

The phraseology of Jane Austen's *Persuasion*: Phraseological units as carriers of meaning

Bettina Starcke
University of Trier

Abstract

Phrases are carriers of meaning in language. One way of extracting the literary meaning of a text is therefore to look at its most frequent phrases and their collocates. In this article, I will extract the most frequent phraseology of Jane Austen's novel Persuasion and present a detailed analysis of its two most frequent phrases. I will then use the data generated by these analyses to illustrate how computer assisted techniques can reveal new shades of meaning in a text even after nearly two hundred years of literary criticism. This demonstrates that corpus stylistics is not only a linguistic discipline, but that it also provides valuable data for the interpretation of literary works.

1 Introduction

It is a truth universally acknowledged that literary meaning in a text is conveyed by the language of the text. In order to decode the meaning of a text, we therefore have to look at its language. And since phrases have been found to be units of meaning in language (e.g. Sinclair 2004; Stubbs 2005b), looking at the phraseology of a text is an appropriate basis for decoding the literary meaning of a novel. This will be illustrated by an analysis of the most frequent phrases in Jane Austen's novel *Persuasion*.

I will therefore first extract the novel's most frequent recurrent 3-word sequences (called 3-grams, *she could not* being the most frequent 3-gram of the novel) and its most frequent 3-phrase-frames (a 3-gram with one variable slot, *the * of* being the novel's most frequent 3-frame). This terminology has been adopted from Fletcher (2002) whose software *kfNgram* was used to extract the phrases. Second, I will analyse the semantic and syntactic patterns co-occurring with the novel's two most frequent phrases. The software used to extract the concordance lines is Scott's (1999) software suite *WordSmith Tools*. Third, the

findings from these analyses will be the basis for a literary interpretation of personal relationships in the novel and of its overall atmosphere.

1.1 *The theory*

Corpus linguistics is an empirical discipline which assumes a correlation between the frequency of a linguistic structure and its importance for the structure and the meaning of a text (Teubert 2005). So if for instance one phrase frequently occurs in a general language corpus such as the *British National Corpus*, we can assume that it is frequent in general English. It therefore fulfils a particular function in the language. The second basic assumption is that the more frequent an item is, the greater is its significance for the structure or the content of a text. The analysis of the frequency of linguistic items is the only available objective evidence for the significance of this item in the text. This also applies to literary discourse where frequent phrases and their collocates contribute to the meaning of a text and to its structural organisation. While evidence for this proposition has been given in numerous corpus linguistic studies with regard to general English, this article shows that the most frequent phrases not only contribute to the organisation of the language but also to the encoding of meaning in a text. I will demonstrate this by showing that the use of the most frequent phrases in *Persuasion* is often limited to single characters of the novel or that the phrases often occur in similar kinds of situations. They are therefore one means of characterising the protagonists and events in the plot and of creating an overall atmosphere in the novel. This is the case even though Austen is unlikely to have consciously distributed the phrases to specific situations in the plot. Nevertheless, their frequency and distribution throughout the novel contribute to the literary character of the text and convey literary meaning. To decode how this meaning is expressed in the text is the aim of my phraseological analyses.

Phraseology has become a major area of interest in corpus linguistics. A large amount of research has focused on multi-word units as units of meaning in language (e.g. Sinclair 1996; Partington 2004). Stylistics has not taken much notice of the notion that literary meaning is frequently expressed by phrases and not single words. Naciscione (2001) is one of the few linguists who has worked on a phraseological analysis of literary works. In her analysis of works by Chaucer and D.H. Lawrence, she looks for recurrent phraseological units with partly or completely figurative meaning. While this excludes an automatic extraction of phrases, her work nevertheless shows that also in literary texts, units of meaning frequently consist of more than one word. This notion will be further devel-

oped in the present paper on the basis of phrases automatically extracted from the text.

Corpus linguistic techniques in general have only rarely been used in the analysis of literary texts, though see Tabata (2002), Stubbs (2005a) and, for an early computational study of Jane Austen's novels, Burrows (1987). Techniques used in this work include analysing wordlists and identifying recurrent semantic fields in the texts. This information then serves as the basis for interpretations of literary effects in the data.

Drawing literary conclusions from corpus linguistic data on *Persuasion* is the first aim of this paper. This includes demonstrating how the atmosphere of the novel is created by linguistic means. The second aim is to explain some intuitive reactions on the text from its readers. These aims will be realised by looking at the most frequent phraseological units in the novel and their collocational patterns.

1.2 The novel

A brief summary of the plot of the novel should enable readers who are not familiar with the novel to follow the turns of the plot that will be mentioned in the course of the phraseological analysis.

Anne Elliot, the protagonist, is the second daughter of Sir Walter Elliot, an impoverished, vain and rank-conscious baronet. She is living with her father and older sister Elizabeth at Kellynch Hall before financial difficulties force them to rent their house to Admiral and Mrs Croft and move to Bath. In the course of accepting the Crofts as tenants, it turns out that Mrs Croft is the sister of Captain Wentworth to whom Anne had been engaged nearly eight years ago. The engagement did not last long since Anne soon broke it off on the advice of her motherly friend Lady Russel. Because of his lack of influential friends, the captain did not then seem to have a promising career in the navy and Lady Russel was convinced that Anne should be married to a more wealthy man.

When Sir Walter and Elizabeth Elliot first settle in Bath, Anne stays behind with her younger sister Mary who is married to Charles Musgrove and is living close to Kellynch. At the same time, Captain Wentworth, now a rich man and back in England after having been at war, visits his sister Mrs Croft at Kellynch and he and Anne are forced by social conventions in the neighbourhood to meet regularly. While Anne is still in love with Captain Wentworth, he starts flirting with the two sisters of Mary's husband, Henrietta and Louisa Musgrove, and Anne's and Captain Wentworth's acquaintance is dominated by avoidance of each other and lack of communication. Meanwhile the Musgroves are convinced that Captain Wentworth will marry one of the Musgrove sisters but Henrietta

soon becomes engaged to her cousin Charles Hayter, and Louisa, after some complications in the plot including a serious accident, becomes engaged to a friend of Captain Wentworth's, Captain Benwick. Unexpectedly to Anne, Captain Wentworth is now single again.

On moving to Bath to join her father and sister, Anne is introduced to her cousin Mr William Elliot who falls in love with her. Perceiving this, Captain Wentworth, who has meanwhile also come to Bath, realises that he still loves Anne. When the Musgrove family comes to Bath a little later, Anne and Captain Wentworth again regularly meet, but his jealousy of Mr Elliot and general insecurity with regard to Anne's feelings prevent him from talking much to her. When Anne realises what his feelings are, she takes an active part in forming the relationship between them by telling Captain Harville, a friend of Anne and Captain Wentworth, about the constancy of a woman's feelings – fully aware that Captain Wentworth would overhear them. As had been her plan, he realises that she still loves him and that his jealousy of Mr Elliot is unfounded. By the end of the novel, Captain Wentworth and Anne have again become engaged to be married.

1.3 The first analyses

The first step in the analysis was to extract and classify the novel's (circa 83,400 tokens) most frequent 3-grams and 3-frames.

The top 16 3-grams are:

Table 1: 3-grams

<i>she could not</i>	54	<i>it was a</i>	28
<i>she had been</i>	36	<i>as soon as</i>	23
<i>i am sure</i>	35	<i>had not been</i>	23
<i>a great deal</i>	34	<i>he did not</i>	21
<i>he had been</i>	34	<i>in the world</i>	21
<i>it would be</i>	32	<i>that he had</i>	21
<i>could not be</i>	29	<i>the Miss Musgroves</i>	21
<i>i do not</i>	28	<i>there had been</i>	21

Looking at the grammatical structure of these 3-grams, there are

- twelve phrases that include an auxiliary verb; ten of these phrases contain a third person pronoun and seem to introduce facts, possibly as part of a narrative
- eight personal pronouns
- five grammatical negatives.

The top 16 3-frames are:

Table 2: 3-frames

<i>the * of</i>	723	<i>had been *</i>	255
<i>of the *</i>	429	<i>the * and</i>	230
<i>* of the</i>	429	<i>she had *</i>	227
<i>to be *</i>	378	<i>* she had</i>	227
<i>* to be</i>	378	<i>* it was</i>	220
<i>* in the</i>	324	<i>it was *</i>	220
<i>in the *</i>	324	<i>* Captain Wentworth</i>	196
<i>* had been</i>	255	<i>Captain Wentworth *</i>	196

Only nine of these 16 3-frames have distinctly different semantic and grammatical structures. Fourteen occur once with the * at their beginnings and once at their ends. These frames will be called *doubles*. This dominance of doubles within the phrase-frames poses the methodological question of whether to count them as one or two frames. Their semantic and syntactic similarity appears to be a reason for blending the two frames into one for analytic purposes. But in view of the automatic extraction of the frames, I have chosen to retain doubles as two separate frames since they are part of the original data the software has generated.

Further semantic and grammatical patterns are noticeable in the list of 3-frames:

- six 3-frames include forms of the lemma BE which might indicate a declarative and possibly narrative structure of the text with the past perfect in some phrases indicating events prior to the fictional present which is set in the past tense (*to be **, ** to be*, ** had been*, *had been **, ** it was*, *it was **)
- six 3-frames denote physical, quantitative or qualitative relationships (*the * of*, *of the **, ** of the*, ** in the*, *in the **, *the * and*)
- four 3-frames include pronouns (*she had **, ** she had*, ** it was*, *it was **).

Grammatical negatives, which occur in nearly one third of the 3-grams, do not occur among the 3-frames.

Contrasting the results of the analysis of the 3-grams and the 3-frames, the greater dependence on the content of the data of n-grams compared to the

phrase-frames is visible. Phrase-frames indicate dominant grammatical and syntactic patterns which are partly variable. They provide information on the grammatical structure of the data, n-grams on the content of the data.

Comparing the 3-grams and 3-frames, there are five overlaps:

Table 3: Comparison between 3-grams and 3-frames

3-gram	3-frames
she had been	<i>she had</i> * (<i>been</i> is the most frequent realisation of the *)
he had been	* <i>had been</i> (<i>he</i> is the second most frequent realisation of the *)
it was a	<i>it was</i> * (<i>a</i> is the most frequent realisation of the *)
in the world	<i>in the</i> * (<i>world</i> is the most frequent realisation of the *)
there had been	* <i>had been</i> (<i>there</i> is the third most frequent realisation of the *)

It is striking that four phrases occurring on both lists include a form of the lemma BE. As pointed out above, this hints at a dominant declarative and possibly narrative structure of the novel. The narrative structure is most clearly indicated by the two tenses in which the lemma BE occurs in the phrases. The presence of both the past tense and the past perfect indicate a temporal sequence which has been named by both E.M. Forster (1927) and Labov (1972) as one of the main characteristics of a narrative. I will return to that point later in the analysis. The frequent co-occurrence of BE with both the male and the female third person singular pronouns also indicates the novel's focus on people and their actions and emotions. Summing up these observations, the focus of the novel seems to be on its characters and on the chronological development of the plot. While this is not news to literary critics, additional information on the protagonists can be extracted from the concordance lines of single phrases. I will illustrate this later in the analysis.

2 *More detailed analyses*

To substantiate these claims, I will now discuss in more detail the novel's two most frequent 3-grams, *she could not* and *she had been*. The first phrase occurs

on average every 4.7 pages of the book, the second on average every 7.1 pages. On the assumption that frequency and significance are linked, this frequency indicates that the two phrases are major contributors to the organisation of the text. Furthermore, both phrases are realisations of two of the most frequent 3-frames of the novel so that also their variants contribute to the organisation of the novel's discourse.

2.1 *she could not*

The concordance lines of *she could not* (54) show that the main collocations are with

- psychological verbs, nouns or adjectives (41) as for example in:

ready acknowledged it to herself, and	she could not think differently, let hi
at was to be their sort of intercourse,	she could not foresee. It was all quite
ill usage between him and his friend.	She could not endure that such a friend
other; and checking the surprise which	she could not but feel at such an appea
rying, but very painful agitation, as	she could not recover from, till enabled

- the reception or production of language (20) as for example in:

d thoughtful and not disposed to talk,	she could not avoid hearing many undesir
father and Elizabeth caught her ear.	She could not distinguish, but she must
cter of his manner was embarrassment.	She could not have called it either cold
en existence or when hope is gone."	She could not immediately have uttered a
m Lyme, which found their way to Anne,	she could not tell how, and brought a r

- indicators of time (20) as for example in:

enant." Anne could listen no longer;	she could not even have told how the po
nting him. This was almost cruel. But	she could not be long ungrateful; he wa
ave said, that "now Miss Anne was come,	she could not suppose herself at all wa
h they had now been acquainted a month,	she could not be satisfied that she rea
the prospect of an hour of agitation.	She could not quit that room in peace wi

- and grammatical negatives, negative affixes and quantifiers (15) in addition to the grammatical negative within the phrase as for example in:

sccovery made her perfectly speechless.	She could not even thank him. She could
h a degree of earnest admiration, which	she could not be insensible of. She was
o love, had found he did not love her.	She could not endure the idea of treache
st do. She must not appear at first.	She could not stand it. I have been thi
or she saw nothing to excite distrust.	She could not imagine a man more exact

This identification of the main collocates of the phrase now allows us to support the hypotheses on the novel's content and structure formulated above. Again, these conclusions are based on the assumption that the frequency of an item and its significance in the data correlate.

The dominance of psychological words as collocates of the novel's most frequent phrase indicates, as literary critics have frequently discussed, that the characters' inner lives are one of the major foci of the novel. While the focus on the characters is indicated by the recurrent personal pronouns, the psychological words, which in most cases have *she* as their agent, hint at a description of their agents' mental state. In most concordance lines (31 of the 41), the agent of the psychological word is identical with the agent of *she could not*. While Anne Elliot, the protagonist, is the agent in 45 of the 54 occurrences of the phrase, she is also the agent for 32 of the 41 psychological words co-occurring with the phrase. We can therefore conclude that Anne's emotions and reflections are dominant in the novel. They shape the reader's perception of events since her perspective is the most frequent.

Looking at the concordance lines in still greater detail, the main topic of Anne's reflections can also be inferred: Captain Wentworth. In 20 out of the 32 occurrences in which Anne is the agent of the psychological words, Captain Wentworth is the object of her reflections. This shows that the relationship between Anne and Captain Wentworth is a major topic in the novel, also presented from Anne's point of view.

While these conclusions on *Persuasion* are neither new nor original and have been frequently commented on, corpus linguistics can offer insights into the text that go further than that. The fact that Anne's preoccupation with Captain Wentworth frequently co-occurs with the negation *not* in *she could not* indicates that Anne's reflections on Captain Wentworth are either restricted for some reason or are not satisfactory for her. These rather vague conclusions will be further discussed and substantiated with additional data later in the analysis.

The collocation of *she could not* with expressions referring to the reception and production of language also hint at the novel's focus on the relationship between its protagonists, with the novel describing their inner lives, thoughts and interactions. The fact that *she could not* contains a grammatical negative suggests that not only Anne's reflections on Captain Wentworth are often unfavourable, but that also communication between them is often unsuccessful. This is (1) indicated by the co-occurrence of the negative *not* within the 3-gram with expressions describing the reception and production of language and (2) by the explicit negation of these communicative events in 12 of their 20 occurrences. Examples of the explicit negation of communication are the following concordance lines:

, when her spirits had nearly failed.
not doing it so well as Lady Russell.
umstances, recollections and feelings,
scovey made her perfectly speechless.
enant." Anne could listen no longer;

She could not call herself an invalid no
She could not speak the name, and look s
she could not hear that Captain Wentwort
She could not even thank him. She could
she could not even have told how the po

Again, Anne is the agent in 10 of the 12 explicitly negated attempts at communication and Captain Wentworth their object in 8 of these 10 cases: Anne's emotions, thoughts and attempts at communication are focused on Captain Wentworth. Their unsuccessful nature is characteristic of Anne's current relationship with him, which is dominated by a lack of communication concerning past and present events and feelings. This is enforced by social conventions "because she is a woman and because it was she who had originally broken off the engagement" (Hardy 1984: 112). Anne's emotions towards Captain Wentworth are therefore characterized by insecurity concerning his feelings and her inability, partly enforced by social conventions, to act according to her feelings.

However, it is not only Anne being female and the recognition of social conventions that trigger her inability to communicate and to act. An analysis of the agents and objects of *she could not* shows that Captain Wentworth himself prompts this inability. In 28 occurrences of *she could not*, Anne is the agent and Captain Wentworth the object. Once Captain Wentworth is talking about the inability of another character, Henrietta Musgrove, but remains capable of action himself. We therefore have 29 occurrences of the phrase where Captain Wentworth is capable of action but other characters who function as either agent or object of the phrase are unable to act. This is especially remarkable since Captain Wentworth is either agent or object of the phrase in only 30 of its 54 occurrences. He therefore triggers an inability in the other characters in nearly one hundred percent (29 out of 31 occurrences) of the relevant cases. This shows that his mere presence has an impact on the other characters' thoughts and actions. This confirms linguistically that he is "a fine young man, with a great deal of intelligence, spirit and brilliancy" (p. 24) who attracts women of all ages. Since all the characters affected by him in this way are female, their inability to act functions as indicators that they are attracted to Captain Wentworth. This will be demonstrated in detail with regard to Anne in the following section.

The fact that Captain Wentworth triggers an inability to act and communicate with the female characters is particularly noticeable with regard to Anne. The figures presented above have revealed a close link between Anne and Captain Wentworth. But while Anne is portrayed as unable to communicate and to act in his presence, this is not the case when he is absent. Throughout the novel,

she is presented as capable and active. This is exemplified when she suggests drastic savings after her father's bankruptcy (p. 10f.), when she is visiting her neighbours at Kellynch for her farewell without the rest of her family (p. 37), and when she is nursing her nephew Charles after an accident (p. 51ff.). Her portrayal as unable to act and to communicate is restricted only to situations in which Captain Wentworth is present and not emotionally involved or when her family or Lady Russel are present. When he seems to be emotionally involved (it later turns out that he was not), she is capable of action also in his presence. This is the case after Louisa's accident when he "seemed to look to her [Anne] for directions" (p. 110), when Captain Wentworth leaves a concert at which Anne is present because he is jealous of Mr Elliot (p. 189f.), or when jealousy of Mr Elliot prevents him from speaking to Anne and she ensures that he hears her talking about women's constancy of romantic feelings (p. 233ff.). Otherwise Anne is unable to act and to communicate when meeting Captain Wentworth. This is the case for example when "she could not shake him [her nephew Charles] off" (p. 79) and Captain Wentworth takes her nephew off her back. Her inability to speak and to act in his presence, as exemplified above, is a linguistic hint that she is still in love with him. Her ability to think and to act seems to be gone, and even though today this is a rather clichéd way of representing someone in love, we nevertheless have to suppose that Anne conforms to the female ideal of her time by being weak and in need of help and protection. The fact that Captain Wentworth himself triggers this inability in the female characters only confirms his desirability in their eyes.

The frequent occurrence of additional grammatical negatives within the collocational span of *she could not* serves a double function: they portray the characters as unable to act and contribute to the overall atmosphere of the novel. The latter is supported by a general tendency toward negatively connotated terms in *Persuasion*. As mentioned at the beginning of the analysis, the 16 most frequent 3-grams of the novel contain five grammatical negatives and further grammatical negatives are collocates of the phrases (as demonstrated for the phrase *she could not*). In addition, several negatively connotated words appear within the collocational span of *she could not*. Examples for negatively connotated collocates of *she could not* occur in the following concordance lines:

y that, busy as she was about Charles,	she could not shake him off. She spoke
The worst is over!" Mary talked, but	she could not attend. She had seen him.
grove had not a word to say in dissent;	she could not accuse herself of having
rying, but very painful agitation,	as she could not recover from, till enabled
, when her spirits had nearly failed.	She could not call herself an invalid no

This dominance of grammatical negations and negatively connotated words as collocates of the most frequent phrase contributes to making *Persuasion* what literary critics have called “the most wistfully melancholy of the [Austen’s] novels” (Hardwick 1973: 208). This theme of melancholy has been further expanded by literary critics who add that “*Persuasion* explores the passing of time, the fading of beauty and the failure of persons of influence to learn from the lessons time teaches” (Gill & Gregory 2003: 354). Both these comments reflect the sombre and melancholy atmosphere of the novel which is created and supported by the frequent use of grammatical negatives and negatively connotated words as exemplified above.

The fourth major class of collocates of *she could not* are markers of time. They contribute to the narrative structure of the novel since a temporal sequence is an essential requirement for a narrative. This has been emphasised by both E.M. Forster (1927) and Labov (1972) in their famous definitions of a narrative:

A plot is also a narrative of events, the emphasis falling on causality. ‘The king died and then the queen died’ is a story. ‘The king died, and then the queen died of grief’ is a plot. (Forster 1927: 60)

and

We define a narrative as one method of recapitulating past experience by matching the verbal sequence of events which (it is inferred) actually occurred. (Labov 1972: 359f.)

The co-occurrence of the novel’s most frequent phrase with expressions of time therefore confirms my earlier classification of 3-grams containing auxiliary verbs as introducing “facts, possibly as part of a narrative”.

The collocation of *she could not* with markers of time fulfils a second thematic function. It contributes to one of the main themes of the novel, the passing of time and the consequent fading of beauty. This theme has been identified by literary critics such as Duckworth (1986). Their intuitive perception of the importance of time in the novel is now supported by linguistic evidence which also closely links the theme with the protagonist Anne Elliot. Since she is the agent in most of the sentences containing *she could not*, the emphasis of the sentences is on her. Therefore, the collocation of the phrase with expressions referring to the passing of time create a connection between her and the theme. The co-occurrence of references to time with grammatical negatives and negatively connotated words indicates that this development is negative for her. Already at

the very beginning of the novel, the reader learns that “[a] few years before Anne Elliot had been a very pretty girl, but her bloom had vanished early” and that “now (...) she was faded and thin” (p. 4). This explicitly marks the connection between the protagonist and time. Their continued connection throughout the novel is mostly implicit. While frequent reference is made to Anne’s poor looks and their improvement in the course of the novel, the intuitive focus in the novel is rather on the fictional present and not past or future. But the linguistic connection between Anne and the past reminds the reader implicitly that her loss of beauty seems to have coincided (the reader can never be sure of that) with her loss of Captain Wentworth eight years ago and that only their renewed acquaintance seems to have brought it back. The references along the chronological plotline therefore fulfil important functions in terms of the content of the novel and its logical development.

The fact that references to time co-occur with grammatical negatives and negatively connotated words creates the sombre atmosphere which pervades the novel. Since the markers of a melancholic atmosphere co-occur with the novel’s most frequent phrase, they are also linked to its agents. We therefore have a direct linguistic link between Anne, the most frequent agent of the phrase, and the melancholic atmosphere of the novel. This allows us to draw conclusions on her state of mind which is presumably also sombre and melancholic. While this atmosphere is perceptible in much of the novel, its existence has been frequently attributed to the time of year when the novel is set and to the miserable life the protagonist leads with her family. The fact that it is part of the language and closely linked to a single character has been frequently overlooked by critics of the novel.

The analysis of the most frequent phrase in *Persuasion* has shown that main lines of the plot and the novel’s atmosphere can be inferred from a linguistic analysis. Linguistic evidence has been presented for thematic and structural features: the relationship between protagonists and the narrative structure. This reveals that the relationship between Anne and Captain Wentworth is at the centre of the plot. While this is also intuitively perceptible, for instance the change in Anne from a capable to a helpless woman has, to my knowledge, not been commented on by literary critics.

2.2 *she had been*

Concordance data on the novel’s second most frequent phrase *she had been* (36) also contribute to a deeper understanding of the content and the narrative structure of the novel. Again, the first step is to identify the main collocates of the phrase. These are:

- psychological verbs, nouns, adverbs and adjectives (26) as in:

t had that appearance. It seemed as if	she had been one moment too late; and as
e good principles and instruction which	she had been anxiously giving her daugh
governing principle. Once only, when	she had been grieving over the last ill-
t appreciation or real taste. In music	she had been always used to feel alone i
d, would do her good. He almost wished	she had been prevailed on to come home

- indicators of time (16) in addition to the temporal sequence already indicated by the phrase being past perfect while the novel is written in the past tense:

ngly more guarded, and more cool, than	she had been the night before. He wan
hat they had had a delightful evening.	She had been the only one of the set abs
uced at length an account of the scene	she had been engaged in there, soon afte
all of peculiar attachment to him, but	she had been too dependent on time alone
was very far from easy about it, when	she had been at home a few hours. On go

- grammatical negatives, negative affixes (6) and negatively connotated words:

, the few steps of unnecessary intimacy	she had been gradually led along. She
d thoughtful and not disposed to talk,	she could not avoid hearing many undesir
manners had not suited her own ideas,	she had been too quick in suspecting the
reserve, and would not use her ill; and	she had been particularly fortunate in
been mistaken with regard to both; that	she had been unfairly influenced by app

These collocations closely resemble those extracted for *she could not*. Both phrases collocate with

- psychological verbs, nouns, adverbs or adjectives
- indicators of time
- grammatical negatives and negatively connotated words.

Both phrases therefore support the focus on the characters' inner lives, the narrative structure of the text and the atmosphere of the novel.

The dominance of psychological words in the concordance lines again indicate a focus on mental concepts and processes. While the agent of *she could not* varies, the agent of *she had been* is mostly stable. In 32 of the 36 occurrences of the phrase, Anne is the agent and only nine concordance lines even include names other than Anne's or pronouns other than *she/her*. There is a clear focus on Anne and the mental processes and concepts referred to within the colloca-

tional span of *she had been* are mainly hers. In combination with the finding that Anne is also the most frequent agent of *she could not*, the analysis of the novel's two most frequent phrases shows that Anne's perspective on events and actions is dominant in the novel. Therefore the reader mostly receives her views on people, events and places.

Markers of time are frequent collocates of *she had been*. This fulfils two functions within the novel. First, the explicit mentioning of time, and especially its passing, establishes the chronological sequence of events in the novel. This is made explicit by markers of time such as

- soon after
- one moment too late
- before
- quite long enough.

This is not surprising since, as mentioned above, a sequence of events is a constitutive factor of a narrative and therefore a structural feature of a novel. But their frequent occurrence in the text as collocates of the novel's most frequent phrases makes them particularly dominant in *Persuasion*.

Another interesting feature of time markers collocating with *she had been* is that they are mostly deictic expressions, and therefore interpretable only from their immediate context within the novel. This ensures that the chronological development of the plot is consistent. It relies on internal factors for its plausibility. In addition, historical dates are referred to at the beginning of the novel, when the Elliot's entry in the Baronetcy is cited (p. 1), and in the course of Sir Walter's search for tenants, the current year 1814 is given. This fixed time frame has been remarked on by literary critics as singular for Austen's novels (Sutherland and Le Faye 2005; Modert 1986). Lodge (1986) also mentions the importance of time in the novel:

[E]ven in *Persuasion*, by general consent the most 'poetic' or 'romantic' of the novels [by Austen], the seasonal symbolism that attaches to the heroine's progress from an autumnal mood of resignation to a joyful 'second spring of youth and beauty' [p. 121 Penguin Edition] arises metonymically out of the actual seasonal span of the action (172).

It appears that literary critics have focussed on the exact historical and seasonal dating in *Persuasion*, but not on the fact that minor events are mainly linked by deictic expressions and thereby form a network of references and cohesive links.

This is the case despite the fact that it is mainly these deictic time references which indicate the passing of time between the main dates such as Michaelmas and Christmas. They therefore have great structural importance for the novel as a sequence of events.

The second function which markers of time fulfil is their contribution to the critics' perception of an atmosphere of melancholy and loss in the novel. This perception is based on the collocation of the phrase with grammatical negations and negatively connotated words as shown above. Further contributors to this perception are the mentions of the fading of beauty (mostly Anne's as exemplified above) and the death of characters (we hear about the death of eight (!) characters in the novel). The collocation of the novel's two most frequent phrases with negatives linguistically explains why for instance Jenkyns (2004: 175) calls *Persuasion* "a novel about lonely endurance" and why Auerbach (2004: 233) says that

[m]any Austen readers find it impossible to read this final novel without being poignantly conscious of Austen's own declining health and the fact that she never lived to see *Persuasion* published.

These collocations of the novel's most frequent phrases ensure that the atmosphere of the entire novel is perceived as sombre and melancholic.

Explicit markers of time are supported in their function of marking the passing of time by the phrase *she had been* itself. Since the novel is mainly narrated in the past tense, the occurrence of the past perfect, especially in the second most frequent phrase of the novel, is in itself a marker of time. The past perfect refers to a point in time before the fictional present. The temporal structure of the novel is therefore based on both explicit markers of time and the implicit sequence of tenses in the narrative.

3 Conclusion

The analyses above can be summarized as follows:

- Explicit and implicit markers of time as collocates of the novel's two most frequent phrases
 - denote the temporal relationships that are a constituent part of the novel's narrative;

- create, in connection with the collocation of the phrases with grammatical negatives and negatively connotated words, the novel's melancholy atmosphere which has been noted by literary critics.
- The dominance of psychological words as collocates of both 3-grams indicates the novel's focus on the characters' inner lives. The fact that Anne is mostly their agent shows that her thoughts and perceptions are at the centre of the novel. This conclusion is supported by the dominance of free indirect thought in the concordance lines of *she could not*.
- Anne's relationship with Captain Wentworth, the frequent object of the phrases and sentences containing *she could not*, is characterised by a lack of communication. In addition, Anne, who is usually portrayed as capable in the novel, becomes helpless and in need of protection from Captain Wentworth in his presence. This relationship is reversed only when Captain Wentworth is emotionally involved in an action. The inability of the female characters to act in his presence is triggered by Captain Wentworth himself.

These observations all concern developments and themes of the entire novel. The phrases and their collocates function as cohesive ties across the entire text. This detailed look at the text contributes new insights into a text that has been analysed for nearly two hundred years. This is, for example, the case with Anne's ability or inability to act depending on whether Captain Wentworth is present or not and of Captain Wentworth being the trigger for this behaviour. Having a close look at the language of a (literary) text, especially by way of corpus techniques which analyse the language of the entire novel, is therefore a valuable contribution not only to the literary criticism of *Persuasion*, but also to linguistics in so far as it shows that even in a well analysed text new features can be found by looking closely at its language.

Acknowledgement

I am grateful to Michael Stubbs for comments on earlier versions of this paper.

References

- Auerbach, Emily. 2004. *Searching for Jane Austen*. Madison/London: The University of Wisconsin Press.
- Austen, Jane. 1818. *Persuasion*. London: Penguin. 1994.

- Burrows, J.F. 1987. *Computation into criticism. A study of Jane Austen's novels and an experiment in method*. Oxford: Clarendon.
- Duckworth, Alistair M. 1986. Nature. In J. D. Grey (ed.). *The Jane Austen handbook*, 317–319. London: Athlone.
- Fletcher, William H. 2002. *kfNgram*. [Computer Software accessed 25 October 2005] <http://miniappolis.com/KWiCFinder/kfNgramHelp.html>
- Forster, E. M. 1927. *Aspects of the novel and related writings*. Cambridge: Edward Arnold. 1974.
- Gill, Richard and Susan Gregory. 2003. *Mastering the novels of Jane Austen*. Basingstoke & New York: Palgrave Macmillan.
- Hardwick, Michael. 1973. *The Osprey guide to Jane Austen*. Reading: Osprey.
- Hardy, John. 1984. *Jane Austen's heroines: Intimacy in human relations*. London & New York: Routledge & Kegan Paul.
- Jenkyns, Richard. 2004. *A fine brush on ivory: An appreciation of Jane Austen*. Oxford: Oxford University Press.
- Labov, William. 1972. *Language in the inner city: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
- Lodge, David. 1986. Jane Austen's novels: Form and structure. In J. D. Grey (ed.). *The Jane Austen handbook with a dictionary of Jane Austen's life and works*, 165–178. London: Athlone.
- Modert, Jo. 1986. Chronology within the novels. In J. D. Grey (ed.). *The Jane Austen handbook with a dictionary of Jane Austen's life and works*, 53–59. London: Athlone.
- Naciscione, Anita. 2001. *Phraseological units in discourse: Towards applied stylistics*. Riga: Latvian Academy of Culture.
- Partington, Alan. 2004. "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9(1): 131–156.
- Scott, Mike. 1999. *WordSmith Tools*. 3.0. Oxford: Oxford University Press. [Computer Software]
- Sinclair, John. 1996. The search for units of meaning. *Textus* 9: 75–106.
- Sinclair, John. 2004. New evidence, new priorities, new attitudes. In J. Sinclair (ed.). *How to use corpora in language teaching*, 271–99. Amsterdam: Benjamins.
- Stubbs, Michael. 2005a. Conrad in the computer: Examples of quantitative stylistic methods. *Language and Literature* 1(5): 5–24.

- Stubbs, Michael. 2005b. The most natural thing in the world: Quantitative data on multi-word sequences in English. Conference presentation at *Phraseology 2005, Louvain*.
- Sutherland, John and Deirdre Le Faye. 2005. *So you think you know Jane Austen? A literary quizbook*. Oxford: Oxford University Press.
- Tabata, Tomoji. 2002. Investigating stylistic variation in Dickens through correspondence analysis of word-class distribution. In T. Saito, J. Nakamura, S. Yamazaki (eds.). *English corpus linguistics in Japan*, 165–182. Amsterdam & New York: Rodopi.
- Teubert, Wolfgang. 2005. My version of corpus linguistics. *International Journal of Corpus Linguistics* 10(1): 1–13.