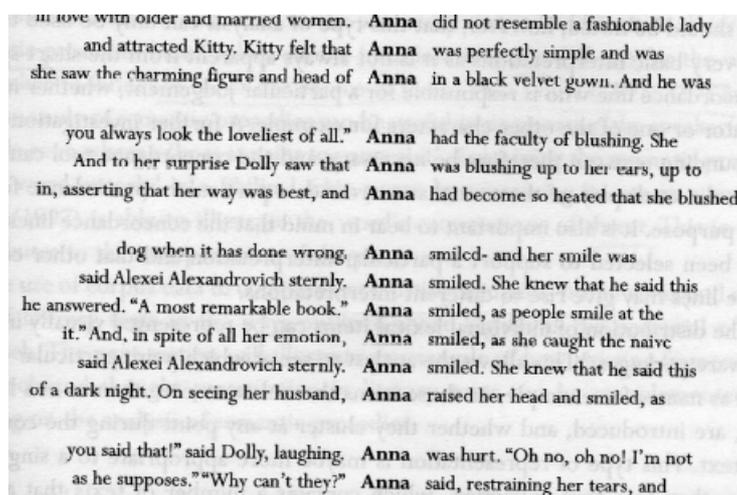


Fig. 1 – concordances of *Anna* (Adolphs 2006: 67)

#### 4 CORPUS LINGUISTICS AND TRANSLATION

Corpus linguistics is now fairly extensively used in translation studies, as summarised by Laviosa (2002), amongst others. As regards the teaching of translation, “[t]he detailed comparison between the Source Text and Target Text brings to the fore not only the ‘unique’ difficulties involved in a particular translation, but also an array of choices, strategies and alternative solutions,” as Agorni reminds us (2007: 200). Corpus linguistics provides a way of doing this, where a translation or translations of the Source Text exist. The sensitivity thus acquired may then be applied prescriptively when approaching a new literary translation.

#### 5 CORPUS STYLISTICS

Stylistics focuses on what makes a text or group of texts distinctive and investigates deviations from linguistic norms that trigger artistic effects and reflect creative ways of using language. Nevertheless, creativity also needs to be measured, since “[t]he more a critic wishes to substantiate what he says about style, the more he will need to point to the linguistic evidence of texts; and linguistic evidence [...] must be couched in terms of numerical frequency” (Leech & Short 1981: 47).

A particular advantage of applying Corpus Linguistics techniques is that quantitative evidence can be provided. This combination of stylistics and corpus linguistics has come to be known as corpus stylistics, in which:

[t]he computer does not provide a single method of text analysis, but offers a range of exploratory techniques for investigating features of texts and corpora. The findings of

corpus stylistics (comparative frequencies, distributions and the like) sometimes document more systematically what literary critics already know (and therefore add to methods of close reading), but they can also reveal otherwise invisible features of long texts. (Stubbs 2005: 22)

As Mahlberg says, “[c]orpus stylistics can provide quantitative data as well as analyse an individual text by providing various options for the comparison of one text with groups of other texts to identify tendencies, intertextual relationships or reflections of social and cultural contents” (2007a: 221).

Corpus stylistics can help to develop descriptive tools to identify and characterise the linguistic and stylistic features that make a text distinctive, summarised by Leech and Short as lexical and grammatical categories, figures of speech, cohesion and context (1981: chapter 3).

Previous research involving corpus stylistics include Stubbs’ study (2005) of the theme of vagueness and uncertainty in Conrad’s *Heart of Darkness* where he considers and uses corpus techniques to look at other lexical items denoting uncertainty, e.g. *something* and *some sort*, which support the status of vagueness as core theme. Turci (2007) uses similar techniques to investigate word forms of the lemma DARK in the same work to look at the portrayal of imperialism.

Other studies have investigated the characterisation of authorship, such as Burrows (1987), who analyses frequency profiles of individual words in Jane Austen novels in relationship to the specific idiolects of particular characters, and Hoover (2002), who identifies authorial style by means of statistical analysis of highly frequent words and word clusters in a corpus of novels by different authors.

Barnbrook (1996) shows how frequency lists of *Frankenstein* can be analysed to extract all nouns that might refer to Frankenstein’s creations, e.g. *monster*, *creature*, then making concordance searches in order to differentiate individual items according to their context of use.

Semino and Short (2004) use a corpus-based approach to investigate how well the model of speech and thought presentation outlined in Leech and Short (1981) worked on written text types other than the novel, by means of a systematic and detailed annotation of a corpus of written fictional and non-fictional narratives for speech, writing and thought presentation categories. Along similar lines but from a Systemic-Functionalist perspective, Taylor Torsello (2007) compares Projection in literary and non-literary texts.

Using a corpus stylistics approach, Michaela Mahlberg (2007a) carried out a study of local textual functions as stylistic features of text in novels by Dickens and particularly in *Bleak House*. She did this by investigating clusters as pointers to meanings and textual functions, comparing clusters in *Bleak House* with clusters in other Dickens' novels. To see whether these stylistic features were typical of Dickens or of contemporary language in general, she also compared them with a corpus of data from the 19th century, since "[w]here an absolute norm for English cannot be relied on, the next best thing is to compare the corpus whose style is under scrutiny with one or more comparable corpuses, thus establishing a relative norm" (Leech & Short 1981: 51).

She then grouped the clusters she identified into functional groups which seemed to be especially relevant to Dickens' style and found that certain functional groups played a particularly important role in characterisation. Her findings were that the link with particular texts was more visible with the longer clusters (e.g. 8 or 9 word clusters). For example, the cluster "*not to put too fine a point on it*" was found to occur 14 times in the novel *Bleak House* and not at all in the others, and is associated with a particular character.

Studies using the corpus stylistics approach like Mahlberg's have picked up on features which have already been intuited by critics. In other words the computer is asked to do something because of the influence of a previous discussion of literary text or by an existing approach to it, thus providing added evidence for themes already identified by literary critics.

However, a computer analysis may also throw up new areas of interest, including "the importance of collocational meaning, which has hitherto not been thought of by critics" (Louw 1997: 247). For example, "[t]he study of unusual collocations with high frequencies in the corpus of writers' whole texts may reveal their predilection for particular collocations; and repeated usual collocations in their works could be considered the characters' idiolects" (Hori 2004: 12).

## 6 CORPUS STYLISTICS AND TRANSLATION

Contrastive linguistic research has been done into similarities and differences between the Source Text and the Target Text, as exemplified by Kenny (2001), who uses corpora to analyse lexical creativity in English translations of German literature in order to investigate normalisation. Lexical cohesion, listed by Leech and Short as one of the stylistic features which makes a text distinctive, lends itself particularly to corpus analysis and was one of the

features investigated by Munday in his study on translation shifts, since “[I]n cases of lexical cohesion, the concordance, by giving rapid access to *all* instances of an item, can be a major help for the analyst in checking for consistency” (1998: 8).

However, few studies have yet been done which bring together corpus stylistics and translation. One of the few researchers to do so is Michaela Mahlberg (2007b), who suggests that, since a human observer is not always aware of the features that are important to a text, local textual functions may be a helpful descriptive tool for translators to use in order to identify subtle patterns in text prior to the process of translation itself, since “corpora can provide information on contextual features of meaning, and [...] information that is difficult or even impossible to retrieve on the basis of small text samples and/or intuition alone” (2007b: 115:).

Mahlberg investigated the 6-word cluster *with his hands in his pockets* that she found to be significant as a Label used for characterisation. She found that the phrase in *Bleak House* and in the Dickens’ corpus as a whole has a highly contextualising function. On comparing these findings with the German translation, she noticed that the translator did not pick up on every detail of the textual functions, which suggests that he had underestimated the stylistic contribution of this phrase.

It must be said however that it is not clear whether this was deliberate or because translator was unaware of the stylistic features that had been brought to light with a concordance analysis.

## 6.1 STYLISTIC FEATURES IN GRAZIA DELEDDA AND TRANSLATIONS: a work in progress

Following the methodology used by Mahlberg, I am currently investigating a number of works by the Sardinian writer and Nobel Prize winner Grazia Deledda using Corpus Linguistics techniques.

My aim is first of all to see what stylistic features emerged from individual and multi-wordlists of a corpus of her work. The next objective is to compare these findings with English translations of her novels, and finally to see whether any suggestions for best translation practice could arise from these analyses.

Important stylistic features would ideally be highlighted by comparison with a larger reference corpus but even a simple examination of the word list for the corpus of novels itself throws up some interesting starting points, some of which are listed below.

### 6.1.1 figures of speech

Something which emerges from a single and multi-wordlist of Deledda's work is the high frequency of words and phrases suggesting the use of simile. The explanation for this could be, as suggested by Leech and Short, that "[t]o give us the illusion of direct experience, a writer often turns [...] to affective language, and the analogical language of simile and metaphor" (1981: 184). It may also be that these are 'loans' from the author's native Sardinian and further linguistic research is required to investigate this hypothesis.

Some examples are shown in Figure 2.

e gli alberi, succhiati dal vento, **pareva** non dovessero mai più scuotersi per la loro stanchezza. Come la farfalla sangue gli sembravano causati dal roteare della terra; gli **pareva** che il moto perpetuo di tutte le cose lo trascinasse sfondo della porta gli passarono in mente; gli **pareva** di trovarsi ancora a Nuoro, in una sera fantastica, e volgeva or di là or di qua la spilla della cravatta che **pareva** un piccolo girasole. "E ti ho persino invidiato,

Fig. 2 – concordances of *pareva* (Deledda corpus)

### 6.1.2. body parts

Certain body parts, eg. *mani*, *occhi*, *viso*, *testa*, occur with particularly high frequency in the Deledda corpus. Indeed, the "use of a bodily part instead of a person as an actor in a clause is a fairly common device for suggesting that the part of the body involved acts of its own accord. This changes the way in which we ascribe motivation for acts" (Leech & Short 1981: 191). This has also been picked up more recently by Simpson (2004).

Some examples are shown in Figure 3 below:

"Ma tu hai lasciato quello che avevi?", domandò in fretta senza sollevar gli **occhi**. Giacinto non rispose subito; mentre donna Ruth lo guardava fisso con **occhi** lucidi, ed Efix traeva dalla bisaccia un bariletto di vino, e pur così "Parlano più i tuoi **occhi** che la mia lingua, Grixè." "Il fuoco ti mangi le palpebre!" "E state zit Giacinto sollevò rapido gli **occhi**, riabbassandoli tosto. **Occhi** pieni di disperazione. "E perché nascere?" "Oh "Troppo lontano." E Giacinto, con gli **occhi** distratti, riprese a narrare delle favolose ricchezze dei Signori del "La **occhi** la saetta, don Predu!", imprecò Natòlia, pur guardandolo con **occhi** languidi per tentarne la conquista.

Fig. 3 – concordances of *occhi* (Deledda corpus)

### 6.1.3 characterisation

As has been said, a concordance search for the protagonist's name provides an immediate picture of actions and feelings occurring in the textual vicinity of the character. The way in which a character is depicted in a story often expressed in verbs and adverbs illustrating his or her actions and in adjectives used to describe the person. Thus a simple concordance search can provide a general idea of how a character is presented in a story or what recurrent actions or features are associated.

Deledda's novel *La Madre* revolves around 3 characters in a remote Sardinian village: the mother of the young parish priest, Paul, the priest himself, and Agnes, the lonely woman through whom the lives of both mother and son are wrecked.

An examination of concordances for a key character in the novel, the madre herself, allows us to focus on the verbal groups close to the node word. We notice in Figure 4 that the concordances contain mostly Mental processes, very few Material processes, and Verbal processes which are all *sottovoce*, helping to construct the figure of the mother as a silent spectator to the almost Greek tragedy which unfolds before her:

il Re Nicodemo, gravemente malato". La madre sollevò gli occhi uscendo dalla sua pena. Ricordò che il Re Parlava come un omino, il piccolo sagrista; e la madre del prete s'intenerì. Dopo tutto, se Antioco parlava così "I più giovani!", disse come fra sé la madre. "Ma perché non sanno. Poi si possono pentire. Possono anche La madre sorrise; eppure una visione rapida di bei bambini sparsi per la casa la tu egli impose, senza sollevare gli occhi; ma la madre vide le ciglia di lui sbattersi e il viso farsi bianco come s'era non senza prima aver osservato che viso faceva la madre del prete. La madre del prete gli sorrise e gli accennò di avevano paura a star soli: la madre ogni tanto vedeva gli occhi vaganti e inquieti del figlio fermarsi, divent "È ossessa, ecco!", mormorò la madre, arrossendo di vergogna. Allora senz'altro Antioco la fece entrare, anzi "È stata molto al sole?", domandò. "Non è questo", disse la madre sottovoce: "io credo che sia posseduta dallo

Fig. 4 – concordances of *madre* (Deledda corpus)

These are just a few of the features which emerged as worthy of interest and further investigation using corpus linguistics tools. Whether or not they are peculiar to Deledda would of course have to be verified by resorting to a reference corpus of late 19<sup>th</sup>-century Italian fiction.

## 6.2 Corresponding features in the English translation of *La Madre*<sup>1</sup>

---

<sup>1</sup>Deledda, G. (1923) *The Mother*. M. Steegman (trans.). New York: Macmillan

A parallel aligned corpus would be particularly useful to compare these features in the ST with those in the Target Text, but even without alignment, wordlists and concordances are still useful.

6.2.1 The use of simile in the English translation emerges from multiword frequency lists in phrases such as *as if/though*, amongst other things. Figure 5 shows some examples:

The previous night! He felt **as if** he had been suddenly awakened out of a long sleep, and the distress of h  
“I am growing childish!” she sobbed. She felt **as if** she had grown twenty years older in that one long day of  
I enjoy ringing the church bells and I feel **as if** I was a bird up in the tower. And haven’t I had an amusing time  
He felt **as if** he were revolving round and round like a top, then pulled himself togeth  
a power stronger than her own will kept her down. Her knees trembled **as though** trying to rebel against that  
her knees trembled, but her feet refused to move, and it was **as though** two compelling hands were holding her  
her short, broad figure clothed in rusty black she looked **as though** she had been hewn with an axe, all of a piece  
trunk of blowing out her skirts and the handkerchief over her head, **as though** it were trying to force her back into

Fig. 5 – concordances of *as if/as though* (Deledda GB corpus)

6.2.2 A glance down the frequency list and the corresponding concordances shows how body parts have been used instead of actors in the same way in the English translation. A few are shown in Figure 6:

something evanescent, doubtful. Her large dark **eyes** fell instantly on the basket of fruit standing on the table, then  
as though she knew no other way of hiding the feelings her **eyes** betrayed too plainly. And those **eyes**, and that way  
a one-sided smile that touched but one corner of his lips, his **eyes** were sad as they rested on the village which now

Fig. 6 – concordances of *eyes* (Deledda GB corpus)

6.2.3 A concordance search for the key word *mother* highlights an equivalent use of mainly Mental processes, with Verbal processes expressed with *whisper*, *wail* etc. as shown in Figure 7:

allowed himself for some considerable time now? The **mother** remembered that lately she had several times  
Actually, in the depths of her maternal heart, his **mother** delighted to see him so handsome and strong,  
“Oh Lord, oh Lord!” wailed the **mother**, and her voice sounded like the voice of some other woman speaking. T  
surrounded by the welcoming throng, the **mother** saw him falter as though about to fall, a man supported him  
longing for a joy not of this world, and the **mother** felt the balm of those tears falling on her heart even in this  
While he was talking thus the **mother** thought to herself: “He is only saying these things in order to persuad  
“Why have you not gone to bed?” he asked curtly. His **mother** turned to look at him, her dream-haunted face still  
He did not answer, made no sign that he had heard. His **mother** raised her eyes and beheld him standing tall and  
above her, he heard his **mother** weeping unrestrainedly, as though she were weeping for the dead.  
Why had his **mother** wept so bitterly, after he had sworn an oath to her? What could she hav

Fig. 7 – concordances of *mother* (Deledda GB corpus)

Certainly an equivalence seems to have been obtained, at least superficially, and restricted to the 9-word window. It is of course important however to remember that this type of analysis can give only a very basic interpretation as the surrounding text would also need to be examined.

## 7 CONCLUSION

Current procedures for preparing for literary translation have been summarised, rather loosely, by Boase-Beier as “getting to know the original (by critical reading and reading of secondary works and other works by the same author)” (1998: 35). Corpus stylistics provides a way of doing this, highlighting possible elements to be subjected to a detailed investigation. It would also be profitable to refer back to Leech and Short’s linguistic and stylistic categories to identify particular elements to focus on.

As exemplified by Malhberg, a corpus search for clusters in specific sections of text could be a useful starting point for identifying core patterns of textual functions, identifying those which are specific to the author’s style and perhaps the particular novel.

It would also be useful to compare our literary corpus with a corpus of text contemporary to the writing of the novel, since contemporary corpora like the BNC is clearly inappropriate for comparison with texts from other periods.

To conclude then, a corpus stylistics approach could also be exploited by literary translators in order to begin the task of translation with a more thorough knowledge of the Source Text. More research still needs to be done into the type of investigation that is required and in defining the type of reference corpus.

It would also seem feasible to use a corpus stylistics approach descriptively to evaluate the success of a particular translation or compare different translations of the same text, though again more research is needed into how far translation competence can be evaluated using corpus linguistics techniques.

## Bibliography

- Adolphs, S. (2006) *Introducing electronic text analysis. A practical guide for language and literary students*. London & New York: Routledge
- Agorni M. (2007) "Translation Teaching and Methodology", in D.R. Miller & M. Turci (eds.), *Language and Verbal Art Revisited. Linguistic Approaches to the Study of Literature*. London: Equinox Publishing Ltd.
- Barnbrook, G. (1996) *Language and computers: a practical introduction to the computer analysis of language*. Edinburgh: Edinburgh University Press
- Boase-Beier, J. (1998) "Can you train literary translators?", in P. Bush & K. Malmkjær (eds.), *Rimbaud's Rainbow. Literary Translation in Higher Education*. Amsterdam: John Benjamins Publishing Company.
- Burrows, J. F. (1987) *Computation into criticism. A study of Jane Austen's novels and an experiment in method*. Oxford: Clarendon.
- Hoover, D.L. (2002) 'Frequent word sequences and statistical stylistics', *Literary and Linguistic Computing*, 17(2): 157-80
- Hori, M. (2004) *Investigating Dickens' Style. A Collocational Analysis*. Basingstoke: Palgrave MacMillan
- Hunston, S. (2002) *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press
- Kenny, D. (2001) *Lexis and creativity in Translation. A corpus-based study*. Manchester: St. Jerome Publishing
- Laviosa, S. (2002) *Corpus-based Translation Studies. Theory, Findings, Applications*. Amsterdam: Rodopi.
- Leech, G. N. & M. H. Short (1981) *Style in fiction: a linguistic introduction to English fictional prose*. London: Longman.
- Louw, B. (1997) "The role of corpora in critical literary appreciation", in A. Wichmann, S. Fligelstone, T. McEnery, G. Knowles (eds.), *Teaching and Language Corpora*. London: Longman
- Mahlberg, M. (2007a) "Corpus stylistics: bridging the gap between linguistic and literary studies", in M. Hoey, M. Mahlberg, M. Stubbs, W. Teubert (eds.), *Text, discourse and corpora*. London: Continuum.

- Mahlberg, M. (2007b) “Corpora and translation studies: textual functions of lexis in Bleak House and in a translation of the novel into German”, in M. Gatto & G. Todisco (eds.), *Translation. The State of the Art/ La Traduzione. Lo stato dell'arte*, Longo, Ravenna.
- Munday, J. (1998) ‘A computer-assisted approach to the analysis of translation shifts’, *Meta.*, XLIII, 4, 1998
- Semino, E. & M.H. Short (2004) *Corpus Stylistics. Speech, writing and thought presentation in a corpus of English writing*. London: Routledge
- Simpson, P. (2004) *Stylistics: A resource book for students*. London: Routledge
- Stubbs, M. (2005) ‘Conrad in the computer: examples of quantitative stylistics methods’, *Language and Literature*, 14 (1), 5-24.
- Taylor Torsello, C. (2007) “Projection in literary and in non-literary texts”, in D.R. Miller & M. Turci (eds.), *Language and Verbal Art Revisited. Linguistic Approaches to the Study of Literature*. London: Equinox Publishing Ltd.
- Tribble, C. & G. Jones (1990) *Concordances in the classroom. A resource book for teachers*. London: Longman
- Turci, M. (2007) “The meaning of ‘dark\*’ in Joseph Conrad’s *Heart of Darkness*”, in D.R. Miller & M. Turci (eds.), *Language and Verbal Art Revisited. Linguistic Approaches to the Study of Literature*. London: Equinox Publishing Ltd.