Università degli Studi di Padova

Padua Research Archive - Institutional Repository

"Pride and Prejudice" on the Page and on the Screen: Literary Narrative, Literary Dialogue and Film Dialogue

(Article begins on next page)

Pride and Prejudice on the Page and on the Screen: Literary Narrative, Literary Dialogue and Film Dialogue[1]

*Francesca Bianchi**, University of Salento and Sara Gesuato, University of Padua*

Abstract
This study explores the similarities and differences in content between the dialogic and narrative parts in Jane Austen's *Pride and Prejudice*, and between the novel's dialogues and those in its 1940 and 2005 film adaptations. These four datasets were semantically tagged and compared to one another by using qualitative and quantitative methods. The findings show how, in covering conceptual areas largely complementary to those of the narrative, the dialogues in the novel perform various communicative functions. The investigation also points to how dialogues are adapted to the semiotic needs and goals of its film adaptations.

Keywords: Austen; corpus-informed analysis of literature; film adaptations of novels; dialogues; subtitles; semantic tagging

1. Introduction

Broadly speaking, a novel is the verbal representation of non-verbal phenomena and circumstances, a report of imaginary events and a description of fictitious situations, conveyed in the fabric of a text. This verbal narrative conveys human motivation and values in "propositions which attempt to develop perception", throwing "them out onto the external world, elaborating a world out of a story" (Dudley 1984: 101). Its linguistic code imposes constraints on the rendering of its content: for example, episodes and entities may or may not be concurrent or co-present, respectively, in the fictional world, but in the text they can only be introduced sequentially (McFarlane 1996: 27). Moreover, their relative prominence is reflected or manipulated in the foregrounded or

---

backgrounded syntactic structures of the sentences, and the ways in which they are defined, described and classified co-vary with the author's specific lexico-stylistic choices. That is, a novel presents its narrative through its wording, which is never conceptually neutral, but rather shapes content. As Leech and Short (2007: 100-106) illustrate, in a fictional work, even minute variants of a given sentence, from which what is apparently the "same" event can be inferred, have stylistic values. That is, semantic, syntactic and grapho-phonological variations convey different senses, a phenomenon of which writers are acutely aware (pp. 107-108).

A film is also a non-neutral representation of events and situations, one out of many possible alternative versions of the same circumstances. Unlike novels, films work "from perception toward signification, from external facts to interior motivations and consequences, from the givenness of a world to the meaning of a story cut out of that world" (Dudley 1984: 101). A film's signs sensuously and perceptually represent what in the text is rendered conceptually and symbolically (McFarlane 1996: 27). Furthermore, unlike novels, films are multi-modal interpretations of stories. They rely on various semiotic codes: sounds and music, still and moving images, and of course words, mostly in dialogues (McFarlane 1996: 26, 29). These codes interact, often simultaneously (McFarlane 1996: 27), weaving a variegated and multi-dimensional narrative fabric.[2] As a result of this, films communicate at several levels (Boozer 2008: 1).

There is an intricate interplay between verbal and visual modes of expression in films. Film dialogue informs viewers about the theme, plot, characters and circumstances of a story (Zettl 2008: 340). It guides and supports the audience in the interpretation of the storyline and in understanding relationships between characters, including characters' attitude, mood and personality (Bianchi 2015: 242). Likewise, filmic images define the background and setting in which action takes place, and render characters physically concrete. They also provide sociocultural references through the characters' appearance and demeanour. Besides aesthetically affecting viewers (Zettl 2008), images act on the conscious and subconscious mind, conveying spatiotemporal

---

[2] The interplay of these elements in video production and broadcast is described in Zettl (2008).

and/or sociocultural information and recalling additional images, knowledge, collective and individual memories to viewers' minds (Bianchi 2015: 240). Film adaptations of literary works are therefore complex cultural constructs.[3] They require multimodal equivalents of the narrative exposition, which includes metaphors and interior character observations (Boozer 2008: 7).

Research generally takes into consideration the various components of film adaptations and how these are creatively employed by film makers. The most typical question addressed is how faithfully it represents the content and sequencing of the original text (Dudley 1984: 96-97; McFarlane 1996: 8). However, attention has also been drawn to its dynamic intertextuality (e.g. its literary influences, cultural milieu and context; Aragay and López 2005: 203; Cobb 2008: 281-283), its authenticity (i.e. its rendering of the essence or spirit of the novel; Dudley 1984: 100) and its use of the source text to serve its own ideology (Orr 1984, quoted in McFarlane 1996: 10). Scholars have therefore identified several modes of relation between literary texts and their film adaptations: borrowing, intersection and fidelity of transformation regarding the letter or the spirit of the text (Dudley 1984: 98-100; McFarlane 1996: 8-9); literal reading, general correspondence and distant referencing (Boozer 2008: 9); transposition, commentary and analogy (Wagner 1975: 222, quoted in McFarlane 1996: 10-11); fidelity, reinterpretation and using the source material as the occasion for original work (Klein and Parker 1981: 9-10, quoted in McFarlane 1996: 13); and transfer versus adaptation (McFarlane 1996: 12). Other issues discussed are the aesthetic value and cultural meanings of the literary and filmic texts in their temporal contexts (Boozer 2008: 10) and the collaborative nature of filmmaking (Cobb 2008: 284-285, 289; Parrill 2002: 50), which "undercuts the idea of the director of the adaptation as the translator of the novel" (Cobb 2008: 285).

Specific issues that have been investigated are: the rewriting of the novel "in dramatic form" (Stovel 2011) and the application of visual techniques practised by writers (McFarlane 1996: 6); the treatment of time (Parrill 2002: 53) with the possible omission (Parrill 2002: 59) or

---

[3] It is beyond the scope of this paper to review film adaptation research in general. Relevant works include McFarlane (1996), Lu (1999) and Scholz (2013). Parrill (2002) provides an analysis focused on film adaptations of Austen's novels.

addition of scenes (Aragay and López 2005: 211); the use of camera angle techniques to convey various perspectives (Stovel 2007; McFarlane 1996: 17-20) and especially to render characters' gaze (Aragay and López 2005: 206); the bridging of the historical (i.e. linguistic and cultural) gap between the original text and the new audience with the consequent (de-)emphasization of given topics, notions or phenomena (Stovel 2011); the treatment of sound in its informational and orientational function (Zettl 2008: 351), especially for rendering the expressiveness and indirect communication which, in novels, relies implicitly on the body (Hudelet 2005: 176) and for expressing the "materiality of the voice" (Hudelet 2005: 177); the actors' various interpretations of the characters' roles; and the casting of actors against type (Stovel 2011).

Despite their specific characteristics, a novel and its film adaptation are comparable artefacts in that their signs and units elicit a chain reaction of relations that give rise to an unfolding structure, the elaboration of a fictional world (Cohen 1979, quoted in Dudley 1984: 103). A novel and its film adaption, therefore, share narrativity (McFarlane 1996: 12), that is, the use of consecutive signs, whether verbal or visual.

Nonetheless, narration is carried out through partially different means in the two types of fictional works: verbal only in the novel, and verbal and visual in the film (Stovel 2011). More specifically, in novels, the telling of the story is in the hands of the narrator and the characters. Their voices advance the plot and illustrate the topics chosen by the author through "different layers of consciousness" (Mahlberg et al. 2016: 436). In films, instead, it is the characters' voices that dominate; a narrator—in the form of a voice-over or written contextualising information—is generally absent or quantitatively secondary to dialogues (McFarlane 1996: 16), while it is the non-verbal components of the film that support the characters' dialogues in constructing a narratorial viewpoint.

While the interplay of narration and dialogue in novels has received the attention of scholars, the relationship between dialogues in film transpositions and dialogues and narration in the source novel is still an under-investigated topic. This is of interest because, given that novels and films share only one formal component, namely the characters' dialogues, and that, as stated by Stovel (2011), "[a]dapting a novel for

the screen involves translating a purely linguistic medium to a primarily visual one", it is difficult to predict what in films will be rendered through words rather than images. In this study we wish to partly explore this issue by comparing and contrasting the most frequent notions expressed in a novel, Jane Austen's *Pride and Prejudice*, and two of its film transpositions—the 1940 film directed by Robert Z. Leonard, and the 2005 film directed by Joe Wright. In particular, we address these research questions: 1) What conceptual areas are covered in the dialogic versus the non-dialogic (i.e. narrative) parts of the novel? 2) What conceptual areas are covered in the dialogues in the films? 3) How different are the novel's dialogues and the films' dialogues in terms of the conceptual areas they cover? To these ends, we examine the lexical-semantic fields characterising the dialogic and non-dialogic parts of the novel, taken separately, and compare them to those in the film dialogues, as represented in the film subtitles. Our analyses adopt a corpus-informed approach combining quantitative and qualitative methods (see Section 3), and involve the automatic semantic tagging of the texts, the automatic extraction of their key domains and a close reading of the concordance lines in which they are instantiated.

In the rest of the paper we contextualise our study (Section 2), we report on our data collection and analysis procedures (Section 3), and we present and discuss our findings (Section 4), drawing conclusions from them (Section 5).

2. Contextualisation

Studies of literature can be examined from a critical or a corpus-informed perspective.

2.1 Critical studies

As is well-known, *Pride and Prejudice* is a romantic novel. It traces the emotional development of its heroine, Elizabeth Bennet, who gradually, and painfully, learns the difference between appearances and reality in her judgement of people's character and behaviour. However, the novel's appeal lies mostly in its portrayal of the sociocultural values and practices of the British Regency period, which Austen subtly makes fun of, from the very first sentence of the novel (Aragay and López 2005:

207; Parrill 2002: 45). That is, the author shows how members of the upper classes of the time are guided and constrained by cultural conventions, social norms, upbringing and class membership in their thoughts, behaviour and especially in their interactions and relationships. In so doing, Austen also shows their affectations, contradictions and biases, but with benevolent humour.

Literary scholars have examined in depth and interpreted at length the content and style of fictional works. In the case of *Pride and Prejudice*, they have identified the following as the crucial themes of the novel: love and marriage, money and social class differences, family relationships and individual growth, as experienced and understood in the microcosm of the late eighteenth century English gentry (for a succinct review, see Bianchi and Gesuato 2019).

Besides examining the thematic focus of fictional works, literary critics have also long drawn attention to the complementary role played in novels by the narrator's versus the characters' utterances, through which the novel communicates its values and worldview. In particular, it has been pointed out that dialogue and narrative in a novel may represent two poles of a continuum comprising several modes of expression. For example, in his analysis of Jane Austen's *Emma*, Hough (1970: 203-205) distinguishes five types of discourse: 1) the authorial voice, which occurs in reflective passages directly addressed to readers and establishes some complicity with them; 2) the objective narrative, found in passages where the narrator presents factual information about the story and the characters in a neutral way; 3) coloured narrative, where the narrator offers reflections or observations presented from a given character's point of view; 4) free indirect style, when a character's mode of expression is presented embedded in the narrative; and 5) direct speech. Hough observes that it is in the fluid and frequent shifts between different modes of expression that Austen conveys irony, presenting two viewpoints simultaneously (pp. 210). The same strategy also enables her to inconspicuously present accurate, plausible information, "on which a correct judgement could have been made" (p. 213), while letting the reader temporarily "go astray". The author also notes the strong similarity between the language of the narrative sections and the characters' speech in the novel. More specifically, the objective narrative in *Emma* is authoritative, adopting the form of "decent educated discourse" (p. 207), "abstract, evaluative and formally correct" (p. 208),

thus giving an impression of "majestic impersonality" (p. 209). Hough argues how in *Emma*, the characters display the language of judgement, "thoughtful, well-ordered, analytical and generalizing" (p. 220), "unconcerned with trivia or material circumstances" (p. 221) and characterised by abstract evaluative vocabulary (p. 218), while deviations from this norm are a sign of the characters' social, emotional and/or intellectual inferiority, and of the narrator's disapproval of them (pp. 217-218, 220-221). Overall, Hough's account of *Emma* indicates that the dialogic and narrative planes of expression in Austen's fictional prose provide modulated and mutually reinforcing representations of her fictional worldview.

Several of Austen's works have been adapted for the screen, and scholars have investigated these adaptations too. Studies exploring the content of these works have identified the following themes: romance and female autonomy in *Pride and Prejudice* (Aragay and López 2005: 201, 204, 205); the notions of femininity and masculinity (Aragay and López 2005: 201; Wakefield 2007); the intertwining of public and private life (Scholz 2013: 124-127); the importance of money and social standing for women as achieved through marriage (Scholz 2013: 129; 133-134); the representation of family and social life (Scholz 2013: 134); social class divisions (Scholz 2013: 140); the staging of "dancing as both a metaphor and a model for marriage" (Stovel 2007) and as the dramatisation of courtship (Stovel 2006: 196); and the partnership-rivalry between, and the individuality of, the text and the film adaptation (Snyder 2011: 138-139, 144-145, 151; McFarlane 1996: 6-7).[4]

2.2 Corpus studies

Corpus-linguistic methods have already been used in the analysis of literary works and film dialogues, including Austen's novels and their

---

[4] Surprisingly, however, the screenplay of films is often disregarded as an object of investigation, even though it is "the film's narrative springboard" which "guides screen choices for story structure, characterization, motifs, themes, and genre" (Boozer 2008: 4). Possible reasons are that it undergoes multiple revisions by various people; that it has no independent existence as a finished product; and that it is not meant for the audience (Boozer 2008: 2, 4).

film adaptations.[5] Some investigations have explored the aboutness of texts. These studies have confirmed previous critical interpretations of works of fiction, on the basis of quantitative lexical-phraseological data. For example, Fischer-Starcke's (2010) analysis of the lexis of *Pride and Prejudice* shows that the novel focuses on the themes of family and interpersonal relationships, women and men, the military, mental concepts and emotions, love and courtship, and communication.

Other studies have focused on film dialogue. For example, Bianchi's (2016) automated and manual semantic analysis of the 2005 film *Pride & Prejudice* reveals that this is focused on such themes as positive and negative feelings (e.g. romantic love, family affection, pride), family and interpersonal relationships, social life events (e.g. dancing, talking, visiting), and social norms (e.g. manners, obligations). Similarly, Bianchi and Gesuato (2019) evidenced how the 1940 and 2005 films share themes related to the plot, namely feelings, interpersonal relationships and sensitivity to social values, but also how the earlier film is more focused on the topic of merits and demerits, while the more recent one on that of individual relationships.

Nowadays, corpus linguistic tools enable scholars to selectively focus their attention on the dialogic versus the narrative planes of fictional discourse. For example, Mahlberg et al. (2016) showed that the dialogic versus narrative components of a novel may have distinct stylistic properties (e.g. Mahlberg et al. 2016), with speech effectively differentiating characters (p. 452) and narrative parts describing the fictional world (p. 454).

---

[5] Corpus-linguistic analyses often involve identifying the most frequent words or word combinations in texts and tracing their dispersion across chapters, or extracting from texts the words that are unusually frequent compared to other texts and grouping them into semantic-fields. When applied to works of fiction, such an approach may for example help to identify the topics discussed and the points in the text where they are first introduced, the stylistic/register choices peculiar to given characters, the authorial point of view as evidenced in evaluative terms, or the stylistic distinctiveness of an author or genre. The advantages of a corpus-linguistic approach to the study of literary texts are that it provides "an empirical basis for interpretation" (Stubbs 2001: 125)—thanks to the identification of patterns (Stubbs 2001: 140)—and that its validity can be checked and refined through replication (Stubbs 2001: 123).

To our knowledge, however, no corpus linguistic study has investigated the possibly distinctive contribution made by characters' speech—as opposed to the narrator's discourse—to the thematic make-up of a fictional story, let alone whether and how the topics of characters' speech change when transferred from the page to the screen. To partially fill this gap, we set out to employ corpus methods to identify and describe the conceptual areas covered in the dialogic and non-dialogic components of *Pride and Prejudice*, and to compare the dialogues in the novel to the characters' dialogues in its screen adaptations.

3. Materials and Methods

We analysed and compared Jane Austen's novel *Pride and Prejudice* and its 1940 and 2005 film versions. The 1940 film is often classified as a screwball comedy (Stovel 2013; Parrill 2002). This is a Hollywood genre that was very popular from the early 1930s to the mid-1940s. It is characterised, among other things, by its emphasis on funny, farcical situations and characters (Gehring 2002) and by fast-talking, witty repartee (Otnes and Pleck 2003). The 2005 film, on the other hand, is a British romantic comedy (Chan 2007; Martin 2007; Woodworth 2007) and also a heritage film insofar as it shows spectacular landscapes and formal events in ornate interiors; however, it does not totally adhere to the heritage tradition, considering that it also represents the realistic details of agricultural economy (Dole 2007). The two film transpositions therefore belong to different periods, genres and cultural traditions (US versus UK).

We first created the corpora necessary for the linguistic analysis. We accessed the full text of the novel *Pride and Prejudice* through CLiC, a web corpus concordancing tool specifically designed for the analysis of literary texts (Mahlberg et al. 2016; clic.bham.ac.uk), and which contains a wide selection of 19th century novels. The texts in the CLiC database are annotated so as to distinguish the characters' utterances—i.e. the dialogues, labelled *quotes* on the CLiC platform—from other parts of the text, i.e. the non-dialogues, labelled *non-quotes* on the CLiC platform. The non-dialogues include both pure narrative segments and suspensions, the latter being those narrative segments that interrupt characters' quoted speech. CLiC users can search the full text of a novel,

or specific parts of it. In the current study, we used the CLiC database to extract the dialogues and non-dialogues from *Pride and Prejudice*, thus creating two corpora, which we will here call *novel_d* (50,348 words) and *novel_nd* (65,080 words), respectively. From the two films, we instead collected the dialogues by extracting the English subtitles from their official DVDs. This led to the creation of the *1940 corpus* (15,220 words) and the *2005 corpus* (14,992 words).

Next, we carried out an automated analysis of the four corpora by using Wmatrix (Rayson 2003). This is a corpus analysis and concordancing tool that performs automatic Part-of-Speech tagging and semantic tagging (i.e. classification of terms according to semantic fields). Semantic tagging brings two major advantages over other corpus-analysis tools. One is that it groups all inflected forms of a given word together within the same semantic field (Archer and Rayson 2004), and the other is that each semantic field groups together all the words that are relevant to the same semantic space, including low-frequency words, which would most likely be overlooked in searches carried out with other methods (Rayson 2008: 543). In Wmatrix, semantic tagging is executed by the Semantic Analysis System developed at the Lancaster University Centre for Computer Corpus Research on Language, called *USAS*. Its lexical processing is based on semantic lexicons and frequency statistics. This means that the attribution of a word to a semantic field is a decontextualized process, with all the limits that follow (see below). However, it is enhanced by the fact that Wmatrix recognises not only single terms, but also multi-word expressions (e.g. *Caroline Bingley*, *having to*, *in love*, or *took place*). Therefore, when a group of words collectively encode one unit of meaning, they are assigned together to a given semantic field (e.g. *Caroline Bingley*: Z1: Personal names; *having to*: S6+: Strong obligation or necessity).

Wmatrix can also produce keyword lists (i.e. lists of unusually frequent words), key Part-of-Speech lists (i.e. lists of unusually frequent Part-of-Speech tags) and key concept lists (i.e. lists of unusually frequent semantic tags). Key domains (a.k.a. key concepts) are an extension of the keyword notion, and identify unusually frequent semantic areas that emerge in a given corpus when its tagged semantic fields are statistically compared against those of another corpus. The degree of outstandingness (i.e. unexpected prominence or non-prominence) of a given item (e.g. semantic tag) in the corpus under investigation is called *keyness*, and is

established through statistical analysis—chi-square and log-likelihood (LL) being the most frequently used tests for this purpose.[6]

Specifically, we used Wmatrix to semantically tag the four corpora and extract key domains within them. In keeping with previous studies on literary works (e.g. Culpeper, Archer and Rayson 2009) and film dialogues (Bianchi 2016; Bianchi and Gesuato 2019), we used LL as a statistical measure of keyness, and focused on key domains with LL > 15.13 (p < 0.0001 1d.f.). The semantic classification provided in the USAS tagset is not always fully satisfactory, as exemplified in the second column of Table 1. For example, figurative uses of terms are not recognised as such, but rather classified literally (see examples [1], [5] and [9]). Also, a word that potentially encodes more than one sense is automatically assigned to a given semantic field only on the basis of one (possibly the most frequent) of its semantic traits, that is, without consideration for the context in which it occurs, which might require the activation of other semantic traits (see examples [7], [3], [11] and [12]).

For the above-mentioned reasons, we found it necessary to perform a manual analysis of the concordance lines[7] of the terms belonging to the key domains identified in the novel. This involved paying attention to the co-text of the given terms and the situations or events being recounted in larger stretches of text surrounding those terms.[8] Concordance lines were inductively reorganised into larger conceptual areas cutting across key domains.[9] One coder intuitively classified each line after repeated

---

[6] For a quick explanation of these statistics applied to keyness, see for example, the following page by David Brown: http://www.thegrammarlab.com/?p=193.

[7] Concordance lines are lines of text of a pre-defined length which show a given word or phrase in the middle of the text string (i.e. as the so-called *node term/expression*) and some co-text to its left and right. Concordance lines make it possible to read a text "vertically", for a clearer view of the semantic-grammatical relation between the node term/expression and its co-text.

[8] More generally, as the creator of Wmatrix observes, "[c]areful manual analysis of concordances of key words and key domains is obviously required to check for mistagging and poor dispersion of high frequency items" (Rayson 2008: 544).

[9] We use the term *conceptual area* to identify a portion of semantic space that given lexical items are similarly relevant to and which may or may not represent a theme, i.e. a subject for discussion. Therefore, *conceptual area* is our label for the semantic areas we identified through a manual analysis of the automatically extracted key domains.

readings, while the other coder checked the classification. When in disagreement, the co-text and context of the concordance lines were discussed together; when agreement was not reached, the lines were assigned to the category *Other*. We coded one corpus at a time. Whenever new conceptual areas emerged, we revised our previous codings. When a word or multi-word unit appeared to belong to multiple conceptual areas, we assigned it to all of them (e.g. *Jane* is the name of a specific character and was always classified as such; furthermore, the instances of *Jane* used as vocative were also classified as examples of spoken traits).

Some examples of our manual classification are provided in Table 1. Throughout the paper, conceptual areas are reported in small caps.

Finally, we contrasted the four corpora at the level of conceptual areas from a combined qualitative and quantitative perspective. To this aim we considered two factors. The first is the relative weight of a given conceptual area, as expressed by the frequency of its instantiation, i.e. the percentage of the words or multi-word units representing it over the overall number of words in the relevant corpus. The other factor is the number and type of semantic fields contributing to a given conceptual area, as these are indicators of lexical variation. The results of the analyses are reported in the following section.

Table 1. Examples of manual re-classification of concordance lines into conceptual areas

Semantic tag	Concordance line	Conceptual area
I1.1-: Money: Lack	*if that is all, I have a very <u>poor</u> opinion of young men who live in*	JUDGEMENT
I1.1-: Money: Lack	*he is rich, and many others are <u>poor</u>. I speak feelingly. A younger so*	DESCRIPTION OF PEOPLE
E2+: Like	*ever, the happiness of a most <u>beloved</u> sister? I have every reason*	JUDGEMENT
E2+: Like	*are very few of us who do not <u>cherish</u> a feeling of self-complacency*	MENTAL ACTIONS, FEELING AND ATTITUDES
B1: Anatomy and physiology	*hearing Miss Bingley's voice, Elizabeth was <u>roused</u> by receiving from her a cold inquiry*	MENTAL ACTIONS, FEELING AND ATTITUDES

Semantic tag	Concordance line	Conceptual area
B1: Anatomy and physiology	*subsisted between Mr. Darcy and herself. She was <u>roused</u> from her seat*	MATERIAL ACTIONS
Z6: Negative	*a design against a girl who is <u>by no means</u> unprotected or friendless, and*	EPISTEMIC MODALITY
S6+: Strong obligation or necessity	*If I can perceive her regard for him, he <u>must</u> be a simpleton, indeed*	EPISTEMIC MODALITY
A9+: Getting and possession	*offer of marriage in this way, you will never <u>get</u> a husband at all*	MARRIAGE
A9+: Getting and possession	*A glass of wine; shall I <u>get</u> you one? You are very ill*	MATERIAL ACTIONS
X2.2+: Knowledgeable	*George Wickham has treated Mr. Darcy in a most <u>infamous</u> manner*	JUDGEMENT

4. Results and Discussion

In this section we present and comment on the findings of our analysis, starting from the quantitative data. The analysis will show the general thematic and lexical complementarity of the dialogic and narrative parts of the novel.

4.1. Dialogues versus non-dialogues in the novel

The analytical procedure described in Section 3 applied to the *novel_d* (dialogues) and *novel_nd* (non-dialogues) corpora highlighted the shared and unshared prominent conceptual areas. Eight conceptual areas appear in the key domains retrieved from the dialogues only; five conceptual areas characterise non-dialogues only; and nine conceptual areas appear in the key domains of both corpora, though not with the same degree of prominence. These are illustrated and discussed in the sections below. In the tables, the first column lists the various conceptual areas identified; the second column reports a few sample words illustrating the conceptual areas (more detailed instances in the form of sentences are included in the narrative illustrations that follow the tables); and the remaining column/s include(s) the USAS code(s) that contributed to each conceptual area for the given corpus.

4.1.2 Conceptual areas specific to novel_d *only*
The conceptual areas that appear in the dialogues only are listed in Table 2.

Table 2. Conceptual areas specific to the dialogues only

Label	Sample words	Relevant key domains
SPOKEN TRAITS	*I, my, you, yes, no, my dear, are*	Z8, Z4, A3+, E2+, A5.1+
FUTURE EVENTS, INTENTIONS AND HYPOTHESES	*if, will, would*	Z7, T1.1.3, A7+, S6+, A9+
PEOPLE GENERALLY	*people, woman, character*	S2; S2.1
DESCRIPTION OF PEOPLE	*belong, destitute, poor*	A9+, I1.1-
EXPERIENCE	*have*	A9+
DEONTIC MODALITY	*allowed, must*	S7.4+, A5.1++, S6+, A9+
DYNAMIC MODALITY	*can*	A7+
TIME IN GENERAL	*time*	T1

As to be expected, a primary feature of the *novel_d* corpus is the presence of spoken traits (conceptual area SPOKEN TRAITS), recorded by the software in five key domains and exemplified by the following types of words or expressions: personal pronouns (e.g. *I*, *my*, *you*); exclamations (e.g. *oh*; *you see*; *yes*; *no, my dear*); discourse markers (e.g. *Well, he certainly is very agreeable*); verbs indicating present time (e.g. *They are in the same profession*); verbs serving as predicates to pronominal subjects *I, we,* or *you*; and deictic forms, which identify participants to the here-and-now of conversation (e.g. *And I am happy to say*).

The remaining conceptual areas appearing only in the dialogues suggest that, when interacting, characters perform the following actions:
• Talking about future events, intentions and hypotheses (conceptual area FUTURE EVENTS, INTENTIONS AND HYPOTHESES): e.g. *Perhaps he must, if he sees enough of her*; *Elizabeth will soon be the wife of Mr. Darcy*; *Another time, Lizzy, I would not dance with him, if I were you*; *It will be no use to us, if twenty such should come, since you will not visit them*.
• Talking about people and their personality traits (conceptual area PEOPLE GENERALLY): e.g. *there are very few people of whom so much can be said*; *when a woman has five grown-up daughters*; *considering Mr. Collins's character*.

• Describing people (in terms of what they possess or fail to possess) (conceptual area DESCRIPTION OF PEOPLE): e.g. *It is evident that you belong to the first circles*; *for else they will be destitute enough; she is luckily too poor to be an object of prey.*

• Undergoing experiences (conceptual area EXPERIENCE): e.g. *While I can have my mornings to myself, it is enough.*

• Talking about duties, obligations, and abilities (conceptual areas DEONTIC MODALITY and DYNAMIC MODALITY): e.g. *ice every week, and are never allowed to walk home*; *it would be better for the neighbourhood that he*; *I must confess that he did not speak so well of Wickham*; *No, my dear, you had better go on horseback.*

• Referring to time in general (conceptual area TIME IN GENERAL): e.g. *it is the first time we have ever had anything from him; But to be guarded at such a time is very difficult.*

4.1.3 Conceptual areas specific to novel_nd *only*
The few conceptual areas appearing only in the non-dialogues are listed in Table 3.

Table 3. Conceptual areas specific to the non-dialogues only

Label	Sample words	Relevant key domains
TIMELINE	*began, Sunday, breakfast*	T2-, T2+, T1.3, N4, F1
SPATIAL SETTINGS	*dining-room, gates*	H2, H5, F1, M2
CHARACTERS	*Ms Bennet, Jane*	Z1
DESCRIPTION OF ACTIONS AND EVENTS	*quietly, abruptly*	N3.8+, X5.1+, X5.2+, X3.2-
ABSENCE OF EXPECTED THINGS	*absence*	S1.1.3

As to be expected in narrative text, the prominent conceptual areas characterizing *novel_nd* serve the following functions: indicating points in time, periods of time, or events and occasions associated with particular points in time (conceptual area TIMELINE): e.g. *when the first tumult of joy was over, she began to declare that*; *Elizabeth now began to revive*; *On Sunday, after morning service*; *At last she recollected that*; *Breakfast was scarcely over when a servant*; identifying or describing spatial settings (conceptual area SPATIAL SETTINGS): e.g. *in spite of Mrs. Phillips's throwing up the parlour window*; *and she soon passed one of*

the gates into the ground; *Lydia looking out of a <u>dining-room</u> upstairs*; *The garden <u>sloping</u> to the road*); and identifying the various characters (conceptual area CHARACTERS).

The other conceptual areas that are specific to the *novel_nd* corpus show that Austen also uses narrative text to:

• Describe the manner in which actions and events take place (conceptual area DESCRIPTION OF ACTIONS AND EVENTS): e.g. *and therefore, <u>abruptly</u> changing the conversation*; *She listened most <u>attentively</u>*; *The advice was followed <u>readily</u>*; *Elizabeth <u>quietly</u> answered*.

• Report absence of expected qualities or phenomena (conceptual area ABSENCE OF EXPECTED THINGS): e.g. *sometimes it seemed nothing but <u>absence</u> of mind*.

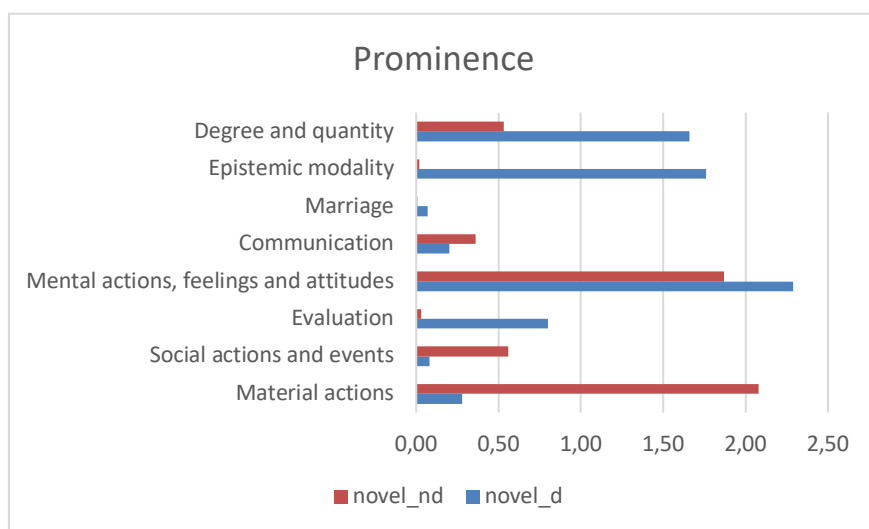*4.1.4 Conceptual areas that appear in* novel_d *and* novel_nd

Table 4 shows the conceptual areas that appear in both corpora. As indicated by the different semantic tags in the third and fourth columns, these conceptual areas are instantiated through different lexical choices in the two corpora.

Table 4. Conceptual areas appearing in both the dialogues and the non-dialogues

Label	Sample words	Relevant key domains in the dialogues	Relevant key domains in the non-dialogues
MATERIAL ACTIONS	*keep, live // advanced, saw*	S7.4+, E2+, A9+	M1, M2, X3.4, B1, X2.4, F1, I3.1, A10+, Q2.2, T2+; Q2.1
SOCIAL ACTIONS AND EVENTS	*known // introduction*	X2.2+, A9+	Q2.1, Q2.2, X3.4, B1, F1, A10+, T2+, M1
EVALUATION	*superior, infamous // lively*	A5.1+, A5.1++, X2.2+, E2+, A13.3	X5.2+
MENTAL ACTIONS, FEELINGS AND ATTITUDES	*fancy, believe // insensible, discovery*	E2+, X2.1, X2.2+, A9+, I1.1-	E4.1-, B1, X2.6-, X3.2-, O4.3, E4.1+, X5.2+, E3+, Q2.2, M2, X3.4, X2.4, A10+, X5.1+, T2+, I3.1

Label	Sample words	Relevant key domains in the dialogues	Relevant key domains in the non-dialogues
COMMUNICATION	*informed // said, told*	X2.2+, A9+, Z4	Q2.1, Q2.2, M2, T2+, M1, E4.1-, N5+
MARRIAGE	*accept // hand*	A9+	B1
EPISTEMIC MODALITY	*By no means, must // assurance*	Z6, A7, A7+, T1, S6+, A13.3, Z4	Q2.2
DEGREE AND QUANTITY	*great, greater // a great deal*	A5.1+, A5.1++, A13.3	E3+, N5+, T1.3-, X5.2+

In addition, the shared conceptual areas show different prominence in the two corpora (Graph 1). Prominence was decided based on the frequency of instantiation of a given conceptual area, that is, the percentage of the words representing it over the overall number of words in the relevant corpus.

Graph 1. Prominence of the conceptual areas appearing in both *novel_d* and *novel_nd*

The shared conceptual areas that are largely dominant in *novel_d* are the following: EVALUATION; MENTAL ACTIONS, FEELINGS AND ATTITUDES; MARRIAGE; EPISTEMIC MODALITY; and DEGREE AND QUANTITY. These conceptual areas serve the following purposes:

• Expressing evaluation (conceptual area EVALUATION): e.g. *and it has the <u>advantage</u> also of being in vogue*; *I am fond of <u>superior</u> society*; *George Wickham has treated Mr. Darcy in a most <u>infamous</u> manner*; *Is not this an <u>agreeable</u> surprise?*; *Miss Darcy will be to receive <u>such a</u> letter!*.

• Reporting mental actions, feelings and attitudes (conceptual area MENTAL ACTIONS, FEELINGS AND ATTITUDES): e.g. *I <u>fancy</u> she was wanted about the mince-pie*; *He can not <u>know</u> what Mr. Darcy is*; *I do not <u>believe</u> Mrs. Long will do any such thing*; and *I <u>had</u> not an idea of it*.

• Talking about marriage (conceptual area MARRIAGE): e.g. *why may I not <u>accept</u> him?; and if I were determined to <u>get</u> a rich husband*; *Lizzie declares she will not <u>have</u> Mr. Collins*.

• Expressing certainty and uncertainty (conceptual area EPISTEMIC MODALITY): e.g. *Your tempers are <u>by no means</u> unlike*; *<u>Perhaps</u> I did not always love him so well*; *There are <u>undoubtedly</u> many who could not say the same*; *I should <u>never</u> be happy without him*; *It <u>must</u> be his own doing*; *Very true, <u>indeed</u>*.

• Using expressions of degree and quantity (conceptual area DEGREE AND QUANTITY): e.g. *she is a very <u>great</u> favourite with some ladies*; *your surprise could not be <u>greater</u> than mine*; and *I am <u>exceedingly</u> gratified*.

When the same conceptual areas are to be found in *novel_nd*, they show distinctive traits. More specifically:

• The conceptual area EVALUATION is instantiated only in a few items (0.08% versus 0.3% *in novel_d*) belonging to a single key domain: e.g. *said in a <u>lively</u> tone*.

• The conceptual areas MENTAL ACTIONS, FEELINGS AND ATTITUDES has a more limited occurrence (1.87% versus 2.29% in *novel_d*), despite the fact that a wider variety of key domains are involved (16 versus 5 in *novel_d*). Sample concordances illustrating them are: *His character sunk on every <u>review</u> of it*; *To this <u>discovery</u> succeeded some others equally mortifying; a tone of gentleness and <u>commiseration</u>*; *she could not be <u>insensible</u> to the compliment of such a man's affection*.

• The conceptual area MARRIAGE has a very limited frequency of occurrence (0.006% versus 0.07% in *novel_d*), being instantiated in a single word: *Darcy approached to claim her <u>hand</u>*.

• The voicing of certainty and uncertainty (conceptual area EPISTEMIC MODALITY) is also extremely infrequent (0.02% versus 1.67%

in *novel_d*), and is represented by a single word: *assurance* (e.g. *he briefly replied, with <u>assurance</u> of his eagerness to promote*).

• Finally, expressions of degree and quantity (conceptual area DEGREE AND QUANTITY) are also less frequent in *novel_nd* (0.53% versus 1.66% in *novel_d*); illustrative examples are: *the <u>rest</u> of the evening passed*; *<u>A great deal</u> more passed at the other table*; *After a <u>short</u> silence*; *welcomed her friend with the <u>liveliest</u> pleasure*.

Similarly, a few conceptual areas are present in both corpora, but dominant in *novel_nd*. These are: MATERIAL ACTIONS; SOCIAL ACTIONS AND EVENTS; and COMMUNICATION. These conceptual areas suggest that Austen uses narration much more than conversation to achieve the following aims:

• Reporting material actions (conceptual area MATERIAL ACTIONS) as well as social actions and events (conceptual area SOCIAL ACTIONS AND EVENTS). Here are some relevant examples of material actions in *novel_nd*: *the person who <u>advanced</u> was now near enough*; *Mr. Bingley was to <u>bring</u> twelve ladies and seven gentlemen with him*; *She <u>saw</u> him start at this*. And here are some examples of social actions or events in *novel_nd*: *she looked forward to her <u>introduction</u> at Rosings*; and *Mr. Collins received and <u>returned</u> these felicitations with equal pleasure*.

• Reporting communicative exchanges (conceptual area COMMUNICATION): e.g. [quote] *<u>said</u> her mother resentfully*; *opening it with the utmost impatience, read as <u>follows</u>:* [quote]); *when Mrs. Bennet was <u>told</u> of this*; *Bennet <u>sent</u> them word that*.

When the same conceptual area are found in *novel_d*, they are less frequently instantiated. In particular:

• The conceptual area MATERIAL ACTIONS and SOCIAL ACTIONS AND EVENTS display limited lexical variety. The former includes almost exclusively the key semantic tag A9+ (e.g. *Colonel Forster is a sensible man, and will <u>keep</u> her out of any real mischief*), plus very few instances of S7.4+: (5 hits; e.g. *He meant to <u>provide for</u> me amply*) and E2+ (4 hits; e.g. *Miss Bingley is to <u>live</u> with her brother*). The latter, instead, includes only X2.2+ (e.g. *they have <u>known</u> her much longer that they have known me*), and A9+ (e.g. *We must <u>have</u> Mrs. Long and the Gouldings soon*).

• Words referring to communicative exchanges (conceptual area COMMUNICATION) are occasionally present (0.12% versus 0.36%), and are exemplified in the following concordances: *As Lydia <u>informed</u> you*; and *I <u>have</u> still another [thing] to add*.

4.1.5 Summary of findings

This analysis shows that the dialogic and narrative parts are largely complementary both thematically and lexically. In fact, the conceptual areas evidenced are either specific to *novel_d* or *novel_nd*, or dominant in one of the two corpora. Furthermore, when a conceptual area is common to the dialogues and the non-dialogues, it is dealt with through different lexical choices, as indicated by the different semantic tags.

In particular, the analysis showed that, as one would expect, the dialogues are rich in some of the lexico-grammatical features typical of spoken language, i.e. pronouns, exclamations, discourse markers, verbs indicating present time, and verbs associated with *I, we,* or *you* as subjects. They are used in the novel to describe future events, intentions, hypotheses, people and their characters, and talk about duties, obligations and abilities; they are also the preferred venue for reporting mental actions, feelings and attitudes, for expressing evaluation, certainty and uncertainty, as well as degrees and quantities, and to talk about marriage. On the other hand, the narrative parts identify the various characters and specify the time and place of actions, describe actions and events, and also the absence of expected things; furthermore, they are specifically used to report material actions, social actions and events, and communicative exchanges.

4.2 Film dialogues versus dialogues in the novel

The same analytical procedure described in the previous sections was also applied to contrast the dialogues of each of the two films (*2005 corpus* and *1940 corpus*) with the dialogues in the novel (*novel_d* corpus). The analyses returned the following picture:

- The key domains of both films illustrate the following conceptual areas: CHARACTERS, SOCIAL ACTIONS, MENTAL ACTIONS, SPOKEN TRAITS, FAMILY TIES, MARRIAGE, MATERIAL ACTIONS, EVALUATION and EPISTEMIC MODALITY (Table 5).

- The key domains of the 2005 film also instantiate the conceptual area EXPERIENCE (Table 6).

- The key domains of the 1940 film also comprise the conceptual areas COMMUNICATION, ANIMALS, WEALTH, COLOURS, FOOD, HOUSEHOLD, WEATHER, TIMELINE, and DEONTIC MODALITY (Table 7).

Three of the conceptual areas (EXPERIENCE, TIMELINE, and COMMUNICATION) are instantiated by a single word appearing one to three times. Such few occurrences do not permit a fine-tuned interpretation of the data. For this reason they will be discarded from the analyses.

The remaining are discussed in the following sections. In particular, we will consider the presence (or absence) of each conceptual area among the conceptual areas in the novel and, in the case of presence, its frequency of occurrence in the four corpora.

Table 5. Conceptual areas appearing in the key domains of the dialogues of both films

Label	Sample words	Relevant key domains in 2005 film	Relevant key domains in 1940 film
CHARACTERS	*Jane, Lady Catherine // Jane, Foster, Bennet*	Z1	Z1, S4, O4.3
SOCIAL ACTIONS AND EVENTS	*balls, dance, entertain // dinner, tea parties, standing*	K1, Z4	F1, W4, M6
MENTAL ACTIONS	*I believe, sorry, leisure // I believe, sorry, tempt, overlook*	Z4, K1	Z4, O4.2+, M6
SPOKEN TRAITS	*no, bye, oh, thank you, Mama // no, bye, oh, thank you, Mama, here*	Z4, S4, Z1	Z4, S4, Z1, L2, M6
FAMILY TIES	*aunt, brother // aunt, brother*	S4	S4
MARRIAGE	*engaged, married // engaged, married*	S4	S4
MATERIAL ACTIONS	*I say, you see // you see, feed, swinging, arrival, stand*	Z4	Z4, F1, O4.2+, M1, M6
EVALUATION	*kind of // kind of, well done, lamb, gloomy, delicacy, abreast*	Z4	Z4, L2, O4.3, F1, O4.2+, M6
EPISTEMIC MODALITY	*I suppose, of course // I suppose, of course, in the way*	Z4	Z4, M6

Table 6. Conceptual area appearing in the key domain of the dialogues of the 2005 film only

Label	Sample words	Relevant key domain
EXPERIENCE	*jilted*	K1

Table 7. Conceptual areas appearing in the key domains of the dialogues of the 1940 film only

Label	Sample words	Relevant key domains
COMMUNICATION	*tell you what*	Z4
ANIMALS	*horses, pigs, bird*	L2
WEALTH	*damask, silver, turtle soup*	O1.1, F1
COLOURS	*blue, pink, green*	O4.3
FOOD	*broth, soup, food*	F1
HOUSEHOLD	*kitchen, Officers' Mess*	F1
WEATHER	*rains, thunderstorm*	W4
TIMELINE	*dinner*	F1
DEONTIC MODALITY	*Come along, go*	M1

4.2.1 Conceptual areas appearing in the key domains of the dialogues of both films

Graph 2 illustrates the presence (or absence) in *novel_d* of the conceptual areas that appear to characterise the dialogues of both films, and their frequency of occurrence in each corpus.
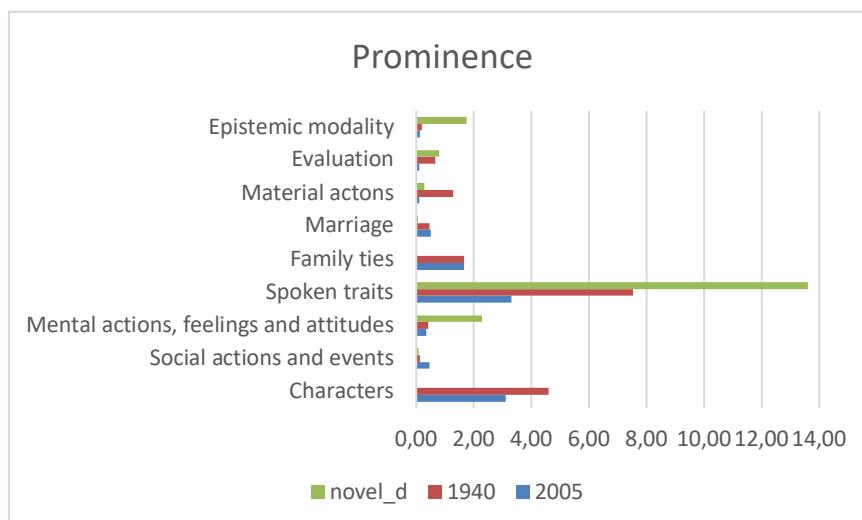
Graph 2. Prominence of conceptual areas appearing in both films, compared to *novel_d*

As Graph 2 shows, some of the conceptual areas emerging in both films can be considered distinguishing features of the film dialogues, given their prominence (i.e. frequency in percentage values) over the dialogues in the novel. These are: FAMILY TIES, MATERIAL ACTIONS, SOCIAL ACTIONS, CHARACTERS and MARRIAGE.

The conceptual area FAMILY TIES did not appear in our analysis of the novel. This does not mean that the novel does not mention family ties—Jane Austen's novel revolves around more than one family (the Bennets, the Lucases, the Bingleys and the Darcies), and words indicating family ties are bound to appear in it—but rather that reference to family ties is somehow equally distributed across the narrative and dialogic parts. The prominence of this conceptual area in the dialogues of both films can only be explained by the need, in films of all types, to verbally account for social and interpersonal relations among characters (Bianchi 2015: 242).

The conceptual areas MATERIAL ACTIONS, SOCIAL ACTIONS, and CHARACTERS, appeared exclusively (see Table 3, Section 4.1) or dominantly (see Graph 1, Section 4.1) in the narrative parts of the novel. This suggests that some of the conceptual areas dealt with by Jane Austen in the narrative part of the text have entered and acquired ample

space in the dialogues in these films. We can put forward some possible explanations as to why this is the case.

Both material and social actions are functional to the plot, and supporting the development of the plot is one of the roles that film dialogues generally perform (Zettl 2008: 196). Therefore their presence in the dialogues cannot not be considered unusual. However, all actions lend themselves to being represented visually. Thus, choosing to convey them verbally rather than visually is probably also functional to other needs. One such need could be to represent a large amount of action performed in different places and by several characters in the limited timespan of the duration of the film (less than two hours). Indeed, a visual representation that is sufficiently explanatory of who did what when and perhaps also why would generally require a longer film sequence than having a character reporting on it verbally. As an example, let us consider the following line which appears in both films: '*My father has gone to London*'. These six words can be easily uttered in about three seconds. In order to convey the same content visually we need to show Mr Bennet on a carriage running along country roads, and to see him reach London. Furthermore, in order to make sure viewers understand that this is London, the film sequence should also show shots of famous London landmarks of the time, say the Tower of London and Saint Paul's Cathedral. The whole scene would last at least a few minutes, and this is only a very simple example. Should the line be slightly more complex and also explain—for example—the reason why the action is done, the delta between the running times needed to convey the content verbally and visually would be much greater. Another possible explanation could be the script writer or film director's desire to give prominence to a specific action. In films, as in everyday life, talking about actions gives them relevance, and this is all the more true should the action be both shown and spoken about.

Furthermore, the presence of conceptual area CHARACTERS, i.e. the very frequent mentioning of characters' names in the film dialogues, may depend on several factors. Once introduced by name and role to the viewers, characters can be very clearly identified by their names; this is especially true in fiction, where names are purposefully selected and different characters rarely bear the same name, as this would generate confusion. Another possible explanation has to do with the fact that these dialogues largely discuss material and social actions. To do so in a clear

and functional way, it is necessary to mention the participant(s) in the action. Moreover, another reason of the frequent use of characters' names may be that this is a way to make the film dialogues sound more natural. Indeed, as Taylor (1999) observed, British and American films have reached a high level of linguistic realism, and the linguistic features that are more faithfully reproduced include the ample presence of vocative expressions. Additionally, the use of vocatives may be motivated by the need to be viewer-friendly, acting as a reminder of characters' identity and interpersonal relationships.

Finally, MARRIAGE was a dominant topic in the novel's dialogues too (Graph 1, Section 4.1), but its greater prominence in the dialogues of both films suggests a specific desire by the script writers or film directors to bring the topic to the fore.

The remaining conceptual areas appearing in the key domains of the dialogues of both films are EPISTEMIC MODALITY, EVALUATION, SPOKEN TRAITS and MENTAL ACTIONS. The presence of these conceptual areas in the film dialogues is not surprising. Indeed EPISTEMIC MODALITY, EVALUATION and MENTAL ACTIONS are almost impossible to represent visually, and SPOKEN TRAITS are a primary feature of dialogue of all types. What is more interesting is their quantitative relation compared to the dialogues in the novel: they have lower relevance compared to dialogues in the novel, where they appeared either as exclusive (Table 2, Section 4.1) or prominent in the dialogues compared to the non-dialogues (Graph 1, Section 4.1). Our explanation of this is that the need to give space in the film dialogues to elements that in the novel were reported in the narrative parts has reduced the running time available for other elements typical of the dialogic parts of the novel—a phenomenon we would like to call *selectivity*.

4.2.2 Conceptual areas appearing in the key domains of the dialogues of the 1940 film only

As shown in Table 7 (Section 4.2), seven conceptual areas appear in the key domains of the dialogues of the 1940 film only. These are: DEONTIC MODALITY, ANIMALS, COLOURS, FOOD, WEATHER, HOUSEHOLD and WEALTH. As Graph 3 shows, six of them—ANIMALS, COLOURS, FOOD, WEATHER, HOUSEHOLD and WEALTH—were not found in the key domains of the novel's dialogues. On the other hand, the conceptual area

DEONTIC MODALITY also appeared in the key domains of the novel's dialogues, while its prominence in the film's dialogues is lower.
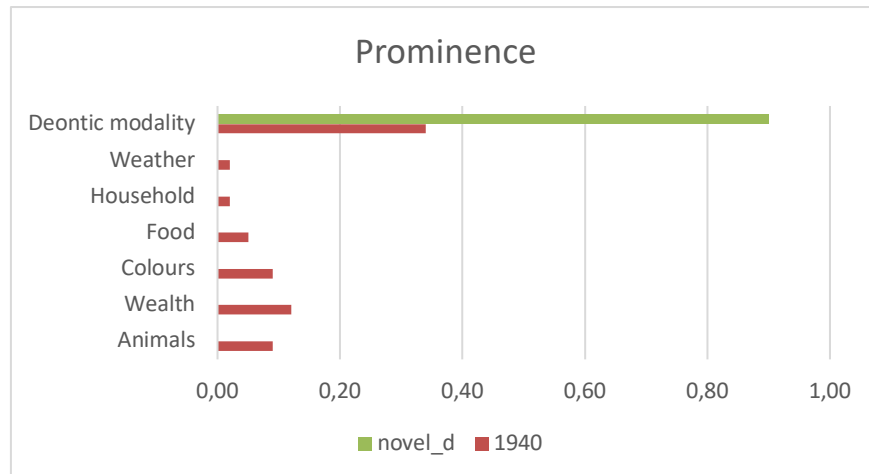
Graph 3. Prominence of the conceptual areas that appear in the 1940 film only, compared to *novel_d*

The conceptual areas ANIMALS, COLOURS, FOOD, WEATHER and HOUSEHOLD are connected to specific events or details in the story, such as the storm that makes Jane ill (WEATHER; ANIMALS), dresses for the ball or of the soldiers (COLOURS), preparing to move to Margot (ANIMALS) and Mrs Bennet's fainting moments (FOOD). These details are functional both to the plot and to the screwball comedy aspects of the film. Instead, explaining the presence of the conceptual area WEALTH requires a longer explanation.

The conceptual area of WEALTH is instantiated in words referring to objects or materials that tend to indicate opulence (e.g. silver, damask, pearl, marble, turtle soup). In the key domains of the novel's dialogues, the idea of wealth is indirectly present in the conceptual areas DESCRIPTION OF PEOPLE (Section 4.1.2) and instantiated in the words *poor* and *destitute*. A look at the wordlists of *novel_d* and *novel_nd* shows that neither corpus includes words of materials that tend to indicating opulence, and that references to money and wealth are scarce and equally distributed in the two corpora, which explains the absence of WEALTH in our analysis of the novel. We can thus say that the 1940 film places particular emphasis on the idea of wealth, expressed by reference

to material objects. An analysis of the visual aspects of the films would clarify whether these objects and material are also shown in the images; in any case, having characters remark the rich material the objects are made of guarantees that their value is immediately understood even by a less attentive audience.

Finally, the presence in this film's dialogue of the conceptual area DEONTIC MODALITY cannot be considered surprising, since orders, suggestions, promises and so on are easily performed verbally. What is more surprising is their absence in the key domains of the 2005 film, given that deontic modality was a characteristic feature of the novel's dialogues. A look at the concordance lines instantiating deontic modality in the 1940 film shows that they are nearly all orders (or harsh invitations) expressed through verbs of movement (key domain M6; e.g. *let's go; Come in, Mr Collins*; *Come, my dears*; *stand up, dear*). In the novel's dialogues, deontic modality is instantiated in a totally different range of verbs and expressions belonging to other semantic fields (S7.4+, A5.1++, S6+, A9+; see Table 2, Section 4.1.2). This suggests that the conceptual areas of DEONTIC MODALITY appear in the 1940 film dialogues for reasons that are unrelated to its presence in the novel's dialogues. Furthermore, the absence of similar phrases in the 2005 film suggests that they may serve to mark the film as a representative of the screwball genre, a genre where characters are black or white personalities and the scene is dominated by action.

4.2.3 Summary of findings

The analyses in this section showed that there is no direct correspondence between the topics dealt with in the dialogues in the novel and those dealt with in the film dialogues. Some of the conceptual areas appearing in the key domains of the film dialogues were characteristic of, or prominent in, the novel's dialogues (MARRIAGE, EPISTEMIC MODALITY, EVALUATION, SPOKEN TRAITS and MENTAL ACTIONS for both films; DEONTIC MODALITY and COMMUNICATION for the 1940 film only). Other conceptual areas were characteristic of, or prominent in, the narrative parts (MATERIAL ACTIONS, SOCIAL ACTIONS and CHARACTERS for both films), and yet others did not appear in the key domains of the novel subsets (FAMILY TIES for both films; ANIMALS,

COLOURS, FOOD, WEATHER, HOUSEHOLD and WEALTH for the 1940 film only).

Our analysis of the instantiations of the conceptual areas led us to attribute the presence of each of them to one or more of the following motivations, many of which are well-known functions of film dialogue: verbally clarifying the social and interpersonal relations among characters (conceptual area FAMILY TIES); supporting the development of the plot (conceptual areas MATERIAL ACTIONS and SOCIAL ACTIONS, ANIMALS, COLOURS, FOOD, WEATHER and HOUSEHOLD); representing a large amount of action performed in different places and by several characters in the limited running time available (conceptual areas MATERIAL ACTIONS and SOCIAL ACTIONS); giving relevance to a specific conceptual area (conceptual areas MATERIAL ACTIONS, SOCIAL ACTIONS, MARRIAGE, WEALTH); clearly identifying characters (conceptual area CHARACTERS); making the film dialogues sound natural (conceptual area CHARACTERS); impossibility of rendering a given conceptual area visually (conceptual areas EPISTEMIC MODALITY, EVALUATION and MENTAL ACTIONS); and characterising the film genre (conceptual areas ANIMALS, COLOURS, FOOD, WEATHER, HOUSEHOLD and DEONTIC MODALITY).

The analysis also showed that selective choices were probably made by script writers in creating the dialogues in order to fit such a long and complex story into the limited running time of a single film (lasting 113 to 127 minutes). This was evident in conceptual areas EPISTEMIC MODALITY, EVALUATION, MENTAL ACTIONS and SPOKEN TRAITS.

5. Conclusion

This study set out to explore whether and how the dialogic component in the novel *Pride and Prejudice* by Jane Austen differs in terms of conceptual areas from the non-dialogic parts, and from the dialogues in its 1940 and 2005 film adaptations. To do so, we compared the dialogic and the narrative parts of the novel, and then the dialogic subtitles of the two films, taken separately, against the dialogues in the novel. We used a corpus-informed method: automatic semantic tagging and the automatic extraction of key domains were accompanied by a manual analysis of the concordance lines instantiating the extracted key domains. This led to a reclassification of the topics as labelled by the software into a lower

number of conceptual areas, which took co-textual information into consideration and cut across key domains.

Comparison between the dialogic and the narrative parts in the novel (Section 4.1) showed that the two are largely complementary, highlighting their rather distinct conceptual areas. Comparison of the film dialogues to the dialogues in the novel (Section 4.2) showed both qualitative and quantitative similarities and differences in the topics covered. Some conceptual areas emerging as prominent in the films are also characteristic of the dialogues in the novel; others appear to have migrated from the narrative parts of the novel to the film dialogues. Still other conceptual areas, though present in the novel, are not characteristic or prominent in any of its two components. Our findings suggest that there is a complex interplay between the dialogic and narrative planes of discourse in a novel on the one hand, and between a novel and its filmic rendering, on the other.

This study illustrates in detail the varied functions that dialogue performs in one specific fictional text, and how these are adapted to the semiotic needs and goals of its film adaptations, highlighting the conceptual relationships existing between the original work and those specific film transpositions.

Given the research slant adopted, which involved comparing and contrasting the novel's dialogues all together first with the novel's narrative components taken as a whole, and then with the films' dialogues, other questions remain unanswered. A more fine-grained analysis might consider distinguishing, on the one hand, between uninterrupted dialogic utterances and those interrupted by suspensions (i.e. narrative segments between dialogues), and on the other, between general narrative parts and suspensions. A more ambitious goal would involve comparing and contrasting the text segments instantiating the different modes of discourse described by Hough (1970) so as to detect their specific semantic focus. In addition, it might be worthwhile to selectively examine each character's speech, both in the novel and the films, so as to determine their specific topics and styles. In parallel, the speech of the same character in the novel versus the films could reveal peculiar phraseological and/or content-related traits. This type of description could lead to character profiling (cf. Culpeper 2002). Finally, an analysis of the visual aspects of the films would ascertain whether some of the conceptual areas found in the dialogues (e.g. material and

social actions) are also present in the images, thus giving rise to intra-semiotic redundancy.

References

Aragay, Mireira and Gemma López. 2005. "Inf(l)ecting *Pride and Prejudice*: Dialogism, intertextuality, and adaptation." *Books in Motion. Adaptation, Intertextuality, Authorship*. Ed. Mireira Aragay. Amsterdam and New York: Rodopi. 181-200.

Archer, Dawn, and Paul Rayson. 2004. *Using the UCREL Automated Semantic Analysis System to Investigate Differing Concerns in Refugee Literature*. Accessed at https://pdfs.semanticscholar.org/bccf/25cfcf7ee4a388ba45b1b7af97c5f18955ef.pdf.

Bianchi, Francesca. 2015. "Integrazione e apprendimento. I prodotti cinetelevisivi come strumento didattico linguistico e culturale per il mediatore e il migrante." *Lingue e Linguaggi* 16: 237-263.

Bianchi, Francesca. 2016. "Subtitling Jane Austen: *Pride & Prejudice* by Joe Wright." *Pride and Prejudice: A Bicentennial Bricolage*. Ed. Caterina Colomba. Udine: Forum. 241-253.

Bianchi, Francesca and Sara Gesuato. 2019. "The adaptation of a novel to filmic needs: Thematic focus in the subtitles of two filmic versions of 'Pride and Prejudice'". *Textus* 2: 93-109.

Boozer, Jack. 2008. "Introduction: The screenplay and authorship in adaptation." *Authorship in Film Adaptation*. Ed. Jack Boozer. Austin: University of Texas Press. 1-30.

Chan, Mary M. 2007. "Location, location, location: The spaces of *Pride & Prejudice*." *Persuasions: The Jane Austen Journal On-Line* 27 (2): no pagination.

Cohen, Keith. 1979. *Film and Literature: The Dynamics of Exchange*, New Haven: Yale University Press.

Cobb, Shelley. 2008. "Adaptable Bridget: Generic intertextuality and postfeminism in *Bridget Jones's Diary*." *Authorship in Film Adaptation*. Ed. Jack Boozer. Austin: University of Texas Press. 281-304.

Culpeper, Jonathan. 2002. "Computers, language and characterization: An analysis of six characters in Romeo and Juliet". *Conversation in Life and in Literature. Papers from the ASLA Symposium. Association Suedoise de Linguistique*

Appliquée (ASLA) 15. Eds. Ulla Melander-Mattala, Carin Ostman and Merja Kyto. Uppsala: Universitetstryckeriet. 11-30

Culpeper, Jonathan, Dawn Archer and Paul Rayson. 2009. "Love— 'A familiar of a devil'? An exploration of key domains in Shakespeare's comedies and tragedies." *What's in a Word-List? Investigating Word Frequency and Keyword Extraction. Digital Research in the Arts and Humanities*. Ed. Dawn Archer. Ashgate: Farnham. 136-157.

Dole, Carol M. 2007. "Jane Austen and mud: *Pride & Prejudice* (2005), British realism, and the heritage film." *Persuasions: The Jane Austen Journal On-Line* 27 (2): no pagination.

Dudley, Andrew. 1984. *Concepts in Film Theory*. Oxford: Oxford University Press.

Fischer-Starcke, Bettina. 2010. *Corpus Linguistics in Literary Analysis. Jane Austen and her Contemporaries.* London and New York: Continuum.

Gehring, Wes D. 2002. *Romantic vs. Screwball Comedy: Charting the Difference*. Lanham, MA: Scarecrow Press.

Hough, Graham. 1970. "Narrative and dialogue in Jane Austen." *Critical Quarterly* 12 (3): 201-229.

Hudelet, Ariane. 2005. "Incarnating Jane Austen: The role of sound in the recent film adaptations." *Persuasions* 27: 175-184.

Klein, Michael and Gillian Parker (Eds.). 1981. *The English Novel and the Movies*. New York: Frederick Ungar Publishing.

Leech, N. Geoffrey and Michael H. Short. 2007. *Style in Fiction. A Linguistic Introduction to English Fictional Prose*. Second edition. Harlow, UK: Pearson Education.

Lu, Shenghui. 1999. *Transformation et Réception du Texte par le Film*. Bern: Peter Lang.

McFarlane, Brian. 1996. *Novel to Film: An Introduction to the Theory of Adaptation*. Oxford: Clarendon Press.

Mahlberg, Michaela, Peter Stockwell, Johan de Joode, Catherine Smith and Matthew Brook O'Donnell. 2016. "CLiC Dickens— Novel uses of concordances for the integration of corpus stylistics and cognitive poetics." *Corpora* 11 (3): 433-463.

Martin, Lydia. 2007. "Joe Wright's *Pride and Prejudice*: From Classicism to Romanticism." *Persuasions: The Jane Austen Journal On-Line* 27 (2): no pagination.

Orr, Christopher. 1984. "The discourse on adaptation." *Wide Angle* 6 (2): 72-76.

Otnes, Cele C. and Elizabeth Pleck. 2003. *Cinderella Dreams: The Allure of the Lavish Wedding*. Berkeley, CA.: University of California Press.

Parrill, Sue. 2002. *Jane Austen on Film and Television. A Critical Study of the Adaptations*. Jefferson, NC: McFarland.

Rayson, Paul. 2003. *Matrix: A Statistical Method and Software Tool for Linguistic Analysis through Corpus Comparison*. Ph.D. thesis, Lancaster University.

Rayson, Paul. 2008. "From key words to key semantic domains." *International Journal of Corpus Linguistics* 13 (4): 519-49.

Scholz, Anne-Marie. 2013. *From Fidelity to History: Film Adaptations as Cultural Events in the Twentieth Century*. New York, Oxford: Berghahn.

Snyder, Mary H. 2011. *Analyzing Literature-to-Film Adaptations. A Novelist's Exploration and Guide*. London: Continuum.

Stovel, Nora Foster. 2006. "From page to screen: Dancing to the altar in recent film adaptations of Jane Austen's novels." *Persuasions* 28: 185-198.

Stovel, Nora Foster. 2007. "An invitation to the dance and a proposal of marriage: Jane Austen's *Emma* and two film adaptations." *Persuasions: The Jane Austen Journal On-Line* 28 (1): no pagination.

Stovel, Nora Foster. 2011. "From page to screen: Emma Thompson's film adaptation of *Sense and Sensibility*." *Persuasions: The Jane Austen Journal On-Line* 32 (1): no pagination.

Stovel, Nora. 2013. "'Will you dance?" Film adaptations of *Pride and Prejudice*." *Persuasions: The Jane Austen Journal On-Line* 34 (1): no pagination.

Stubbs, Michael. 2001. *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford /Malden: Blackwell.

Taylor, Christopher. 1999. "Look who's talking. An analysis of film dialogue as a variety of spoken discourse." *Massed Medias: Linguistic Tools for Interpreting Media Discourse*. Eds. Linda Lombardo, Louann Haarman, John Morley and Christopher Taylor. Milano: LED. 247-278.

Wagner, Geoffrey. 1975. *The Novel and the Cinema*. Rutherford, NJ: Fairleigh Dickinson University Press.

Wakefield, Sarah. 2007. "How masculinity plays: Effects of musicianship in the 1995 film adaptation of *Sense and Sensibility*." *Persuasions: The Jane Austen Journal On-Line* 28 (1): no pagination.

Woodworth, Megan A. 2007. "'I am a gentleman's daughter'? Translating class from Austen's page to the twenty-first-century screen." *Persuasions: The Jane Austen Journal On-Line* 27 (2): no pagination.

Zettl, Herbert. 2008. *Sight, Sound, Motion. Applied media aesthetics*. Fifth edition. Wadsworth Pub. Co., Belmont.

Filmography
Leonard, Robert Z. 1940. *Pride and Prejudice*.
Wright, Joe. 2005. *Pride & Prejudice*.

Webliography
CLiC http://clic.bham.ac.uk/ (last access 28 December 2018).