Lingue e Linguaggi Lingue Linguaggi 27 (2018), 43-65 ISSN 2239-0367, e-ISSN 2239-0359 DOI 10.1285/i22390359v27p43 http://siba-ese.unisalento.it, © 2018 Università del Salento This work is licensed under a <u>Creative Commons Attribution 3.0</u>

REWRITING ROMEO AND JULIET FOR A YOUNG AUDIENCE A corpus-assisted case study of adaptation techniques

FRANCESCA BIANCHI

Abstract – Children's literature has been explored from different perspectives. General agreement seems to exist on the fact that writing for children involves adjusting contents and language (vocabulary and syntax) to the target audience, but no systematic and detailed description of the linguistic strategies used or required to adapt texts to young audiences is available. The current chapter analyses two narrative versions of Shakespeare's *Romeo and Juliet* written in contemporary English by the same author for two young audiences of different ages, and investigates this author's adaptation techniques through corpus-assisted methods. The analyses show that the author has resorted to a clear set of adaptation techniques, with some differences in the two texts. These language and cultural context adaptation strategies (Klingberg 1986) are in perfect keeping with the affective needs and cognitive abilities of each age group as described in theoretical and empirical studies on children's literature and developmental psychology.

Keywords: children's literature; corpus linguistics; stylistic analysis; *Romeo and Juliet*; adaptation techniques.

1. Introduction

Rewriting and adapting classical masterpieces for children is a practice that goes back in time. Originally,

books were adapted for children in order to offer them 'great literature', [but] it was often with a didactic intention. These adaptations therefore reveal much about contemporary attitudes and expectations with regard to young readers and childhood in general. (Beckett 2009, p. 19)

inauaaa

This practice has continued to our times. Contemporary age-targeted versions of famous masterpieces, which are the focus of the current work,¹ fall somewhere in between a literary text and a pedagogical tool with both cultural and linguistic aims.

Children's literature – including adaptations of adult texts – has been largely explored from three perspectives: literary scholars have investigated the position and role of children's literature, or of single authors or books for children, within literature as a historical, cultural and social polysystem (e.g. Hunt 2005; Lundin 2004; Shavit 1986; Vandergrift 1990); other scholars have adopted a pedagogical perspective and focused on how to teach literature or reading abilities to children through children's literature (e.g. Barone 2011; Gamble, Yates 2002); finally, translation scholars have illustrated the main issues connected to translating children's texts into a different target language and for a different culture (e.g. Klingberg 1986; Lathey 2015; Oittinen 2000; Shavit 1986; van Coillie, Verschueren 2006). In these works, there appears to be general agreement that writing for children involves adjusting contents and language (vocabulary and syntax) to the target audience, but to the best of my knowledge no systematic or detailed description of the linguistic strategies used or needed to adapt texts to young audiences is available.

The current work takes a step in this direction, investigating adaptation techniques and using the methods of corpus linguistics. In particular, this study analyses and compares two narrative versions of Shakespeare's *Romeo and Juliet* written in contemporary English for two young audiences of different ages by an experienced literature and language schoolteacher. The study addresses the following research questions:

- 1. What kind of adaptation techniques did this author employ?
- 2. Did the author really differentiate between his two target audiences, and, if so, how?
- 3. To what extent are the techniques used by this author in keeping with scientific observations on the tastes, needs, and abilities of young readers?
- 4. To what extent can corpus linguistics help us to outline the adaptation techniques adopted in texts of this type?

To these aims, Section 2 provides an introduction to children's literature by focussing on specific content and language issues directly related to creating or adapting texts for different age groups. Section 3 offers a compact review of how corpus linguistics can be used in the analysis of language and style in literary texts. Section 4 describes the material and analytical methods

¹ The current work focuses on versions specifically created for children and children's literacy in L1. This excludes graded readers, which – though sometimes used with children – target adult learners studying English as a FL/L2 (Hill 2008).



employed in the current study, while Section 5 discusses the results of the analyses. Finally, Section 6 summarizes the most important results of the current study and, by evidencing some of its limitations, suggests possible paths for further research on the topic.

2. Children's literature: content and language issues

As briefly mentioned in the Introduction, scientific studies on children's literature provide only rather broad information on how texts are or should be written or adapted for children. The considerations that are most relevant to the current study are summarised in the following paragraphs.

Analysis of children's books show three typical ways of writing a story for children (Lathey 2015): adopting the voice of the oral storyteller ("Once upon a time…"); adopting an omniscient narrative stance, but with ironical comments to adult inconsistencies, a sort of "narrative voice that conspires with the child reader to unmask ridiculous aspects of adult expectations and behaviours" (Lathey 2015, p. 19 in the Kindle edition); and adopting a child narrator, i.e. a child speaking to his/her peers. Analyses of adaptations across times have shown that adaptations are generally based on what society believes to be pedagogically and morally appropriate for children (e.g. Klingberg 1986; Shavit 1986), implementing adjustments that can be classified into six macrocategories (Klingberg 1986): cultural context adaptation; modernization; purification; language adaptation; abridgement; and localization. Furthermore, Bell (1986, cited in Lathey 2015) argues that the dominant narrative tense in English is the simple past, and suggests that children's books should use it as their primary narrative strategy.

Appleyard (1991) illustrates the interests and needs of readers considering age-related developmental stages, each stage corresponding to a different approach to reading. He identifies five reader roles across a person's life, from early child to adulthood. In later childhood and adolescence – the age groups targeted in the current study – the reader is described, respectively, as Hero or Heroine, and as Thinker. Readers at both these developmental stages require a narrative structure that is complex enough to hold the child's attention, and characters with whom the child can identify. For later-childhood readers, characters are what they do, which involves presenting them primarily through dialogue and action, plus a few distinctive traits. Furthermore, for this age group, characters should ideally be fairly simple (either heroes or villains). Identification with the characters in the story becomes even stronger among teenage readers, and at this stage it is important that "the characters of adolescent novels match their readers' newfound sense of complexity, but do not exceed it" (Appleyard 1991, p. 106). However, what really distinguishes

jingue e

teenage readers from other age groups is their appreciation for realism and for stories that make them think, which, according to Sellinger Trites (2000) and James (2009), calls for treatment of topics of specific interest to them such as death and sexuality.

Finally, general agreement exists on the need to adjust vocabulary and syntax to the reading abilities of young audiences, but no description of what this entails can be found in the scientific studies on children's literature. For some insight into this issue, we need to look into experimental studies on language acquisition and reading comprehension. Reid and Donaldson's (1977, cited in Gamble, Yates 2002), for example, observed that embedded subordinate clauses, hidden negative clauses, and passive voice are not easily understood by children, while Chapman (1987, cited in Gamble, Yates 2002) singled out cataphoric reference, ellipsis, and conjunctive ties. Long et al. (1997) observed that less-skilled readers have difficulties in making causal inferences. Other researchers have found that less-skilled readers benefit from the presence of section titles (e.g. Cain, Oakhill 1996; Yuill, Joscelyne 1988). Finally, on the lexical level, less-skilled readers have been seen to be hampered by low-frequency words (e.g. Nation, Snowling 1998), while Gamble and Yates (2002) suggest that figurative language may be problematic for children readers.

Adaptations of adult novels to youngsters should consider some or possibly all the factors illustrated above.

3. Corpus linguistics and literary texts

Corpus linguistics – i.e. "the electronic analysis of language data" (Fischer-Starcke 2010, p. 1) – has been employed in the investigation of English texts of all natures, including literary texts. In literary-text analysis, corpus methods may contribute to "document[ing] more systematically what literary critics already know (and therefore add to methods of close reading), but they can also reveal otherwise invisible features of long texts" (Stubbs 2005, p. 22). This section introduces the corpus linguistics tools and analytical methods used in the current study, and briefly illustrates how they have been applied so far to the analysis of language and style in literary texts.

By tools I mean software programmes specifically designed to investigate (concordancers) or add annotations (taggers) to a corpus. Concordancers retrieve all instances of a given word or phrase and display them along with their surrounding co-texts. Taggers enhance a corpus with technical annotations. Two types of automatic annotation have been largely used in the analysis of literature and will be employed in this study: semantic annotation, in which every word in the corpus is matched to a semantic field; and part-of-speech (POS) tagging, in which every word in the corpus is labelled according to the morphosyntactic category it belongs to. An annotated corpus can be investigated starting from the annotated information, as well as from words in the corpus.

In the analysis of literary texts, one particular method appears to have been preferred over others by researchers, and this is keyword analysis, along with extensions of the keyword concept to semantic tags (key domain analysis) and POS tags (key POS tag analysis). In keyword analysis, the word list of a corpus is automatically compared to that of another corpus (called 'reference corpus'), in order to highlight words with outstanding frequency. The degree of outstandingness of the specific word in the corpus is called 'keyness' and is established by statistical methods. Unusually frequent words in comparison to the reference corpus are called 'positive keywords', while unusually infrequent words are called 'negative keywords'. Keywords provide information on the contents of a corpus, in terms of its aboutness, but also its structure and style (e.g. Fischer-Starcke 2010; Scott, Tribble 2006). Different reference corpora generally provide different outcomes (see for example Fischer-Starcke 2010); however, any reference corpus may yield interesting results (Scott and Tribble 2006, p. 65). The automated extraction of keywords is generally accompanied by manual analysis of their concordances, that is lines of texts surrounding the keywords. Concordances help researchers to observe semantic and grammatical patterns; these are technically known as collocations and colligations. The words that co-occur with a node word are called 'collocates'. Analysis of a node's collocates allows researchers to observe its semantic preference, i.e. the node's preferential association with a given experiential domain, and its semantic prosody, i.e. the semantic qualifications the node word derives from its association with other words in the co-text.

Keywords allowed Fischer-Starcke (2009, 2010) to identify the themes of two novels by Jane Austen and enabled Gerbig (2010) to show that travel writing has changed in time not only in terms of the topics addressed, but also as regards the positioning of the travel-writer within the story. In Mahlberg (2010), and Mahlberg and Smith (2010), keywords were the starting point for identifying potentially interesting words for further analyses. Culpeper (2002) used the keywords tool to cross-compare the characters in a play and highlight different lexical and grammatical patterns for each of them. Finally, in an analysis of Shakespeare's Romeo and Juliet, Scott and Tribble (2006), observed that beside illustrating the aboutness of the play, the keywords were able to give insight into some stylistic devices, such as a distinctive use of exclamations 'oh' and 'ah'. Furthermore, Archer, Culpeper, and Rayson (2009) used semantic tagging and the extraction of key domains to compare Shakespeare's love tragedies to love comedies, finding that the two sets of data were characterized by different domains and different love metaphors. Finally, in Murphy (2007), analysis of key POS tags in the soliloquies of 12 works by

Shakespeare clarified that soliloquies provide important stage directions for the characters (third person form of lexical verbs), perform an evaluative function of other characters or actions (general adjectives), reveal the speaker's intentions (first person singular pronoun) and express generalizations (plural common nouns), this last property being possibly related to a moralizing function.

The current study draws inspiration from the papers above and adopts keyword analysis – and its extensions – as its primary method of investigation for the identification of the stylistic features and linguistic traits that characterize two narrative adaptations of *Romeo and Juliet* for young audiences.

4. Materials and methods

This work analyses and compares two narrative versions of *Romeo and Juliet*, distributed by the e-book publisher *No Sweat Shakespeare*,² taken from its children's series (*The Shakespeare for Kids* series), and from the teenage series (*Modern English Shakespeare* series). All the books in those series are written by Warren King, a former schoolteacher and a specialist in bringing Shakespeare to young audiences.³

The e-books in *The Shakespeare for Kids* series – aimed at children aged 8-11 – are collectively described on the website as follows: "[these books] tell the stories of Shakespeare's plays in very simple language, are highly abridged and are not broken up into acts and scenes".⁴ The *Modern English Shakespeare* series – for a more grown-up audience – includes descriptions specific to each volume, and the one of *Romeo and Juliet* reads: "translated as an easy to read, exciting teenage novel. Follows the acts & scenes of original Romeo & Juliet text. Allows you to master the plot, characters & language of Romeo & Juliet".⁵ These different descriptions and the fact that these two narrative versions of *Romeo and Juliet* are presented as two separate editions creates

⁵ https://www.nosweatshakespeare.com/ebooks/modern-romeo-juliet/



² <u>https://www.nosweatshakespeare.com/ebooks/</u>

³ As declared in *Romeo & Juliet - for kids* (section About The Author), Warren King "has been teaching English literature for thirty-five years in English comprehensive and public schools. During the 1980's he was seconded to the national Shakespeare and Schools project to help develop methods of teaching Shakespeare in the classroom to bring the plays to life for pupils of all ages. After the project ended he continued that work as an adviser in to a London Education Authority, where he worked with teachers in creating Shakespeare projects in schools and helping English teachers, both primary and secondary, to make Shakespeare lively, comprehensible and enjoyable for their pupils."

⁴ <u>https://www.nosweatshakespeare.com/ebooks/</u>

expectations on the existence of substantial differences between them, the differences being due to the different target audiences.

The two e-books considered in this study were converted into text-only files, after removing the tables of contents, information about the series and the author, and copyright notices. The cleaned-up texts, now containing exclusively the *Romeo and Juliet* narratives, were automatically tagged with Wmatrix (Rayson 2009), an online corpus analysis and tagging system that also detects multi-word units (such as *give up* and *on his own*) and treats them as single words. Wmatrix includes a highly efficient POS tagging software for English texts (CLAWS; Garside, Smith 1997) and a semantic tagger (USAS; Rayson *et al.* 2004), and integrates several reference corpora to use for the automatic extraction of key words, key POS tags, and key domains.

Table 1 provides a quantitative outline of the two texts under investigation, the children's and the teenager's versions of *Romeo and Juliet*, here called C-R&J and T-R&J for short.

	Children's version (C-R&J)	Teenager's version (T-R&J)
Tokens (Running words)	26,782	27,619
Types	3,529	3,610
POS tags	145	146
USAS tags	332	336

Table 1

Quantitative outline of the children's and the teenager's versions of Romeo and Juliet.

In overall quantitative terms, the two narrative versions show great similarities. The teenage version is only 837 words longer (+3.12%), and shows 81 types more (+2.29%). These texts also feature practically the same number of grammatical forms (POS tags) and semantic fields (USAS tags).

The texts were first compared to three sub-sets of the British National Corpus (BNC): Sampler Written Imaginative; Sampler Written Informal; and Sampler Spoken.⁶ The Written Imaginative sub-set includes drama, poetry and prose fiction. The other two sets were chosen because they include frequent direct dialogue. Comparison with the BNC Sampler Written Imaginative helped to highlight the strategies adopted by the author to adjust the play for young audiences, by evidencing the peculiarities of the texts under investigation compared to other forms of creative writing for a general (usually adult) audience. It was also a first step in observing similarities and differences between the two narrative versions. Comparison to the other two sets of the

jingue e

⁶ <u>http://ucrel.lancs.ac.uk/bnc2sampler/sampler.htm</u>

BNC served the purpose of verifying whether certain features (in particular the presence of everyday vocabulary) were a specific adaptation strategy or a consequence of the presence of ample stretches of dialogue in the texts. Comparisons were performed at the level of words (keyword lists), and POS tags (key POS tags). In the automatic comparisons to the BNC sub-sets, only the positive key items with very high statistical significance (LL > 15.13, p < 0.0001, 1d.f.) were examined. While some keywords were immediately interpretable, others required scrutiny of their concordance lines and an analysis of their co-text, or plotting – a visual feature that shows in what part of the text/story a word appears. For concordancing and plotting, AntCont (Anthony 2013)⁷ was used. The similarities and differences emerging from these comparisons are presented in Section 5.1.

Finally, the children's and the teenager's versions of Romeo and Juliet were compared to each other, for a more systematic analysis of their differences. First, automatic comparison between the two texts was performed using Wmatrix, at the level of words, POS tags and semantic tags. However, even considering all possible log-likelihood cut-off values,8 this form of comparison returned zero key POS tags, only one positive keyword with statistical significance (act) and one key domain (semantic tag K4 - DRAMA, THE THEATRE AND SHOW BUSINESS). This would suggest that the division of the teenage version in Scenes and Acts is the only difference between the two texts. Hypothesizing different degrees of lexico-grammatical complexity between the two versions, I used Sketch Engine⁹ (Kilgarriff et al. 2014) to search the texts for specific grammatical structures (noun phrases, relative clauses, passive voices, that subordinate clauses), but once again no substantial differences could be observed between the two texts. Believing that some difference must distinguish the two editions besides division in Acts and Scenes, I explored the texts further by using the Compare feature in MS Word. This evidenced a series of differences, which were manually analysed and classified. The results of this manual analysis are discussed in Section 5.2.

5. Results

Section 5.1 presents and discusses the results obtained by comparing the children's and the teenager's versions of *Romeo and Juliet* with the BNC. Section 5.2 compares the narrative texts to each other.

⁹ <u>https://www.sketchengine.co.uk</u>



⁷ <u>http://www.laurenceanthony.net/software.html</u>

⁸ <u>http://stig.lancs.ac.uk/cgi-bin/wmatrix3/help.pl#logl</u>

5.1. The narrative versions vs. the BNC

Automatic comparison of the two narrative versions of *Romeo and Juliet* to the Written Imaginative sub-set of the BNC Sampler returned 78 and 77 positive keywords for the children's version and the teenager's version, respectively, and 13 key POS tags above the given threshold for each version. As many as 72 items (92-93%) in the keyword lists are key for both texts, with minor rank differences, while the two POS lists include exactly the same positive key items, with minor rank differences. For this reason, the keywords and POS tags above the threshold are listed in alphabetical order in Tables 2 and 3.

As expected, the positive keywords of the narrative versions (Table 2) provide insights into the aboutness of the texts, but also indications about their structure and style. Similarly, their positive key POS tags (Table 3) primarily provide insights into the stylistic peculiarities of these texts, but not only. As we shall see, the key words and the key POS tags converge in the same directions.

Common to the two texts				C-R&J only	T-R&J only	
as though	face	holy	n't	this	find	act
beautiful	fight	hurry	nurse	Thursday	girls	even
bed	get up	husband	oh	torch	oh no	fellow
beg	give	if	Paris	turned	please	scene
chapel	go	ill	's	wedding	prince	wake up
come on	go on	kill	said	what	thumb	
count	going to	killed	servant	wife		
cousin	gone	let 's	servants	WO		
crying	grave	lips	she	word		
dancing	he	'11	stared	you		
dead	heaven	love	stop	young		
death	her	'm	stopped	your		
desperate	here	man	swear			
die	him	master	sword			
do	his	me	tell			

Table 2Keywords compared to BNC Sampler Written Imaginative.

inguaggi

Tag	Description
APPGE	possessive pronoun, pre-nominal (e.g. my, your, our)
CS21	subordinating conjunction (e.g. if, because, unless, so, for)
PPHO1	3rd person sing. objective personal pronoun (him, her)
PPHS1	3rd person sing. subjective personal pronoun (he, she)
ppio1	1st person sing. objective personal pronoun (me)
PPY	2nd person personal pronoun (you)
UH	interjection (e.g. oh, yes, um)
VBZ	is
VM	modal auxiliary (can, will, would, etc.)
VM21	modal auxiliary (can, will, would, etc.)
vv0	base form of lexical verb (e.g. give, work)
VVD	past tense of lexical verb (e.g. gave, worked)
VVGK	ing participle catenative (going in be going to)

Table 3Key POS tags compared to BNC Sampler Written Imaginative.

5.1.1. Key items pointing to content and structure

Without even checking concordance lines, it is easy to understand that some of the common keywords portray main themes in the play, which are:

- A tragic story: crying; dead; death; desperate; die; grave; ill; kill; killed.
- Love and marriage: *beautiful; husband; lips; love; wedding; wife*.
- A noble setting: *count* (capitalized in the concordance lines); *master; servant; servants.*
- Religion: chapel; heaven; holy.
- Fights: *fight; sword*.
- Family ties: *cousin; husband; wife*.
- Intentions and promises: *going to; let's; if; 'll; swear; wo + n't*; key POS tag VVGK.

Knowledge of the original plot suggests that also other common items in the list relate to content. *Thursday* is the day when Juliet would have to marry Paris, but drinks a sleeping potion instead. The *Nurse* and *Paris* are important characters in the play. Some important events in the plot (Romeo sneaking in at Capulet's party; Romeo and Paris visiting Juliet's tomb) take place at night, and moving at night requires a *torch*. Keyword *dancing* is connected to Capulet's party, where Romeo and Juliet fall in love at first sight, while *hurry*

signals the urgency that marks some of the characters' actions. Finally, Juliet is found (apparently) dead in her *bed*. (A check of the concordance plot of *bed* using AntConc confirms that this keyword is mostly concentrated in a single point towards the end of the text).

If we consider the very few keywords differentiating the children's version from the teenage version, we observe that a few of them, too, point to content and plot. This is the case of keywords *Prince* in C-R&J, and *wake up* in T-R&J. The keyness of the former is a consequence of the key role played by Prince Escalus in the plot. The verb, instead, significantly appears exclusively towards the end of the file when the Friar's expedient is first described and then put into action.

Finally, words *act* and *scene*, which characterize the teenage version, appear as key because only this version is divided into acts and scenes.

5.1.2. Key items relating to adaptation – Presenting characters through dialogue and action

Some keywords appear to be associated with direct or reported speech. In fact, said - ranking third in both keyword lists ordered by LL - always appears in accompaniment to direct dialogue (e.g. 'Right,' he said. 'We're on. Pick a fight with them. I'll be right behind you.' 'That's what I'm afraid of,' said Gregory.). Also *come on* predominantly appears in spoken interaction (96.6%), with the frequent intent to "encourage [a person] to do something they do not much want to do" (53.6%), including stopping their current actions, and less frequently to "encourage them to hurry up" (28.6%), or to signal that "what they are saying is silly or unreasonable" (17.8%).¹⁰ Many of the verbs in Table 2 are frequently used in the imperative form or with subjects I or you, both circumstances being indicative of direct dialogue: get up (66.7% imperative); go on (91.7% imperative); give (48.8% imperative; 36.6% I/you subject; total of 85.4%); stop (55.6% imperative; 7.4% I/you subject; total of 63%); tell (58.8% imperative; 32.4% *I/you* subject; total of 91.2%); go (49.4% imperative; 12% I/you subject; total of 61.4%); not to mention imperative let's. This is also evident in the key POS list, with the presence among key items of tags VM21 (let's), and VV0 (base forms, i.e. imperatives), and of items PPIO1 (me), PPY (you), and APPGE (my; your; etc.). Finally, deictics here and this can be markers of spoken discourse, and, indeed, in the current texts they appear predominantly in dialogic lines.

¹⁰ The Collins English Dictionary (<u>https://www.collinsdictionary.com/dictionary/english/come-on)</u>



Other key items, too, point to the presence of spoken discourse; these are keyword *please* appearing in the CR&J list only, and exclamations, i.e. key POS tag UH, keyword *oh* (common to both texts), and keyword *oh no* present in C-R&J only. (For more details on exclamation *oh*, see the next section.)

Finally, key POS tags VVD (past tense of lexical verb), but also PPHS1 (3rd person sing. objective personal pronoun) and PPHO1 (3rd person sing. subjective personal pronoun) can be interpreted as indicators of the narrative texture into which direct or reported speech is inserted.

Integrating ample stretches of dialogue into the narrative texture can be considered a specific adaptation strategy that performs two fundamental functions: on the one hand, it maintains some of the dialogic features of the original play; on the other hand, it meets the young audiences' need for characters presented through dialogue and action (Appleyard 1991).

5.1.3. Key items relating to adaptation - Making the text dynamic

The unusually high frequency of go is probably due to the fact that the texts under examination are narrative versions of a play. On stage, actors continuously enter and exit the scene, often upon the orders of other characters. This makes the scene dynamic despite the contents of the dialogues. In written plays, stage directions as well as dialogues contribute to providing readers with information about the characters' movements, thus suggesting the idea of dynamic action. In the current narrative versions, this role is performed by lines like those reported in Excerpt 1, where the characters' utterances, including verb go, evoke changes of scene and lots of events.

Let 's	go	to Capulet 's party , Romeo . No-one will mind.
Alright , I'll	go	, said Romeo .
I have to	go	. Come on Juliet , the Count 's waiting .
Go on , darling. said the Nurse .	Go	and meet your love .
Hand me a mask , Benvolio . Let 's	go	Come on , then , said Benvolio .
Let 's	go	, said Benvolio , If we do n't get a move on
Alright then , off we	go	. [END OF CHAPTER]
Tybalt stopped . He turned .	Go	and get my rapier , Boy.
then to a passing servant :	Go	and get more torches
You 'd have to	go	to Friar Lawrence 's chapel for that
Come on, let 's	go	. Hey Mercutio !
That 's enough . Time to	go	. Where 's Horatio ?

Excerpt 1 Extract from the concordancing of keyword *go*.

inauaaa

Thus, the presence of the verb *go* in the keywords is indicative of the author's attempt to convey dramatic action in writing. This makes the text more dynamic and more successful at holding the reader's attention, thus meeting a fundamental need for young readers (see Appleyard 1991).

5.1.4. Key items relating to adaptation – Explication strategies

Some key words and POS tags suggest the idea that the author deliberately tried to be clear and explicit. These are keywords *do*, *as though*, *oh*, *oh no*, *even*, and *face*, and key POS tag CS21.

When used in its auxiliary role (68% of instances), the verb do appears primarily in negative constructions (78.2%) – which contributes to explaining the keyness of n't –; it also appears in interrogative or emphatic constructions, though in much smaller proportions (respectively, 19.3% and 1.7%). Overt negative structures, compared to hidden ones, are much easier to understand for children (see Reid, Donaldson 1977), and the systematic use of the former over the latter can be considered a strategy aimed at explicating rather than implying.

The key POS tag CS21 (subordinating conjunction) includes the following items: as if / as though (38.5%; as though appears also key in the keyword list); even if / even though (23.1%); so that (28.9%); seeing that (3.8%); in case (1.9%); now that (1.9%); and rather than (1.9%). These conjunctions – which in Systemic Functional Grammar (Halliday 2004) are almost all classified as instances of enhancement, of causal-conditional type (even if / even though; so that; seeing that; in case; now that) or manner (as if / as though) – explicate relations between circumstances and events. In other words, they are there to facilitate comprehension of the characters' actions and decisions.

Keywords *oh* (appearing key in both texts), and *oh no* (in C-R&J) are rhetorical ways to emphasize the distress of the characters. Analysis of the concordances of *oh* will clarify this point. In the current texts, *oh* appears 74 times, compared to only 8 instances of *ah*. *Ah* indicates primarily surprise (62.5%); in the remaining cases it appears in a positive environment (all utterances by the Nurse). *Oh*, on the other hand, is used mostly in negative circumstances (63.5%), and more rarely in positive circumstances (18.9%), ambiguous sentences (12.2%), mockery (2.7%), or to indicate surprise (2.7%).¹¹ Thus, in these texts *Oh* seems a rhetorical way to emphasize the

¹¹ Compare these uses to Scott and Tribble (2006), where, in analysing Shakespeare's *Romeo and Juliet*, the authors observed that 'oh' was more likely to be used in a positive environment, than 'ah' (although both appeared in mockery).

distress of the characters, which can be considered an explication strategy to the benefit of young readers.

Keyword *even*, which characterizes the teenage version, performs an emphatic role, too. It appears exclusively in its adverbial function, and is meant to emphasize feelings and circumstances, thus making them more evident.

Finally, concordancing of keyword *face* shows that this word accompanies adjectives 'his', 'your', 'her' or a genitive structure ('s') in 1L position; furthermore, it is followed to the right by verbs 'was' or 'showed' (1R) and nouns or adjectives indicating emotions (2R; 'concern'; 'surprise'; 'twisted'; 'solemn'; 'grim'; 'dripping [water]'; 'stinging'; 'growing [redder]'; 'red'; 'pale'; 'bad'). Thus, in the narrative versions under analysis the description of a character's face appears to be an important means to convey the character's emotions.

5.1.5. Key items relating to adaptation – Simplification and repetition strategies

The verbs *gone, stared, stopped, turned, do* (when used as lexical verbs; 32% of cases), 'm (am), and 's (i.e. is; also key as POS tag VBZ), and the noun man can be interpreted in the light of the author's attempt to "tell the stories of Shakespeare's plays in very simple language". This is true also for personal pronouns and adjectives (*he; her; him; his; me; she; you; your;* POS tags PPHO1, PPHS1, APPGE). With the only exception of *stared*, these words are key also against at least one of the other two reference corpora ('m, 's, you), if not both (give; gone; stop; stopped; tell; turned; he; her; him; his; me; she; your), which means that they are statistically more frequent in these narrative versions than in written informal texts and/or spoken language. This lends support to the idea that the author of these versions has systematically resorted to simple everyday language and used a set of rather general verbs and nouns more frequently than in everyday conversation or informal writing.

Furthermore, words *word* (key in both texts), *find* and *girls* (key in C-R&J), and *fellow* (key in T-R&J) also contribute to this category of common words repetitively used for the sake of simple style.

Finally, a few keywords tend to appear in set phrases. This is the case of *beg*, which appears in phrase 'I beg of you' in 70% of concordance lines, and *word*, which appears in the phrase '[I want/'d like] a/one word with you/one of' in about 21% of cases. Repeating set expressions could be considered a strategy aimed at facilitating understanding by less-expert readers or readers with limited vocabulary.

5.1.6. Key items relating to adaptation – Strategies that help readers to identify with characters

Keyword *young* collocates with a range of nouns – including 'man', 'men', 'Romeo', 'lady(ies)', 'woman', 'servant(s)', 'fellow', 'Count', 'Capulet', 'Montague', and 'couple'. Although also in the original play many characters are described as young, so much emphasis on the young age of the characters may be a device to help the new, young audience to identify with them.

5.1.7. Key items that are difficult to interpret

The few key items that have not yet been discussed are rather difficult to interpret. These are: key POS tag VM; and keyword *thumb* (C-R&J only).

Tag VM (modal auxiliary) expresses primarily intentions (56%; *will*, *shall*, *'ll*, *would*, *won't*), ability and possibility (33.3%; *can*, *could*), and only minimally other types of stances (*may*, *might*, *must*, *should*, for a total of only 9.8%).

Finally, *thumb* (key in the children's version) appears exclusively in the set phrase 'bite [one's] thumb at [someone]', "[a]n archaic insult, often accompanied by the gesture of biting one's thumb at the person being insulted" (The Free Dictionary).¹² Insulting gestures are more typical of young people than senior ones, so reference to this gesture could be a way to emphasize the young age of the characters, so that the young audience can identify with them. However, this is an archaic insult, which might not be easy to understand for children aged 8-11.

5.2. C-R&J vs. T-R&J

This section summarises the results of the current manual analysis and classification of the difference between the two narrative texts, assisted by an automatic comparison between them using MS Word's Compare tool.

A macroscopic difference between the two versions considered in this study lies in the fact that, in place of scenes and acts, C-R&J is divided into sections introduced by a title, and sometimes also by a brief descriptive paragraph providing background information. Indeed, less-skilled readers have been shown to benefit from the presence of section titles (e.g. Cain, Oakhill 1996; Yuill, Joscelyne 1988). The titles in this work have precise functions: specifying the place (*On the balcony; At Friar Lawrence's cell; The tomb*) or time (*Wedding day*) of the action; drawing attention to specific characters or roles (*Montagues and Capulets; Juliet; Uninvited guests;*

¹² <u>https://idioms.thefreedictionary.com/bite+(one%27s)+thumb+at</u>

ingue e

Where's Romeo?); but above all disclosing the gist of the section (e.g. A marriage arrangement; The commitment; Married; The fight; Trapped), the character' feelings (e.g. Bursting with excitement; Oh heavy day), or both gist and feelings (e.g. Disaster; A ray of light; A brave act).

Some differences can also be observed at the level of content. In fact, the teenage version – unlike the children's version – makes frequent reference to love and sex. Women are called *whore, slut*; sexual innuendos are extremely frequent and very easy to grasp (e.g. *Well, the Count will just have to take you in your bed. He'll wake you up with a big fright, won't he?*; Juliet's thought that [t]he darkness would hide her blushes when they made love). Even the pure and ingenuous love between Romeo and Juliet is described through physical images (e.g. *He shivered with the anticipation of touching her; Their bodies touched and they stayed like that, pressed against each other, for a long time before he spoke again)*. The author has clearly adapted the text to adolescents' specific interest in sexuality (see Sellinger Trites 2000; and James 2009).

From a more specifically linguistic or stylistic perspective, the most frequent type of difference is represented by the replacement of a word with a synonym. In fact, many are the cases where a simple word in the children's version corresponds to a more sophisticated word in the teenager's one (Table 4; the corresponding words are underlined).

	C-R&J	T-R&J
1	producing mouth-watering smells	conjuring mouth-watering aromas
2	this was the <u>big</u> moment	this was the <u>critical</u> moment
3	controlling himself <u>really well</u>	controlling himself <u>admirably</u>
4	An officer of the city's police force	An officer of the city's <u>Watch</u>
5	baked bread and frying <u>bacon</u>	baked breads and frying <u>hams</u>
6	disturbed the <u>peace</u>	disturbed the <u>tranquility</u>
7	his thoughts about Tybalt were not the	his thoughts about Tybalt were not the
8	his daughter was very <u>pretty</u>	his daughter was very <u>desirable</u>
9	What do you say to my <u>offer</u>	What do you say to my <u>proposition</u>
10	said	Intimated
11	from the <u>next</u> room	from the an <u>adjoining</u> room

Table 4Examples of lexical substitution.

Examples [1]-[3], [6], and [9]-[11] are cases in which a more common word was used in the C-R&J, compared to a less-frequent word in T-R&J. These – as well as examples [7] and [8] discussed below – are all instances of Klingberg's (1986) language adaptation strategy, differently applied to the



two age groups. Examples [4] and [5], on the other hand, are cultural context adaptations (Klingberg 1986): the terms used in the teenage version (*Watch*, and *hams*) presuppose some knowledge of medieval habits and institutions and are replaced in the child version with cultural equivalents in current British society. Finally, examples [7] and [8] illustrate a preference for concrete terms vs. abstract ones in the child and teenage versions, respectively, probably in keeping with Gamble and Yates's (2002) observation that figurative language may be problematic for younger readers.

The children's version is also characterized by greater explication (Table 5), achieved either by lengthier or clearer descriptions of a concept (examples 12-13) or by adding information (examples 14-18).

	C-R&J	T-R&J
12	buzzing with preparations for the day	humming with <u>activity</u>
	<u>ahead</u>	
13	the Montague men couldn't ignore	the Montagues couldn't ignore
14	Two of them, Gregory and Sampson,	Two of them stepped out
	stepped out	
15	He [] asked his friend, <u>Balthasar</u> .	He [] asked his friend.
16	caught sight of <u>his fifteen year-old</u>	caught sight of <u>Romeo</u>
	<u>cousin</u>	
17	You need to see her among a lot of	You need to make comparisons.
	other girls, and make comparisons	
18	'And I'll bite my thumb at them. If they	'And I'll bite my thumb at them. If they
	take that it will really show them up.' <u>It</u>	take that it will really show them up.'
	certainly would, because biting your	
	thumb at someone was the worst insult	
	you could give to another person.	

Table 5 Examples of explication in C-R&J.

The only observable syntactic difference between the two versions is the presence of a few elliptical sentences in the teenage text, and their total absence in the other text: e.g. *Not in love!* vs. *You're not in love; Not a penny!* vs. *I won't take a penny!*. This seems to suggest the author's awareness that ellipsis is difficult to understand among less-experienced readers (Gamble, Yates 2002), and is yet another example of how Klingberg's (1986) language adaptation strategy can be realized.

Finally, T-R&J includes long stretches of metaphorical or lyrical descriptions. Below are a few examples:

Then come on doomsday, sound your trumpet, because who is alive if those two are dead?'

She wished the god of the sun would whip his horses so that they would carry him faster to the distant west and allow night to fall like a thick curtain.

she had bought a house of love but not yet taken possession of it - she was like some brand new item that hadn't yet been used

He loved nature. He often thought about the soil - about the way that it encompassed the whole of life. It was a grave that took all life into itself when it died but it was also a mother, from which all new life sprang

These are absent in C-R&J, and rightly so, since, as we have seen in Section 2, figurative language is problematic for younger readers (Gamble and Yates 2002).

6. Concluding remarks

This work has investigated adaptation techniques in two versions of Shakespeare's *Romeo and Juliet*, respectively targeting later childhood and adolescence. To this aim, a range of corpus-analytical methods were used.

Comparison of the two narrative versions to specific sets of the BNC sampler (Section 5.1) has evidenced a range of stylistic devices adopted by the author to adjust the texts to the needs of a young audience. In both texts, adaptation revolved around the following strategies:

- Integrating ample stretches of dialogue into the narrative texture. This strategy meets two separate needs: it maintains some of the dialogic features of the original play, and also meets the young audiences' need for characters presented through dialogue and action (see Appleyard 1991).
- Conveying the idea of action and continuous changes of scene, linguistically marked by the presence of the verb *go*. Once again, this strategy not only contributes to maintaining features of the original play, but also to making the plot more dynamic, and thus more suitable to hold the reader's attention, a fundamental need for children (see Appleyard 1991).
- Adopting a range of explication strategies:
 - explicating relations between circumstances and events to facilitate comprehension of the characters' actions and decisions;
 - explicating the character's emotions by conveying them through concrete descriptions of their faces;
 - \circ explicating the character's distress by underlying it with

inauaaai

exclamations 'oh' or 'oh no';

 and preferring overt negative structures to hidden ones, the latter being difficult for children to understand (see Reid, Donaldson 1977).

Explication strategies go in the direction of limiting the need for children to make causal inferences, a difficult cognitive task for the less-skilled (see Long *et al.* 1997).

- Using a set of very common verbs and nouns with a frequency that is well above average. Implicitly, this entails avoiding low-frequency words that in children would hamper fluent reading and comprehension (see Nation, Snowling 1998).
- Repeating set expressions, in order to favour understanding by less-expert readers or by readers with limited vocabulary.
- Implementing strategies that help readers to identify with characters, such as emphasizing the young age of the characters (see Appleyard 1991).

Comparison of the two narrative versions to each other (Section 5.2) has evidenced differences between the two texts, connected to children's and teenagers' different needs. More specifically, in the children's version the author:

- Replaced acts and scenes with titled sections performing a range of specific functions, for the benefit of less-skilled readers (see Cain, Oakhill 1996; and Yuill, Joscelyne 1988).
- Added introductory descriptive paragraphs.
- Used simpler vocabulary (see Klingberg 1986).
- Preferred concrete terms to abstract ones (see Gamble and Yates 2002).
- Used lengthier or clearer descriptions of a concept and added information, for greater clarity.
- Limited the use of elliptic sentences (see Klingberg 1986; and Gamble, Yates 2002).

The teenage version on the other hand includes:

- Long stretches of metaphorical or lyrical descriptions.
- More technical vocabulary, sometimes even requiring knowledge of medieval habits and institutions.
- Frequent reference to women, love and sex, in order to meet adolescents' specific interest in sexuality (see Sellinger Trites 2000; and James 2009).

These language and culture context adaptation strategies are in perfect keeping with the affective needs and cognitive abilities of each age group as described in theoretical and empirical studies on children's literature and

developmental psychology (Section 2), which suggests that the author is indeed an experienced writer.

The analysis has illustrated how the concept of adaptation can be operationalized into in a range of structural, linguistic, and stylistic choices. This description of these strategies could be of use to authors writing or adapting texts for children and teenagers, but also to translators working in the area of children's literature.

From a methodological perspective, corpus linguistics methods of investigation have proven useful in highlighting features common to the two narrative versions under investigation. In particular, keyword analysis based on a statistical comparison between the text under investigation and one or more reference corpora - has made it possible to identify strategies that would not have been (easily) observable in other ways (e.g. preference of overt negative structures to hidden ones; explicating relations between circumstances and events), and has given evidence that the observed features are systematic rather than accidental. The corpus linguistics methods tested in this work, however, have proven almost useless in evidencing differences between the two versions, the reason being that the features that distinguish the children's from the teenage version - eventually observed using the Compare feature in MS Word - did not include semantic, lexical, or grammatical repetitions. It cannot be excluded, however, that differences of these types could be identified using other corpus methods (e.g. analysis of hapax legomena) or the application of specific tools, such as the English Regressive Imagery Dictionary.¹³

Finally, the structural, linguistic, and stylistic strategies described in this study, though representing a highly interesting set of devices, do not cover the entire gamut of resources and practices for adapting a text for young audiences. The academic and professional communities would certainly benefit from a systematic description of the viable adaptation strategies and devices and of their congruence with the needs of different age groups. Such a systematic description may be achieved with a corpus analysis of a much larger corpus of adaptations for children, including texts by several experienced authors.

¹³The English Regressive Imagery Dictionary (RID; <u>http://www.provalisresearch.com/wordstat/RID.html</u>) is composed of about 3,200 words and roots assigned to 29 categories of primary process cognition, 7 categories of secondary process cognition, and 7 categories of emotions. [...] These categories were derived from the theoretical and empirical literature on regressive thought. The dictionary allows to distinguish primordial or regressive thinking (in which images dominate over concepts) from those indicating secondary or conceptual thinking (which highlight understanding of logical continuity and relationships between concepts).

Bionote: Francesca Bianchi is researcher and lecturer at the University of Salento, Lecce, Italy. One of her primary research interests is corpus linguistics and how corpus linguistic methodologies can be applied to discourse and genre analysis. She has applied corpus methods to the analysis of academic papers, audiovisual material, and web material. Her studies have appeared in a range of national and international publications. She has published a volume titled *Culture, corpora and semantics*, a methodological investigation in the use of elicited data and Web data in the analysis of cultural specificities.

Recapito autore: francesca.bianchi@unisalento.it

jingue e o inguaggi

References

- Anthony L. 2013, *Developing AntConc for a new generation of corpus linguists*, Proceedings of the Corpus Linguistics Conference (CL 2013), July 22-26, 2013, Lancaster University, UK, pp. 14-16. <u>http://www.laurenceanthony.net/research/20130722 26 cl 2013/cl 2013 paper final.pdf</u> (14.01.2018).
- Appleyard J.A. 1991, *Becoming a Reader: The experience of fiction from childhood to adulthood*, Cambridge University Press, Cambridge.
- Archer D., Culpeper J. and Rayson R. 2009, Love 'A familiar of a devil'? An exploration of key domains in Shakespeare's comedies and tragedies, in Archer D. (ed.), What's in a word-list? Investigating word frequency and keyword extraction, Digital research in the Arts and Humanities, Ashgate, Farnham, pp. 136-157.
- Barone D.M. 2011, *Children's literature in the classroom. Engaging lifelong readers*, The Guilford Press, New York/London.
- Beckett S.L. 2009, *Crossover Fiction. Global and historical perspectives*, Routledge, New York/London.
- Bell A. 1986, Translator's notebook: delicate matters, in "Signal" 49, pp. 17-26.
- Cain K. and Oakhill J.V. 1996, *The nature of the relationship between comprehension skill and the ability to tell a story*, in "British Journal of Developmental Psychology" 14, pp. 187-201.
- Chapman J. 1987, Reading: from 5-11 years, Open University Press, Milton Keynes.
- Culpeper J. 2002, Computers, language and characterisation: An Analysis of six characters in Romeo and Juliet, in Melander-Marttala U., Ostman C. and Kyto M. (eds.), Conversation in Life and in Literature: Papers from the ASLA Symposium. Association Suedoise de Linguistique Appliquee (ASLA), 15, Universitetstryckeriet, Uppsala, pp. 11-30.
- Fischer-Starcke B. 2009, *Keywords and frequent phrases of Jane Austen's* Pride and Prejudice. *A corpus-stylistic analysis*, in "International Journal of Corpus Linguistics" 14 [4], pp. 492-523.
- Fischer-Starcke B. 2010, Corpus Linguistics in Literary Analysis. Jane Austen and her contemporaries, Continuum, London/New York.
- Gamble N. and Yates S. 2002, *Exploring children's literature*. *Teaching the language and reading of fiction*, Paul Chapman Publishing, London.
- Garside R. and Smith N. 1997, *A hybrid grammatical tagger: CLAWS4*, in Garside R., Leech G. and McEnery A.M. (eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora*, Longman, London, pp. 102-121.
- Gerbig A. 2010, Keywords and key phrases in a corpus of travel writing: from Early Modern English literature to contemporary 'blooks', in Bondi M. and Scott M. (eds.), Keyness in Texts, John Benjamins, Amsterdam/Philadelphia, pp. 147-168.
- Halliday M.A.K. 2004, An Introduction to Functional Grammar, Hodder Arnold, London.
- Hunt P. (ed.) 2005, Understanding children's literature, Routledge, London/New York.
- James K. 2009, *Death, Gender and Sexuality in Contemporary Adolescent Literature*, Routledge, New York/London.
- Kilgarriff A., Bušta J., Jakubíček M., Kovář V., Michelfeit J., Rychlý P. and Suchomel V. 2014, *The Sketch Engine: ten years on*, in "Lexicography" 1, pp. 7-36.
- Klingberg G. 1986, *Childrens' fiction in the hand of translators*, CWK Gleerup, Malmo. Lathey G. 2015, *Translating Children's literature*, Routledge, New York/London.



Rewriting *Romeo and Juliet* for a young audience. A corpus-assisted case study of adaptation techniques

- Long D.L., Oppy B.J. and Seely M.R. 1997, *Individual differences in reader's sentenceand text-level representations*, in "Journal of Memory and Language" 36, pp. 129-145.
- Lundin A. 2004, *Constructing the canon of children's literature. Beyond library walls and ivory towers*, Routledge, New York/London.
- Mahlberg M. 2010, *Corpus Linguistics and the Study of Nineteenth-Century Fiction*, in "Journal of Victorian Culture" 15 [2], pp. 292-298.
- Mahlberg M. and Smith C. 2010, Corpus Approaches to Prose Fiction: Civility and Body Language in Pride and Prejudice, in Busse B. and McIntyre D. (eds.), Language and Style, Palgrave, Basingstoke, pp. 449-467.
- Murphy S. 2007, Now I am alone: A corpus stylistic approach to Shakespearian soliloquies, in Gabrielatos C., Slessor R. and Unger J.W. (eds.), Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching, Vol. 1, <u>http://www.lancaster.ac.uk/fass/events/laelpgconference/papers/v01/Volume01.pdf#pag</u> <u>e=68</u> (14.01.2018).
- Nation K. and Snowling M.J. 1998, Semantic processing and the development of word recognition skills: Evidence from children with reading comprehension difficulties, in "Journal of Memory and Language" 39, pp. 85-101.
- Oittinen R. 2000, Translating for Children, Garland Publishing, New York/London.
- Rayson P. 2009, *Wmatrix: a web-based corpus processing environment*, Computing Department, Lancaster University. <u>http://ucrel.lancs.ac.uk/wmatrix/</u> (14.01.2018).
- Rayson P., Archer D., Piao S.L. and McEnery T. 2004, *The UCREL semantic analysis system*, in *Proceedings of the workshop on Beyond Named Entity Recognition Semantic labelling for NLP tasks in association with 4th International Conference on Language Resources and Evaluation (LREC 2004), 25th May 2004, Lisbon, Portugal*, pp. 7-12. <u>https://www.researchgate.net/publication/228881331_The_UCREL_semantic_analysis_</u> <u>system</u> (14.01.2018).
- Reid J. and Donaldson H. 1977, Reading: Problems and Practices, Ward Lock, London.
- Scott M. and Tribble C. 2006, *Textual Patterns: Key Words and Corpus Analysis in Language Education*, John Benjamins, Philadelphia.
- Shavit Z. 1986, *Poetics of children's literature*, The University of Georgia Press, Athens/London.
- Sellinger Trites R. 2000, *Disturbing the Universe: Power and Repression in Adolescent Literature*, University of Iowa Press, Iowa City.
- Stubbs M. 2005, Conrad in the computer: examples of quantitative stylistic methods, in "Language and Literature" 14 [1], pp. 5-24.
- van Coillie J. and Verschueren W. (eds.) 2006, *Children's Literature in Translation: Challenges and strategies*, St. Jerome Publishing, Manchester.
- Vandergrift K.E. 1990, *Children's literature. Theory, research and teaching*, Libraries unlimited, Englewood, Colorado.
- Yuill N. and Joscelyne T. 1988, Effects of organizational cures and strategies on good and poor comprehenders' story understanding, in "Journal of Educational Psychology" 80 [2], pp. 152-158.