

Shakespeare in Afrikaans: A corpus-based study of involvement in different registers of drama translation

Alet Kruger

Dept of Linguistics (Translation Studies), University of South Africa
krugea@unisa.ac.za

Abstract

This article draws on a corpus-based study of translations of Shakespeare in Afrikaans in order to contribute to the debate on what CTS (corpus-based translation studies) can offer the translation researcher. Corpus linguistic methodologies are versatile and therefore it was possible to apply aspects of Douglas Biber's (1988) multi-dimensional approach to register variation to a parallel corpus of Shakespeare's *The merchant of Venice* in Afrikaans. In particular, the focus is on an investigation of the manifestation of linguistic features of involvement in different registers of Shakespeare translation. 'Features of involvement' refer to those linguistic features which reflect that speaker and listener typically interact with one another (i.e., private verbs, contractions, 1st and 2nd person pronouns, analytic negation, demonstratives, emphatics, discourse markers, causative subordination, amplifiers, questions, time and place adverbials). The overall finding is one of a statistically highly significant difference between the two registers with the stage translation displaying more features of involvement than the page translation. Despite production constraints, the dialogue of a Shakespearean stage translation does indeed exhibit more features of involvement than a page translation.

Introduction

This article draws on a corpus-based study of two very different translations of Shakespeare's *The merchant of Venice* in Afrikaans. In particular, the focus is on an investigation of the manifestation of linguistic features of involvement. 'Features of involvement' refer to those linguistic features which reflect that speakers and listeners typically interact with one another in face to face communication (i.e., private verbs, contractions, first and second person pronouns, analytic negation, demonstratives, emphatics, discourse markers, causative subordination, amplifiers, questions, time and place adverbials). This is usually not the case in written texts unless the author actively strives to simulate

involvement on the part of the reader. The types of translation in this particular corpus are both drama texts written for spoken delivery; the question is whether, despite production constraints, the dialogue of a Shakespearean stage translation does indeed exhibit more features of involvement than a page translation.

What then, does involvement in dramatic dialogue entail and how was this particular research problem solved?

Involvement in drama translation

Ideally, every drama is written, and translated, to be performed. The dramatic text, as written text, addresses a context of performance which requires a change in the mode of discourse – the transformation and transmutation of the written lines into the dynamics of spoken speech, which involves more than the recitation of the lines of the text by actors (Herman 1995, 13). The dialogue in a drama text can either be written to be spoken as if not written when spoken by actors during a performance, or, written to be read as if heard, because in reading a play, the reader is aware of the fact that it is actually intended to be performed, and he acts out the play in his imagination (House 1981, 43, 172). This combination of the written and the spoken medium gives the drama its typical dual nature. As a result, it is possible to regard a drama as a written text and treat it as a literary text only, thereby ignoring its performance potential. Alternatively, it is also possible to treat a drama as a theatrical performance which can only be properly understood and evaluated in the theatre.

There is ample evidence that drama dialogue does not mirror ordinary conversation (cf. Short 1996, 169ff.; Short 1998, Simpson 1998, 41). As Herman (1998, 24) aptly puts it, dramatic speech is ‘tidied up speech’. It is ‘designed to simulate real-life, spontaneous language’ (House 1981, 172) and remains a simulation because it was carefully crafted and edited by the dramatist. However, it draws on aspects of real speech and therefore shows certain similarities such as instances of fragmentation, that is, spontaneous and non-fluent language, as well as ‘linguistic features of interaction and involvement’ (Biber & Finegan 1992, 691). Biber (1988, 43) defines *involvement* as those linguistic features which reflect the fact that speaker and listener typically interact with one another in face-to-face conversation, while writer and reader typically do not. In face-to-face communication, there is normally continuous non-verbal feedback between speaker and listener, e.g. nodding of the head or smiling. There may also be verbal feedback in the form of, for example *um* or *yes*. Owing to this interaction, speakers often refer directly to the listener by means of, for instance, second person pronouns and questions and they express their own thoughts and feelings by means of first person pronouns, affective forms such as emphatics and amplifiers, and cognitive verbs such as *think* and

feel. As a result, real speech often has a distinctly non-informational and imprecise character. This is usually not the case in written texts unless the author actively strives to simulate involvement on the part of the reader.

Traditionally, there are two kinds of drama translation, namely *page translations* and *stage translations* (cf. Kruger 1998; 2000). A Shakespearean page translation is usually a literary text meant to be read, and therefore it is fairly faithful, with minor omissions. On the contrary, a stage translation is usually translated exclusively for performance; the dialogue has been cut, is fairly free and idiomatic with instances of linguistic and cultural adaptation of foreign elements to make the play more accessible. When reading the script of the Afrikaans stage translation of Shakespeare's *The merchant of Venice* (Brown 1985) for the first time, it became obvious that the translator, Tjaart Potgieter (1991) has made every attempt to present the dialogue in this particular translation as ordinary, everyday talk. In contrast, this is clearly not the case in the published page translation of this play that was produced 42 years ago by D. F. Malherbe (1949). The question then was 'What makes a stage translation unusual?' In other words, what sort of textual features did the stage translator employ to simulate involvement between characters and make them sound more like real people in authentic situations? In what way might cuts have affected the occurrence of features that signal involvement in this particular translation?

This problem clearly concerns translational issues and therefore needs to be solved within the domain of translation studies. However, although descriptive translation studies (DTS) seems to accept a variety of methods of analysis, so long as they are empirical and descriptive, and also provides a broad methodological framework within which the comparison can be situated, it cannot provide the tools and techniques to solve the specific research problem stated above. In Kruger (2000) my hypothesis was that the stage translation of *The merchant of Venice* would reveal more spoken language features that signal involvement than the page translation. What was clearly required was an analytical tool that would not only enable the quantification of linguistic features of involvement in three Shakespeare texts (the original and two translations, totalling 59,869 words; Kruger 2000), but also provide a 'norm' of the occurrence of such features in authentic spoken English. Biber's (1988) multi-dimensional approach to register variation could be adapted to suit my purposes. Following Biber, who uses the terms *register* and *register variation* 'as cover terms for the full range of language varieties associated with differences in communicative situation' (Finegan & Biber 1994, 316), page and stage translations are regarded as different registers of Shakespearean drama translation. The section below provides a brief overview of the multi-dimensional approach to register variation, the tool adapted to analyse the linguistic features of involvement in my corpus of texts.

The multi-dimensional approach to register variation

The multi-dimensional approach to register variation, referred to here simply as the MD approach, was developed by Douglas Biber of the Northern Arizona University (US) in 1988. It was originally developed for comparative analyses of spoken and written registers in English (Biber 1988; 1995a; 1995b) and has subsequently been refined and used for cross-linguistic analyses of universal tendencies of register variation in non-western languages (cf. Biber & Hared 1992; also Biber & Finegan 1994). To the best of my knowledge it has not yet been applied to translated texts. Methodologically, the approach uses computer-readable text corpora, as well as computational tools to identify linguistic features in texts, and multivariate statistical techniques to analyse the co-occurrence relations among linguistic features, thereby identifying underlying dimensions of variation in a language. The primary aim of the MD approach is to provide comprehensive descriptions of patterns of register variation, that is, identifying underlying linguistic dimensions of variation and specifying linguistic similarities and differences among registers with respect to those dimensions (Biber 1995a, 19–20).

Biber (1988) found several functional differences between typical speech and writing which are associated with the typical situational characteristics of the two modes, for instance, some linguistic features are used to elaborate information in typical writing, while others are used to mark interaction or to express personal feelings in typical speech. However, no absolute distinctions were found between speech and writing. In terms of its situational characteristics, Biber (1988, 37) found that typical speech is interactive and involved, and depends on shared space, time and background knowledge; typical writing has the opposite characteristics. In terms of its linguistic characteristics, typical speech is structurally simple, fragmented, concrete and depends on exophoric reference; again, typical writing has the opposite characteristics. In line with an earlier study by Chafe (1982), Biber (1988, 21, 43) also found that linguistic variation must be analysed in terms of sets of co-occurring dimensions because they work together to mark some common underlying function. This is why he calls his approach ‘multi-dimensional’. Each dimension comprises an independent group of co-occurring linguistic features, and each co-occurrence pattern can be interpreted in functional terms such as ‘involved’, ‘informational’, and so on.

As my corpus was not tagged, the MD approach could not be followed in this respect. For my purposes, I also did not need to analyse every one of the 67 linguistic features of the MD approach and therefore selected only certain linguistic features from some of the dimensions that comprise the MD approach. I was mainly interested in linguistic features typically found in spoken language, so most of the features analysed were therefore taken from Dimension 1, labelled ‘Informational’ versus ‘Involved Production’ and defined by Biber (1988, 115)

as ‘discourse with highly informational purposes which is carefully crafted and highly edited’ as opposed to ‘discourse with interactional, affective, involved purposes, associated with real-time production and comprehension constraints’. In a diachronic study of two speech-like genres written for spoken delivery (i.e., dialogue taken from plays and from fiction), Biber and Finegan (1992, 689) themselves used only three dimensions of the MD approach.

The MD approach was subsequently adapted as follows to allow for the identification and analysis of linguistic features signalling involvement in drama texts. From Dimension 1 (Involved vs. Informational Production), the following features were chosen: private verbs, contractions, second person pronouns, analytic negation, demonstratives, **emphatics**, first person pronouns, causative subordination, discourse particles, **amplifiers** and questions. These features were chosen because they can be characterised as verbal, interactional, affective, fragmented, reduced in form and generalised in content (Biber 1988, 104) – exactly the kind of features that are found in dramatic dialogue. Type-token ratio was also chosen from this dimension to allow for the measurement of lexical variation or diversity in the corpus. From Dimension 3 (Elaborated vs. Situation-dependent Reference), time and place adverbials were chosen, as the dialogue in drama texts in particular features aspects of the *here* and *now* of a particular dramatic situation.

The first of the research procedures involved preparing the texts for automatic and semi-automatic analysis by means of WordSmith Tools. The names of the actors as well as the stage directions were manually deleted from each text so that only the *actual* dialogue remained. Thereafter, by means of the WordList Tool of WordSmith Tools, alphabetically arranged wordlists (and corresponding frequencies) for each of the texts were generated. The list of items of each linguistic feature supplied by Biber (1988, 221–245) was checked against the lists in Quirk et al. (1985) (the source to which Biber himself refers). This procedure assisted me in identifying those items that appear in *The merchant of Venice* and in recording their frequencies. Translation equivalents and synonyms appearing in similar contexts in the Afrikaans translations were subsequently manually identified from the frequency lists and their frequencies were also recorded. The recorded number of frequencies for each feature (e.g. private verbs) was added to obtain total counts of a particular linguistic feature per text.

Raw frequency counts cannot be used for comparison across texts because texts differ in length and such counts will give an inaccurate assessment of the frequency distribution of features. That is, long texts will tend to have higher frequencies simply because there is more opportunity for a linguistic feature to occur; in these cases, the higher frequency count does not indicate a more frequent use of the linguistic feature. Comparing the frequency per 100 or 1 000

words eliminates this bias (Biber 1988, 14; 75–76; Biber et al. 1998, 265–266). Following Biber (1988), all frequency scores were therefore ‘normed’ or converted to *densities*, that is, to a basis of 1000 words.

The drama translations examined here reveal a striking extent of sameness as regards production circumstances, primary purposes and interactiveness, but they also reveal some important differences. When comparing translations and their originals, one does not usually expect to find substantial differences because no matter how many translations there are, they all draw on the source text. Therefore, even small differences between texts are important and the chi square statistical test can direct the translation scholar to where the biggest differences lie. The chi square test¹ was therefore applied to compare the densities of the relevant features in the page and the stage translation (i.e., to establish levels of significance as well as rank ordering).²

Findings and interpretation

In order to establish the extent to which different registers of drama translation reveal differences in the manifestation of involvement, the Involved Production Hypothesis was formulated as a directional hypothesis to predict that a stage translation would reveal more features of involved production than a page translation. Its function was to test the extent to which two different registers of drama translation (page and stage translation) differentially reflect linguistic features of involvement. The underlying assumption is that because a stage translation has a different *function* from a page translation it is characterised by a higher degree of interaction, by higher effort at maintaining a relationship between speaker and addressee in speaking turns and exchanges.

Table 1 shows the frequency distribution of involved production features between the two registers as well as the computed χ^2 results. The linguistic features have been arranged from highest to lowest χ^2 values obtained and not according to Biber (1988, 102–103) who ranks linguistic features in various dimensions according to computations obtained by means of factor analysis. The actual scores as well as the densities/1000 for the source text are also supplied.

As these results show, the overall difference in frequency of involved production features between a page and a stage translation is statistically *highly significant*** ($p \leq .01$) and the null hypothesis (i.e., *There is no difference in the density of involved production features between a page and a stage translation*) can therefore be rejected. This finding lends support to the fact that translation register, that is, page or stage translation (the independent variables), does influence the use of linguistic features of involvement in drama translation.

In the case of five linguistic features (contractions, discourse markers, first per-

TABLE 1: Frequency distribution of involved production features between the source text and two drama translation registers

Linguistic features	Source text 20 908 words		Page translation 20 688 words		Stage translation 18 273 words		Computed x^2 results
	Actual scores	Density /1000	Actual scores	Density /1000	Actual scores	Density /1000	
Contractions	302	14.4	69	3.3	151	8.3	41.97**
Discourse markers	507	24.2	503	24.3	564	30.9	15.64**
First person pronouns	1496	71.6	1357	65.6	1349	73.8	10.17**
Analytic negation	220	10.5	205	9.9	234	12.8	7.31**
Emphatics	274	13.1	356	17.2	257	14.1	6.19**
Second person pronouns	873	41.8	753	36.4	743	40.7	4.78*
Private verbs	472	22.5	491	23.7	489	26.8	3.63*
Demonstratives	152	7.3	88	4.3	98	5.4	2.51
Causative subordination	6	0.3	72	3.5	51	2.8	1.46
Questions	206	9.9	203	9.8	198	10.8	1
Time and place adverbials	272	13.2	380	18.4	350	19.2	3.3
Amplifiers	114	5.4	80	3.9	76	4.2	2.1
Totals	4894	234.2	4557	220.3	4560	249.5	46.39**

son pronouns, analytic negation and emphatics) the findings were statistically *highly significant*** ($p \leq .01$). In the case of two features (second person pronouns and private verbs) the findings were statistically *significant** ($p \leq .05$). In the case of the other features (demonstratives, causative subordination, questions, time and place adverbials, and amplifiers) the findings were not statistically significant. A brief discussion will follow in the order of their ranking with respect to the x^2 values.

Contractions

According to Biber (1988, 243), 'contractions are the most frequently cited example of reduced surface form. Except for certain types of fiction, they are dispreferred in formal, edited writing; linguists have traditionally explained their frequent use in conversation as being a consequence of fast and easy production'. Research found that contractions are distributed as a cline (most frequently in conversation, least frequently in academic prose, with intermediate frequencies in, e.g. public speeches); more in American than in British writing; and that there is no absolute difference between speech and writing in the use of contractions (Biber 1988, 243). Thus, the use of contractions seems to be tied 'to appropriateness considerations as much as to the differing production circumstances of speech and writing' (Biber 1988, 243).

Contractions occur frequently in real spoken language and in informal writing, and it is therefore expected that they will also occur in texts that are written to be spoken such as drama texts and their translations. Biber (1988, 265) recorded densities ranging from 54,4 per 1000 for telephone conversations to 13,3/1000 for prepared speeches, showing that fewer contractions are used in more 'literate' texts. The results in Table 1 reveal a contractions density count of 8,3/1000 for the stage translation and only 3,3/1000 for the page translation. Although the stage translator worked closely with the producer and the actors, sitting in on every rehearsal, and felt that he did not have to indicate every contraction as the actors would automatically contract words when speaking, he nevertheless indicated certain contractions which are typical of ordinary, everyday spoken Afrikaans on the written play script.

Only the odd contraction is indicated in the page translation because the text is mainly meant to be read as a literary text. D. F. Malherbe, the page translator, was a renowned Afrikaans poet, novelist and playwright and very much concerned with the literariness of his translation. No doubt influenced by the Elizabethan spelling in the source text, 'which was more influenced by pronunciation and less by the original form and derivation of the word, than is now the case' (Abbott 1966, 339), and the prosody of the blank verse, certain forms were elided for the sake of rhythm. In comparison, the source text has a much higher density count (14,4/1000).

Discourse markers

According to Biber (1988, 241), discourse particles are used to maintain 'conversational coherence' and to monitor the information flow in involved discourse. I find Biber's (1988) view of discourse 'particles' restrictive because I take both the grammatical and the pragmatic function of different markers or signals that link discourse into account, and therefore, in agreement with Carstens (1997, 304–308) and Schiffrin (1987), who examines *oh, well, and, but, or, so, because, now, then* as discourse markers, have chosen to use the term *discourse markers* as an umbrella term. According to McCarthy and Carter (1994, 68, 85, 206, 207) discourse markers are generalised 'interactive markers' used to signal new segments of information in conversation – and certainly also in drama dialogue. These markers have 'important interpersonal and text-building functions' by taking, keeping and yielding the turn through speech actions, by empathising or communicating with the listener; and by structuring the message. Of the discourse markers listed by Biber (1988, 241), only *well* and *now* occur in *The merchant of Venice*. It was therefore necessary to identify those discourse markers that do occur in *The merchant of Venice* and the Afrikaans translations, to categorise these as (1) feedback words, (2) exclamations and interjections, (3) vocatives, and (4) courtesy adjuncts, and to record them.

The stage translator seems to have consciously created opportunities for interpersonal interaction and involvement between interlocutors by means of discourse markers, taking, keeping and yielding turns through speech actions, by empathising or communicating with the listener, and by structuring the dialogue. The results in Table 1 show a density count of 30,9 for the stage translation, 24,3 for the page translation and 24,2 for the source text. These density counts are all much higher than Biber's (1988, 265) highest finding for discourse particles – 6,6/1000 for telephone conversations – but it is important to mention that I identified and counted a far greater range of discourse markers than he did. For example, the stage translator inserted far more **feedback words** than the page translator in an attempt to make the dialogue sound more like real speech, for example '*Nou-ja, is Antonio hier?*' (Well then, is Antonio here?). One **interjection** that occurs consistently in the stage translation and which is more or less absent from the page translation is *Toe!*, meaning more or less come on/please/all right, a typical feature of spoken Afrikaans used to encourage someone else to say or do something, for example '*Toe, praat*' (Come on, speak).

The page translator in general used more formal, but dated, **exclamations** and **interjections** such as '*Mag God my bewaar van hulle*' (May God protect me from them) which fit in with his attempt to raise the status of Afrikaans as literary language back in 1949. In contrast, the stage translator used ordinary, everyday expressions such as '*Liewe hemel, skroef op jou kop*' (Good heavens, screw on your head; i.e., come to your senses).

Typically, **vocatives** in dramatic dialogue raise the level of personal involvement between interlocutors. Potgieter (1991) inserted extra vocatives in the stage translation to this effect as regards the following characters: Antonio, Bassanio, Nerissa and Portia. In the stage translation there are double the number of occurrences where *Portia* functions as a vocative compared to the source text and the page translation (Kruger 2000, 324).

The stage translator also obviously attempted to retain the Italian flavour of the Venetian setting of the original play and consistently made use of *signior*, *signiora* and *signiorina* as forms of address in his translation. The page translator does not use this term of address at all, and the word *signior(s)* only occurs three times in the source text itself (Kruger 2000, 325). By using the Italian loan words, Potgieter has avoided the more archaic forms of address used for male characters (*kêrel(s)/jongeheer/jonge* = chap/youngman/boy) that occur in the page translation. Also, by using the Italian terms of address (*signiora*, *signiorina*) and more ordinary terms of endearment for the female characters in the play (*liefling* = my love), he has avoided the more 'poetic' and 'literary' terms such as *my skone* (my beauty) and *edele dame/jonkvrou* (honourable lady/young lady = madam) which appear in the page translation. In contrast to the

formal term of address, *vader* (father), used in the page translation, Potgieter (1991) uses the informal *pa* (dad), typical of spoken Afrikaans. Most importantly, by using the term *sakeman* (businessman) as opposed to the more archaic *koopman* (merchant) in the stage translation, he has managed to give the play a thoroughly modern ring which is absent from the older translation.

The frequent occurrence of the **courtesy adjunct** *asseblief* (please), cleverly combined with a vocative, is noteworthy in the stage translation. This combination makes a request such as the following more personal and more involved as the speaker is directly appealing to the listener to say or do something: '*Asseblief, Bassanio, jy argumenteer met die Jood*' (Please Bassanio, you are arguing with the Jew).

First person pronouns

As regards density, the results in Table 1 show a count of 73,8 for the stage translation and 65,6 for the page translation. This result for the page translation is lower than that of the source text (71,6) which is higher than Biber's (1988, 265) finding of 70,7 per 1000 for telephone conversations. Biber (1988, 269) recorded a first person pronouns count of only 41,8 in prepared speeches, showing that interpersonal involvement is lessened in more 'literate' discourse. A drama (and also a translated drama) consists first and foremost of an *I* addressing a *you*, *here* and *now*. This means that any findings on first person pronouns have to be taken into consideration together with findings on second person pronouns, demonstratives, and time and place adverbials.

Analytic negation

Biber (1988, 245, following Tottie 1983a), sees analytic negation by means of the clause negator *not* (e.g. 'I will buy with you [...] but I will *not* eat with you'; 'Why fear *not* man, I will *not* forfeit it') as an alternative to the more integrative synthetic negation that occurs by means of negative forms and adverbs such as *no*; *none*; *neither*; *never*; *nor*; *nothing* (e.g. 'For Gratiano *never* lets me speak'). Biber (1988, 245) finds that analytic negation is associated with a fragmented presentation of information, resulting in a low informational density. There is twice as much negation overall in speech as in writing, which is ascribed to the greater frequency of repetitions, denials, rejections, questions, and mental verbs in speech.

Table 1 shows a density count of 12,8 for the stage translation and 9,9 for the page translation. The lower density count in the page translation could be ascribed to the fact that Malherbe (1949) did not always include the second *nie* of the double negative form in Afrikaans as he was constrained by the rhythm of

the blank verse in his translation. At times, this causes stiltedness and non-fluency in the dialogue.

Emphatics

According to Quirk et al. (1985, 583), emphatics (emphasisers) belong to the range of subjuncts concerned with expressing the semantic role of *modality*. These adverbials have a reinforcing effect on the truth value of a clause or part of the clause to which they apply. In addition to the force (as distinct from the degree) of a constituent, emphatics do not require that the constituent should be gradable. However, when the constituent emphasised is indeed gradable, the adverbial takes on the force of an intensifier, that is, as amplifier (see Amplifiers below). Quirk and Greenbaum (1973, 214) maintain that emphatics have ‘a general heightening effect’, because they heighten the expression of personal feelings and are ‘characteristic of informal, colloquial discourse, marking involvement with the topic’ (Chafe in Biber 1988, 241). General hedges (e.g. *sort of*) and general emphatics (*just, really*) are not explicit in their hedging or emphasising function (e.g. *really* might stand for any of the following more specific emphatics: *absolutely, certainly, intensely*). Although specific emphatics are more carefully chosen than general emphatics, they serve to emphasise the speaker’s or writer’s feelings and attitudes, and so can mark high personal involvement.

Table 1 shows a density count of 17,2 for the page translation and only 14,1 for the stage translation. This phenomenon could perhaps be ascribed to the fact that the page translation is more faithful to the source text. In general, the Afrikaans translator uses a wider range of emphatics than even Shakespeare in the source text, who, for instance, repeated the use of *certainly* (‘But Antonio is *certainly* undone’), *for certain* and *indeed* (‘Lorenzo *for certain*, and my love *indeed*’ Kruger 2000, 321). The source text shows a high frequency of *do/ did + verb*, e.g. ‘I *do beseech* you’; ‘I *did receive* fair speechless messages’. According to Abbot (1966, 216-217), this construction was used ‘in excited narrative’, however, it also abounds in dramatic dialogue where the speaker wants to emphasise the verb (cf. Quirk et al. 1985, 133) and was frequently used by Shakespeare. This kind of construction does not exist in Afrikaans as emphatic, the verb only is used in present tense or in past tense.

Second person pronouns

Similarly to first person pronouns, second person pronouns also generated a statistically significant degree of difference. As indicated in Table 1, a higher density count was recorded for the stage translation (40,7 vs. 36,4).

Private verbs

According to Biber (1988, 105), private verbs are used for the overt expression of private attitudes, thoughts and emotions. Biber (1988, 242) bases his concept and list of ‘private’ verbs on Quirk et al. (1985, 1181), who claim that ‘the “private” type of factual verb expresses intellectual states such as belief and intellectual acts such as discovery. These states and acts are “private” in the sense that they are not observable’.

Once an alphabetical frequency list was obtained for the source text, I could establish that quite a variety of private verbs appear in *The merchant of Venice* (in present tense, past tense or as past participle), for example, *accept, assume, believe, consider, determine, doubt, dream, expect, fear, feel, find, forget, hear, hold, hope, judge, know, learn, mean, note, observe, perceive, presume, prove, reason, remember, see, sense, show, signify, suppose, suspect, think, understand*. I compared this list with whatever private verbs I could manually pick out of the frequency lists obtained for the translations and then generated concordances with WordSmith’s Concord Tool for each individual entry so as to distinguish between verbs and nouns.

Table 1 reveals a density count of 26,8 for the stage translation and 23,7 for the page translation. Biber (1988, 105) found that private verbs generated the largest factorial weighting on Dimension 1 of his model of register variation, indicating interaction and involvement through a high expression of personal feelings, thoughts and emotions in discourse. He recorded density counts of 35,6 for telephone conversations, 35,4 for face-to-face conversations but only 17,6 for prepared speeches, showing that speakers tend to use fewer private verbs in more public discourse (Biber 1988, 265, 264, 269). In my corpus of texts we deal with contrived speech in drama texts where an *I* expresses his or her feelings, thoughts and emotions to a listening *you*. Playwrights generally do not employ more private verbs than deemed necessary to produce appropriate and meaningful communication otherwise the authenticity of the dialogue is questioned. It is therefore interesting that the findings in Table 1 not only reveal higher density counts but also a wider range of private verbs in the translations when compared to the source text.

Demonstratives

The demonstratives (*this, that, these* and *those*) have number contrast and can function both as determiners and pronouns. The general meanings of the two sets can be stated as ‘near’ and ‘distant’ reference (Quirk et al. 1985, 372), e.g. ‘What says *this* leaden casket?’; ‘To cut the forfeit from *that* bankrupt there!’; ‘God defend me from *these* two’; ‘*those* were Jacob’s’. Following Biber (1988, 241), demonstrative pronouns were excluded.

Quirk et al. (1985, 372) point out that demonstratives have definite meaning and therefore their reference depends on the context shared by speaker/writer and hearer/reader. They consider the use of demonstratives under the headings of ‘situational’ reference (i.e., reference to an entity outside the text/an exophoric referent), ‘anaphoric’ reference (co-reference to an earlier part of the discourse), and ‘cataphoric’ reference (co-reference to a later part of the discourse). The ‘near’ demonstratives can have both anaphoric and cataphoric reference, while the ‘distant’ demonstratives can have only anaphoric reference (Quirk et al. 1985, 375).

In drama, deictic reference presupposes the existence of a speaker referred to as *I*, a listener addressed as *you*, a physically present object indicated as *this/these* or *that/those*. In this respect the role of gesture is often crucial: ‘disambiguation of indexical expressions, especially demonstratives, frequently depends upon an accompanying and specifying kinesic indicator allowing the object of the deixis to be ostended [...] the language of the drama calls for the intervention of the actor’s body in the completion of its meanings’ (Elam 1980, 141–142).

The results in Table 1 show density counts of 5,4 in the stage translation and 4,3 in the page translation. The density counts for demonstratives in the translations neither exceed the density count recorded for the source text (7,3) nor the count of 11,1 that Biber (1988, 264) recorded for face-to-face conversations where speakers typically prefer demonstratives to articles (cf. Ochs in Biber 1988, 241).

Causative subordination

According to Biber (1996, 395), extensive use of subordination is associated with production constraints characteristic of speech. That is, subordination is apparently an important strategy for expressing fuller content under real-time production constraints, when there is little opportunity to elaborate content through exact lexical choices. *Because* is the only subordinator to function unambiguously as a causative adverbial (Biber 1988, 236). Other forms, such as *as*, *for*, *since* can have a range of functions, including causative (cf. Quirk et al. 1985, 1103–1107).

The findings in Table 1 reflect a density count of 3,5 in the page translation and 2,8 in the stage translation. Similarly to emphatics, causative subordination also generated a higher density count in favour of the page translation. It is interesting to note that the translations have higher density counts than the original which has a density count of only 0,3. Shakespeare used *because* to indicate causation in both the blank verse and the prose sections of the dialogue, e.g. ‘he tells me flatly there is no mercy for me in heaven, **because** I am a Jew’s daughter’ (III.v.25).

These findings lend support to Biber's (1988, 107) findings that causative subordination in general seems to be found more in speech than in writing, when there is little opportunity to elaborate through precise lexical choice. In drama texts, however, where we are dealing with contrived speech and where the texts were carefully edited after production, the use of causative subordination is not that frequent. To an extent, the Afrikaans translators were constrained by the use of *because* as causative adverbial in the source text, but they nevertheless inserted *want* (because) or *omdat* (because) whenever they wanted to elaborate implicit meanings in the source text, or when they wanted to make implicit links indicated by means of punctuation marks in the source text, explicit.

Questions

According to Biber (1988, 227), questions, like second person pronouns, 'indicate a concern with interpersonal functions and involvement with the addressee'. I therefore included all questions (including *yes/no* questions excluded by Biber 1988, 227) in my corpus. They were identified by means of concordances of a combination of wildcard symbols and the question mark ('*?'). The findings in Table 1 reveal a density count of 10,8 in the stage translation and 9,8 in the page translation, that is, the two translators employed very much the same number of questions.

Time and place adverbials

Table 1 shows density counts of 19,2 in the stage translation and 18,4 in the page translation. Although there is very little difference in the use of time and place adverbials in the two registers, the two translations reveal a wider range of time and place adverbials than the source text which generated a density count of only 13,2. Biber (1988, 102–103) found that time and place adverbials scored negative factorial weights in his analysis because they depend on referential inferences by the addressee as regards text-internal and text-external references. In a drama, however, deixis allows the dramatic context to be referred to as an 'actual' and dynamic world in progress. What we have are references by the speakers to themselves as speakers, to their interlocutors as listener-addressees and to the spatio-temporal coordinates (the here-and-now) of the utterance itself by means of such deictic elements as demonstratives and spatial and temporal adverbs (Elam 1980, 139). Except for first person pronouns and second person pronouns which generated higher density counts in favour of the stage translation, there is not a big difference between demonstratives and time and place adverbials in the two registers.

Amplifiers

According to Quirk et al. (1985, 589), the intensifier subjuncts are broadly

concerned with the semantic category of degree. For them, the term *intensifier* does not refer only to means whereby an increase in intensification is expressed. Rather, an intensifying subjunct indicates a relatively low or high point on an abstractly conceived intensity scale. The scale is seen as applying to a predicate or to some part of a predicate, such as the predication, the verb phrase, or even an item within the verb phrase. The verbs in question are largely expressive of attitude – which is exactly the kind of verbs found in the dialogue of a drama text. Quirk et al. (1985, 590) distinguish between two subsets of intensifiers, namely **amplifiers**, which scale upwards from an assumed norm, and **downtoners**, which have a lowering effect on the force of the verb or predication. Intensification is realised mostly by adverbs, but occasionally also by noun phrases and prepositional phrases. Downtoners were not taken into account as I did not take into account any of the linguistic features comprising Factor 7 of Biber's (1988, 102–103) factorial structure.

Quirk et al. (1985, 590) divide **amplifiers** into (a) **maximisers**, which can denote the upper extreme of the scale, and (b) **boosters**, which denote a high point on the scale. I followed Biber (1988) and did not make this distinction. According to Biber (1988, 241), 'the relation between emphatics and amplifiers is similar to that between hedges and downtoners: emphatics simply mark the presence (versus absence) of certainty while amplifiers indicate the degree of certainty towards a proposition'. Biber (1988, 240) states that 'amplifiers boost the force of the verb, indicating in positive terms, the reliability of propositions. They can signal solidarity with the listener in addition to marking certainty or conviction towards the proposition.' Amplifiers such as the following were identified in the ST: *entirely, heartily, much, quite, strongly, very, well, an infinite deal of*.

The findings in Table 1 show density counts of 4,2 for the stage translation and 3,9 for the page translation. It is interesting to note that although the source text generated the highest density count (5,4), the two translations have a wider range of amplifiers than the source text (Kruger 2000).

Type-token ratio

The standardised type-token ratio (i.e., a running average based on consecutive 1000-word chunks of text), generated by WordSmith Tools in Table 2 is highest in the source text (40,93) and shows that Shakespeare made use of many different lexical items or 'types' in the original play. This more varied vocabulary in the source text reflects extensive use of different words that have very specific meanings.

TABLE 2: Differences in type/token ratio between source text and translations

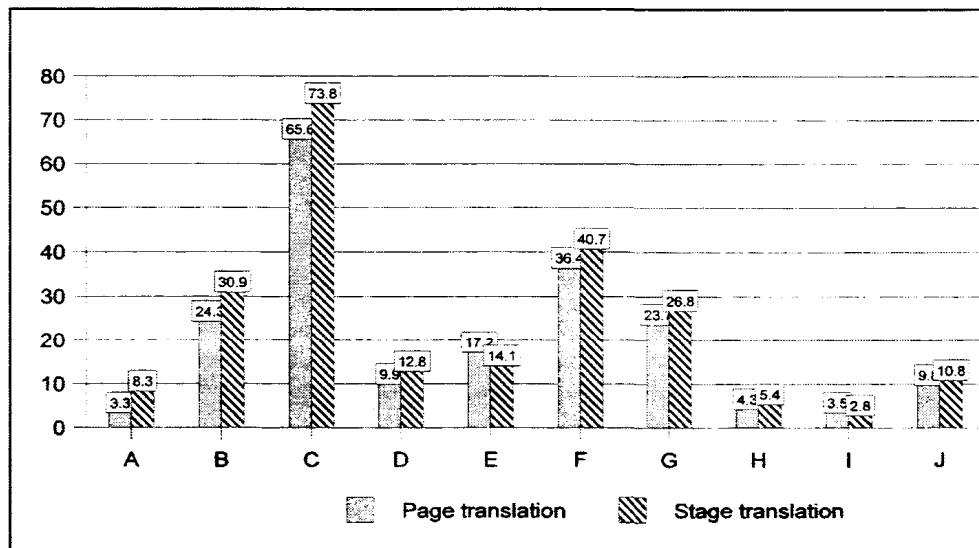
	Source text	Page translation	Stage translation
Number of tokens	20 908	20 688	18 273
Types	3 278	3 052	2 562
Type/token ratio/100	157	148	140
Standardised type/token ratio per 1000 words	40.93	38.46	35.81

The texts analysed here are all drama texts and the dialogue was carefully crafted and edited to simulate real-life conversation. It is therefore acceptable that all the texts reveal a fairly high type-token ratio. This is not the case in real conversation and unplanned speech which is characterised by low information focus and low lexical variety because 'precise lexical choice is a very difficult production task and is thus rarely accomplished in speech' (Chafe & Danielewicz in Biber 1988, 104).

The lower standardised type-token ratio generated by the stage translation, that is, 35,81 compared to 38,46 in the page translation, could be ascribed to the fact that the stage translator consciously attempted to make the dialogue more authentic by simplifying it, that is, repeating lexical items used before rather than introducing new forms and omitting sections, thus moving towards the 'oral' rather than the 'literate' pole on this continuum.

The results of Table 1 are graphically presented as densities per 1 000 as follows in Figure 1.

Overall, the stage translation generated higher density counts for each of the involved production features analysed, the only exception being emphatics (E in Figure 1) and causative subordination (I in Figure 1). These results differ mainly from Biber's (1988) findings in terms of ranking order: for example, the biggest **difference between the page and stage translation** was revealed in respect of **contractions** (A in Figure 1), and not private verbs (G in Figure 1). This statistically highly significant result shows that contracting pronouns and auxiliaries, and eliding words on the play script is the most common way of imitating real conversation in dramatic dialogue. In contrast, the fact that the page translator hardly indicated any contractions shows that he obviously intended for his text to be read and appreciated as a literary text. The second statistically highly significant finding as regards the **difference between the page and the stage translation** concerns **discourse markers** (B in Figure 1). The high density count and wide range of discourse markers displayed by the stage translation revealed that this is an easy and efficient way to signal involvement

**Key:**

- | | | |
|----------------------|------------------------------|----------------------------|
| A: Contractions | B: Discourse markers | C: 1st person pronouns |
| D: Analytic negation | E: Emphatics | F: 2nd person pronouns |
| G: Private verbs | H: Demonstratives | I: Causative subordination |
| J: Questions | K: Time and place adverbials | L: Amplifiers |

FIGURE 1: Density distribution/1000 of involved production features in drama translation

and interaction in dramatic dialogue. The use of ordinary, everyday expressions ensured that the characters speaking on stage will sound more authentic.

Conclusion

The main finding presented here is that drama translation register (page or stage translation) does have a constraining effect on involvement in drama texts. The overall finding was one of a statistically highly significant difference between the two registers with the stage translation displaying more features of involvement than the page translation. Although this finding is by no means earth-shattering, in essence it supports Biber and Finegan's (1992, 695) finding that both dialogue in fiction and dialogue in drama follow 'the same general pattern of drift towards more oral styles', as was the case with three written registers (fiction, essays and personal letters) analysed in an earlier study by Biber and Finegan in 1989. The features of involvement analysed here cluster together sufficiently to reveal that in comparison with an Afrikaans page translation of a Shakespeare play, a recent stage translation displays a definite tendency towards a more oral, more involved and more situated style, reflecting no doubt a general modern trend towards

creating more appropriate and accessible texts. The dialogue in a Shakespearean stage translation is more speakable than that of a page translation precisely because it comprises more spoken language features.

Various translation scholars all over the world have begun to apply corpus linguistic methodology to translated text in an attempt to study translation as a textual phenomenon *per se*. However, the application of corpus linguistic tools to a corpus of translated Shakespeare texts in order to examine and describe the relationship between register and involved production features is new and contributes to our understanding of the textual features that distinguish stage from page translations. In turn, this research can also benefit descriptive linguistics, notably comparative stylistics and register variation studies, by bringing new perspectives to these 'older disciplines' from the vantage point of translation studies. As remarked by Laviosa (2004, this volume), the combination of descriptive translation studies and more rigorous linguistic tools and analytical frameworks from corpus-based translation research can only benefit translation studies as discipline because corpus linguistics provides the researcher with the tools to study large bodies of language, the raw material of any kind of translation, and therefore 'no theorizing about translation, and certainly no training for translators, can proceed on the basis of dismissing this essential component of translation' (Baker 1996, 16–17).

Notes

- 1 'A statistical calculation used to test how well the distribution of a set of observed data matches a theoretical probability distribution. The calculated value is equal to the sum of the squares of the differences divided by the expected values.' Microsoft® Encarta® Reference Library 2004. © 1993–2003 Microsoft Corporation.
- 2 The chi square (χ^2) statistical test could be applied in standard fashion because it was possible to identify both the number of occurrences of the linguistic features being focused on and also to set them against the number of remaining words in the relevant text, the *residual* (cf. Erickson & Nosanchuk 1983, 69). The features in focus and the residual can then be displayed in a *two-by-two contingency table* and the χ^2 formula applied to the various cells of the table as exemplified below for the findings on contractions (cf. Erickson & Nosanchuk 1983, 247ff.; Butler 1985, 112ff.; Mulder 1993, 158ff.):

A	B	A+B
C	D	C+D
A+C	B+D	A+B+C+D=N

$$\text{then } \chi^2 = \frac{N(AD-BC)^2}{(A+B)(C+D)(A+C)(B+D)}$$

In the case of contractions we can see that the final finding was statistically highly significant:

	Stage translation	Page translation	TOTALS
Contractions	69	151	220
Other words/ residual	20619	18122	38741
TOTALS	20688	18273	38961

$$\begin{aligned} \text{then } \chi^2 &= \frac{38961(69 \cdot 18122 - 151 \cdot 20619)^2}{(220)(38741)(20688)(18273)} \\ \chi^2 &= \frac{135232034713324000}{3221972796588480} \\ \chi^2 &= 41.97^{**} \end{aligned}$$

References

- Abbot, E. A. 1966 [1870]. *A Shakespearean grammar*. New York: Dover Publications.
- Baker, M. 1996b. Linguistics and cultural studies: Complementary or competing paradigms in translation studies? In *Übersetzungswissenschaft im Umbruch: Festschrift für Wolfram Wilss*, ed. A. Lauer, H. Gerzymisch-Arbogast, J. Hallo and E. Steiner, 6–19, Tübingen: Gunter Narr.
- Biber, D. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.
- . 1995a. *Dimensions of register variation*. Cambridge: Cambridge University Press.
- . 1995b. On the role of computational, statistical, and interpretive techniques in multi-dimensional analyses of register variation: A reply to Watson. *Text* 15 (3): 341–370.
- Biber, D., S. Conrad and R. Reppen. 1998. *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D. and E. Finegan. 1992. The linguistic evolution of five written and speech-based English genres from the 17th to the 20th centuries. In *History of Englishes: New methods and interpretations in historical linguistics*, ed. M. Rissanen, O. Ihalainen, T. Nevalainen and I. Taavitsainen, 688–704, New York: Mouton de Gruyter.
- . 1994. *Sociolinguistic perspectives on register*. Oxford: Oxford University Press.
- Biber, D. and M. Hared. 1992. Dimensions in register variation in Somali. *Language Variation and Change* 4:41–75.
- Brown, J. R. ed. 1985. *The merchant of Venice* (New Arden Shakespeare.) London: Methuen.
- Butler, C. 1985. *Statistics in linguistics*. Oxford: Basil Blackwell.
- Carstens, W. A. M. 1989. *Norme vir Afrikaans: Enkele riglyne by die gebruik van Afrikaans*. Pretoria: Academica.
- Chafe, W. L. 1982. Integration and involvement in speaking, writing, and oral literature. In *Spoken and written language: Exploring orality and literacy*, ed. D. Tannen, 35–54, New Jersey: ABLEX Publishing Corporation.
- Elam, K. 1980. *The semiotics of theatre and drama*. London: Methuen.

- Erickson, B. H. and T. A. Nosanchuck. 1983. *Understanding data*. Milton Keynes, UK: The Open University Press.
- Finegan, E. and D. Biber. 1994. Register and social dialect variation: An integrated approach. In *Sociolinguistic perspectives on register*, ed. D. Biber and E. Finegan, 315–349, Oxford: Oxford University Press.
- Herman, V. 1995. *Dramatic dialogue: Dialogue as interaction in plays*. London: Routledge.
- Herman, V. 1998. Turn management in drama. In *Exploring the language of drama*, ed. J. Culpeper, M. Short and P. Verdonk, 19–33, London: Routledge.
- House, J. 1981. *A model for translation quality assessment*. Tübingen: Gunter Narr.
- Kruger, A. 1998. Shakespeare translations in South Africa: A history. In *Translators' strategies and creativity*, ed. A. Beylard-Ozeroff, J. Králová and B. Moser-Mercer, 107–115, Amsterdam: John Benjamins.
- . 2000. Lexical cohesion and register variation in translation: *The Merchant of Venice* in Afrikaans. Unpublished D Litt et Phil thesis. University of South Africa, Pretoria.
- Laviosa, S. 2004. Corpus-based translation studies: Where does it come from? Where is it going? (This volume.)
- Malherbe, D. F. 1949. *Die koopman van Venesië*. Johannesburg: Afrikaanse Pers-Boekhandel.
- McCarthy, M. and R. Carter. 1994. *Language as discourse: Perspectives for language teaching*. New York: Longman.
- Microsoft[®] Encarta[®] Reference Library 2004. 8 1993–2003 Microsoft Corporation.
- Mulder, J. C. 1993. *Statistical techniques in education*. Pretoria: HAUM Tertiary.
- Potgieter, T. 1991. 'Die sakeman van Venesië.' Unpublished manuscript.
- Quirk, R. and S. Greenbaum. 1980. *A university grammar of English*. London: Longman.
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Schiffrin, D. 1994. *Approaches to discourse*. Oxford: Blackwell.
- Short, M. 1996. *Exploring the language of poems, plays and prose*. London: Longman.
- . 1998. From dramatic text to dramatic performance. In *Exploring the language of drama*, ed. J. Culpeper, M. Short and P. Verdonk, 6–18, London: Routledge.
- Simpson, P. 1998. Odd talk: studying discourses of incongruity. In *Exploring the language of drama*, ed. J. Culpeper, M. Short and P. Verdonk, 34–53, London: Routledge.