Orality and noun phrase structure in registers of British and Kenyan English

Diana Hudson-Ettle
Bayreuth University Language Centre

Tore Nilsson
Dalarna University College

1 Introduction

Orality has become the focus of many corpus studies in recent years, especially since the publication of Biber's major contributions to that field (cf Biber 1988). One important finding has been that texts are not easily classified into an oral vs written dichotomy. Rather, it is more adequate to regard texts as aligned on a continuum, with informal conversations and formal academic prose representing the extreme ends of this continuum. The present paper will address several questions which are related to the orality debate:

- In how far can we find a correlation between noun phrase complexity and orality?
- In particular, in how far are (the frequency of) adjectival premodication and text type related?
- Do we get different results if we analyse texts written by speakers of English as a second language (ESL), eg English as it is used in Kenya, instead of British English?

In addition, we will address questions of corpora comparability. The corpora we have selected have not been compiled with a view to making them comparable to each other; hence we must ask which of our results are due to 'real' variation within and between text types/English varieties, and which of our results are due to our corpora being skewed in one way or another.

1.1 Hypotheses

Varantola (1984), Jucker (1992), and Nilsson (2001) say that variation in the structure and frequency of noun phrase premodification is partly a linguistic

reflection of the level of technicality of the subject matter (specialized vs general language), and partly due to the formality of the text. We intend to show that the distribution of one type of noun phrase premodification, namely adjectival premodifiers, differs also with respect to the degree of orality in the text. We will do this by placing the texts examined on an orality continuum based on extra-linguistic criteria and by showing that the position of each text type on this continuum correlates with the frequency of adjectival premodification in noun phrases.

For the purposes of this investigation we have chosen the following four text types: social letters, creative writing, reportage feature articles and institutional editorials from a corpus of English from East Africa (International Corpus of English, East African component, henceforth ICE-EA), specifically from Kenya. Intuitively, one would classify the first two text types as more 'oral' in their nature than the last two text types. For Kenyan English in particular, Hudson-Ettle (1998) has evidenced that creative writing shows more 'oral' features than other written Kenyan text types. In its syntactic complexity the creative text type was found to be similar to spoken data in that it has the shortest units, the highest number of simple clauses and fewest clauses with two or more subordinate structures.

We will analyse the four text types mentioned above according to the features discerned by Biber (1988) and others for classifying text types along the orality continuum and check whether our intuitions about the orality of the text types are confirmed. We will then analyse all four text types with respect to the frequency of adjectival premodifiers in noun phrases. What we expect to find is that social letters and creative writing, as the more 'oral' text types, contain less adjectival premodification than the more formal text types of reportage feature articles and institutional editorials.

As a second step, we will have a closer look at differences between British and Kenyan English. Certain grammatical features in the English of Kenya deviate from the British norm, such as the use of the definite and indefinite articles, the expansion of the progressive form to stative verbs, non-count nouns in plural form, and preposition usage (Hudson-Ettle 1990 specifically for English in Kenya, and Schmied 1991: 64 for 'African English'). Kachru (1982: 46) also lists similar characteristics of grammatical deviation in 'African varieties of English'. Hudson-Ettle (1998) undertook a comparative analysis of the syntactic structures of specific text types of the Kenyan component of ICE-EA and found differences from studies based on British English. To date there has been no comparative study of noun phrase premodification.

In the context of the orality debate, we hypothesise that English in Kenya (and perhaps English as a second language, in general) is more 'oral' in its characteristics than the Standard British variety. Kachru (1982: 46) includes in his description 'a tendency to use complex sentences', Zuengler (1982: 117) lists expressions which would be considered 'archaic or overly formal' as common in Kenyan English, and Schmied (1991: 53) claims that 'African speakers of English tend to reproduce characteristics of written English even in the spoken form'. Hudson-Ettle (forthcoming) shows, however, that corpus analysis provides evidence that *that*-clauses are more frequent in the Kenyan written texts than in the other written databases of English analysed and suggests that these might be preferred to more complex syntactic strategies. We would like here to investigate whether the frequency of adjectival premodification is another indicator of structural elaboration and consequently a linguistic feature relevant to the issue of orality. For the present investigation we will look again at the Kenyan reportage feature articles and compare them with British feature articles.

1.2 The Kenyan data

ICE-EA consists of approximately one and a half million words of spoken and written English from Kenya and Tanzania collected between 1991 and 1996. The text types in the spoken component include dialogues and monologues, from the private and public spheres, ranging from informal conversations, classroom discussions, broadcast interviews and discussions to news broadcasts and speeches. The written component consists of published/printed and non-printed data, ranging from social and business letters, examination essays and legal transcriptions to popular writing (mainly from newspapers), editorials, academic writing, and fiction. An additional category – written as spoken – contains transcriptions of parliamentary and legal procedures.

The ICE-EA subcorpus analysed for the present paper was taken solely from the Kenyan component because, as opposed to Tanzanian English, the English used in Kenya is a 'true' ESL variety (Platt et al 1984). English serves in Kenya as a lingua franca between people of various ethnic groups, particularly in the bigger towns and the capital, Nairobi, which is why it was possible to acquire social letters in English here (but not in Tanzania where Kiswahili fulfils this function). In secondary education English is the medium used, and the language acquisition process is continued outside the classroom when it is used to communicate. It is also passively acquired through reading the newspapers and watching television.

For the present study we selected four different subcorpora from the Kenyan corpus, which will be described in more detail here. The first subcorpus we have

focused on consists of social letters. Students in Kenya were willing to furnish the collectors with copies of private letters they had received. These letters are however very very short and rarely written by one author. Rather than grouped together artificially as texts of 2,000 words, they are in one file but named individually as W1B-SK01 to W1B-SK50 (for Kenyan Social). They were all collected in 1996 and consist altogether of 22,079 words.

The second subcorpus of interest consists of creative writing. ICE-EA contains two genres of creative writing – novels and short stories. For the present paper the short stories (W2F11K to W2F020K) were analysed. These were published between 1990 and 1992 and add up to 20,086 words altogether.

The third subcorpus we have selected consists of newspaper reportage. From the start, we decided to divide the ICE-EA newspaper reportage data into three types, first-page splash articles, feature reports and so-called brief texts (cf Schmied and Hudson-Ettle 1996), the latter constituting an additional specialized corpus for each country but not included in ICE-EA (extended corpus material). Several texts written by the same journalist were taken as subtexts to form a text of 2,000 words. The Kenyan reportage articles (W2C011K-W2C020K) were published between 1990 and 1994 and total 20,098 words.

The fourth subcorpus selected for this paper contains persuasive writing. The persuasive category was, according to ICE, to contain ten press editorials. An individual editorial does not consist of 2,000 words but, again, texts from the same newspapers and for the most part from the same year were grouped together, making a total of 20,041 words (W2E001K-W2E010K). Attention was also paid to representing a variety of newspapers. The preponderance of texts from the *Daily Nation* and *Sunday Nation* compared to those from the *Kenya Times*, *Sunday Times*, *The Standard*, and *Weekly Mail* is indicative of the dominance of these two papers in Kenya. Note that in East Africa, however, another persuasive category plays an important role linguistically and socially, which we refer to as 'personal columns' in contrast to the 'institutional' editorials. This study is, however, concerned only with institutional editorials. We will not take 'personal columns' into account because we wanted to keep our East African data as comparable to our British data as possible.

1.3 The British data: Feature articles in UPC and UCTT

In addition to the text types taken from the Kenyan part of the ICE-EA outlined above, the present study includes two sets of British feature articles taken from The Uppsala Press Corpus (UPC)[1] and the Uppsala Corpus of Travel Texts (UCTT). These two components are included in the study to enable comparison partly of Kenyan and British feature articles (regarding orality and adjectival

noun phrase premodification), and partly to investigate issues of corpus comparability. The feature articles represented in UPC deal with various topics (see below), whereas the topic (subject matter) is controlled in UCTT. Consequently, it will be possible to draw certain conclusions regarding the influence of subject matter on the linguistic patterns attested.

UPC consists of newspaper texts from 1994, sampled in such a way as to form as similar as possible a collection of newspaper language to that used in LOB, but with more recent material (for a full presentation of UPC, see Westergren Axelsson 1998: 21ff, 226ff). In the following, the relevant components of UPC are presented in overview. UPC contains samples from three of the major text categories included in LOB, namely (A) Reportage, (B) Editorials, and (C) Reviews. For the purposes of the present study, the first 20 samples (A01-20) of the Reportage component have been used, totalling 20,622 words. The topics covered in the component selected for this investigation are Political (six samples), Sports (two samples), and Society (two samples). The newspapers represented in the selection are *The Sun* (one sample), *Daily Mail* (three samples), *The Times* (one sample), *Morning Star* (one sample), the *Daily Telegraph* (two samples), and *Daily Express* (two samples). The selection thus covers different topics, and 'broadsheet' as well as 'tabloid' newspapers are included (cf Jucker 1992: 47ff).

The Uppsala Corpus of Travel Texts (UCTT) was compiled between 1996 and 1998. It consists of three types of texts with a common subject matter, namely travel and tourism. UCTT has three components: (A) tourist brochures, (B) newspaper articles, (C) travel guide books. The total number of words in the corpus amounts to ca 500,000. For the purposes of the present study, a sub-section of the articles, component (B), consisting of 15 samples (complete articles) comprising a total of 20,511 words has been used.[2] These articles all deal with accounts of visits to various places or stays at different hotels. The texts are highly personal in tone, and apart from descriptions of the places visited, the articles are also characterized by a narrative style and contain personal reflections of the places visited.

2 Orality

2.1 Spoken language vs written

In comparative linguistic studies of samples of speech and writing starting in the late fifties, speech has been found to contain fewer words per sentence, fewer syllables in each word, more words of one syllable, with more references to people, fewer attributive adjectives and less variety in the lexicon. The twenty-five

most frequently used words were found to be very similar for the oral and written samples, nineteen being on both lists (Gruner et al 1967: 450). They included determiners, demonstratives, pronouns and the conjunction *and*. Their frequency varied according to medium, however, there being, for example, more personal pronouns in the spoken than in the written samples. Neither nouns, descriptive words, adjectives nor adverbs were represented in the list. DeVito's (1967) analysis of lexical items revealed that speech contained more words referring to the speaker (eg *I*, *my*), more indefinite quantifying words (*much, many, some*), more allness terms (eg *all*), more qualification terms (*if*, *but*) and more indications of opinion (*seems, appears*). These are some of the features indicating orality that we will be looking at in the analysis.

Halliday (1979) comes to the conclusion that each medium of language is complex in its own way. The complexity of written language is lexical, while that of spoken language is grammatical. Grammatical items such as determiners, pronouns, most prepositions, conjunctions, some classes of adverb, and finite verbs are more frequent in spoken texts, while there is a higher number of lexical words in written texts. Halliday locates the difference between written and spoken language as being one of intricacy, 'the intricacy with which the information is organised' (Halliday 1989: 62). In spoken texts grammatical intricacy takes the place of lexical density, measured as the number of lexical items per clause. Later Halliday writes:

> Typically, written language becomes complex by being **lexically dense**: it packs a large number of lexical items into each clause; whereas spoken language becomes complex by being **grammatically intricate**: it builds up elaborate clause complexes out of parataxis and hypotaxis.
>
> (Halliday 1994: 350)

Chafe (1982: 37) recognises the fact that we can 'integrate a succession of ideas into a single linguistic whole' in writing (and follow these in reading) but that this is not possible in speech where our ideas are expressed separately and follow one another. He suggests that this is why written language is more complex and has an 'integrated quality which contrasts with the fragmented quality of spoken language' (1982: 38). In addition to the stringing together of 'idea units' (Chafe's units of analysis), without connectives, the high frequency of certain conjunctions (*and*, *but*, *so*, *because*) is seen to be evidence of fragmentation[3]. The integrated quality of written language is indicated by the higher frequency

of nominalizations, prepositional phrases usually with *of*-genitive subjects and objects, present and past participles, attributive adjectives, conjoined verb, adjective or noun phrases, and chains of prepositional phrases, *that*[4] and to complement clauses and relative clauses.

Another major difference between spoken and written language is determined by the addressee/audience. Whereas speakers 'interact with their audiences, writers do not' (Chafe 1982: 45). Chafe introduces the concepts of detachment and involvement, the former being a quality of the writer, the latter of the speaker. Devices employed to indicate distance or detachment are the passive voice and nominalizations. Involvement is seen where there are more first person references and references to the speaker's mental processes *(I think)*, and colloquial expressions such as *you know*, *I mean*, which indicate a monitoring process on the part of the speaker (1982: 47). Other indicators are the emphatic particles *just* and *really* and expressions of vagueness and hedging (*and so on*, *sort of*) as well as direct quotes.

2.2 A continuum rather than a dichotomy

In contrast to most previous studies which had suggested dichotomous distinctions rather than a continuum, Biber's notion of the dimension permits description of a 'continuous range of variation' (Biber 1988: 23). Quantitative methods are employed to determine the distribution of linguistic features in texts. Frequency counts of specific linguistic features 'give an exact, quantitative characterisation of a text so that different texts can be compared in very precise terms' (1988: 13). But, he points out, these counts cannot identify linguistic dimensions. To determine a linguistic dimension a 'consistent co-occurrence of pattern among features' is needed. Using quantitative techniques, he identifies the groups of features that co-occur in texts and then interprets them according to their function. Once patterns or sets of features have been established, it is of interest to know 'what functional or situational parameters relate to the co-occurring sets of features, influencing their systematic use across a range of texts.' (1988: 19).

Texts are assigned to types according to their linguistic characteristics and co-occurrence of features, and the function of the individual text type is then determined on linguistic grounds. This 'multi-feature/multi-dimensional' approach is based on automated counts of linguistic features and factor analysis. Factor analysis uses the frequency counts of linguistic features to identify sets that co-occur in texts and that are highly correlated. It is assumed that the co-occurring features share a communicative function. For example, Biber's first of six dimensions is referred to as 'Involved versus Informational Production'. The

linguistic features which co-occur on this dimension are positive in this respect or negative. That is, those 'marking a high level of interaction and personal affect' (eg first and second person pronouns, private verbs, WH questions) and 'a generalized and fragmented presentation of content' (eg hedges, the pronoun *it*) are positive here, whereas the 'highly informational' with 'almost no concern for interpersonal involvement or affective content' (eg few verbs, high frequency of nouns and prepositions) are negative.

So, according to this previous research (Biber 1988; Chafe 1982; Halliday 1989) both spoken and written texts display evidence of orality such as lack of planning and elaboration, fragmentation, involvement and interaction, which is indicated by the relative frequency of specific linguistic features. Results showed that the traditional distinction between spoken and written texts could not be upheld as it is not a valid dimension of textual variation. Text type is found to be of more significance than the medium.

2.3 Analysis of orality features

Whether adjectival premodification of nouns is a feature which correlates with orality can only be determined when other features associated with orality occur at the same time with a higher frequency than in other text types. On the basis of such relevant linguistic features as those found in previous research described above, we have examined the four Kenyan text types to determine their position on a cline of orality.

Table 1: Indicators of orality in ICE-EA text types

Indicator	short stories	social letters	feature articles	editorials
Total words	20,086	22,079	20,098	20,041
Sentence length (words)	12.49	18.84	21.42	25.2
Type/token (standardized)	40.99	41.27	43.12	45.07
Clausal subordinators (per 1,000 words)				
if	3.14	4.80	1.14	2.59
because	10.5	1.54	1.04	1.05
much/many	1.29	3.17	1.84	2.39
some	1.39	4.12	2.74	2.24
all	3.78	4.80	4.13	4.39

but	5.53	7.47	4.83	4.04
just/really	1.54	5.84	1.24	1.50
personal pronouns (per 1,000 words)				
I	27.28	44.16	1.49	0
me	8.91	11.78	0.25	0
you	8.96	29.98	0.85	0.05
we	2.04	6.84	2.44	4.24
us	1.00	2.76	0.99	1.45
+ contractions				
I	1.84	4.85	0.05	0
you	0.30	1.6	0	0
we	0.05	0.41	0	0
Total 1st/2nd person	**50.38**	**102.38**	**6.07**	**5.74**
he/him/his	26.98	6.97	9.45	7.04
she/her/hers	31.46	4.62	6.12	0.25
they/them/their	8.16	4.76	11.05	11.93
he/she/they	0.30	0.81	0	0.05
Total 3rd person	**66.90**	**17.16**	**26.62**	**19.27**

Table 1 shows that there is a recognisable cline as far as sentence length is concerned. The average sentence in the creative writing is 12.49, in the social letters 18.84 words, in the features 21.42, and in the editorials 25.2 words. A greater variety of lexemes in the features, and especially the editorials can be seen from the type/token figures, the progression being short stories (40.99), social letters (41.27), features (43.12) and editorials (45.07). Previous research, as discussed above, pointed to the shorter length of sentences in spoken texts and the greater variety of lexemes in written data.

Hudson-Ettle (1998: 174) found that the optional adverbial clauses in the oral texts analysed were mostly finite (with a conjunction) and this text type had the highest percentage of finite and lowest percentage of non-finite adverbial clauses. She found also that the choice of causal subordination is characteristic of spoken texts in the Kenyan data – 46 per cent of all causal subordination occurred in the oral data – mostly finite and introduced by *because*. In the present study, there are also more occurrences per thousand words of this con-

junction in the Kenyan social letters (1.54) than in the two newspaper text types (1.04, 1.05) but considerably more in the short stories (10.5).

In the 1998 study, more conditional clauses were found in the oral data than in all the other text types (1988: 149), and these were introduced almost exclusively by *if.* In the present study, both the short stories (3.14) and the social letters (4.18) contain more *if*-clauses than the other two categories (1.14 and 2.59).

Other features considered as more frequent in oral data in the research mentioned above were also taken into account and the results for the social letters are higher for *much/many, some, all*[5], *but, just, really* than in all the other texts. The references to the speaker's mental process and colloquial expressions seen by Chafe (1982: 47) as indicators of involvement, *I think/I know/I mean/you know,* are not listed in the table but occurred 54 times (2.45 per thousand words) in the social letters and 19 times (0.95) in the short stories but not once in the other text types. *Sort of* was in general not frequent but there were ten instances (0.45) in the letters, zero in the short stories and features and it occurred only twice in the editorials. *Seems* and *appears* were infrequent and not a distinguishing feature.

First and second person pronouns also function as indicators of orality (cf Chafe 1982: 47; Biber 1988: 225) and are infrequent in texts where there is less personal involvement and a low degree of interaction with an addressee. This is confirmed here by a simple computer count of these lexemes in each text type. The first and second person pronouns *I, me*, *we, us* and *you* are most frequent, with or without contracted verb forms, in the social letters (102.38 per 1,000 words). Only half as many instances (50.38), yet far more than in the features (6.07) and editorials (5.74), are to be found in the creative writing. A particular form of personal involvement can nevertheless be detected in the editorials in the use of *we* and *us*, where the writer usually discusses issues of importance to the society of which he (rarely she) is part, or speaks for the newspaper he is representing, but never as a private person.

(1)   He said the challenge **we** have is whether **we Kenyans** will be able to put **our** nation above tribe; W2E003K

(2)   Not long ago, **we** asked in these columns that the Government tell Kenyans what elaborate and necessary mechanisms it had put in place to ensure that food would be distributed in such a manner that it would reach all Kenyans. W2E001K

The occasional use of the first and second person pronouns in the features can be explained by the composition of the corpus. The newspaper articles include spo-

ken language (in inverted commas) of people interviewed by the reporters. In Hudson-Ettle (1998: 182) a word count illustrated that, on the one hand, personal involvement was strongest in the spoken data but that it was also evident to a varying degree in the three other written text types analysed.[6] The distribution of third person pronouns is not related to orality but is obviously a distinct feature of narrative writing, the short stories having 66.9 instances, which is more than twice as frequent as in the features (26.62), which also have a narrative purpose, and more than three times as many as in the social letters and the editorials.

3 Adjectival premodification of nouns

3.1 Description of the analysis

The present study focuses on the use of one type of noun phrase modification only, namely premodification by adjectival elements. Apart from central and peripheral adjectives found in attributive position (Quirk et al 1985: 402ff), present and past participles functioning as noun phrase premodifiers are also included in the analysis. Biber et al (1999: 574f) make a distinction between adjectives, participial modifiers and nouns. They note that '[p]articipial modifiers are typically adjectival, but in some cases they have the character of noun rather than adjective modifiers, and in yet other cases their word-class membership is unclear'. The stance taken in the present paper is to treat participial premodifiers on a par with adjectives, and to refer to this group of noun phrase premodifiers as adjectival. The three structural types are exemplified in (3)–(5), taken from the general feature articles in UPC:

(3)   … by avoiding the *high* inflation which undermined the economy in the late Eighties. (UPCA02NP83)

(4)   … before he had ever been made *investigating* officer. (UPCA01NP148)

(5)   … received *widespread* full page coverage. (UPCA01NP276)

Furthermore, the study deals only with noun phrases with common noun heads. Proper noun and pronominal heads have been excluded as the potential for such noun phrases to have premodification is very limited (Biber et al 1999: 579f; Nilsson 2001).

It was pointed out above that structural elaboration of the noun phrase is an indication of specialized language. Miller and Weinert (1998: 139) state that 'the general picture is that most complex [noun phrases] are found in dense technical texts, […] with the least complex structures in tabloid newspapers, comics, and

books for young children'. As regards the text types discussed by Miller and Weinert (1998), we note that texts that (intuitively) can be placed at the lower end of the orality continuum, such as academic monographs, legal documents, and broadsheet newspapers, are found to have comparatively complex noun phrases. In contrast, text types that can be placed towards the higher end of the orality scale, eg comics and books for young children, typically have noun phrases of rather low complexity. This observation supports the hypotheses put forth in the present study, namely that text types that show evidence of orality features (see above) should also have a comparatively large proportion of less complex noun phrases.

In the present study, the usual terminological distinction between 'simple' and 'complex' noun phrases is not applicable. The most common practice (see eg Quirk et al 1985; Biber et al 1999) is to classify noun phrases as simple when they do not have pre- or postmodification, whereas a complex noun phrase has either pre- and/or postmodification. The presence of determiners is normally not taken into consideration in this division. As we are concerned with adjectival premodification only in the present study, we will make a distinction between 'noun phrases without adjectival premodification' and 'noun phrases with adjectival premodification'. The number of adjectival elements will also be taken into consideration in the quantitative analyses. In contrast to this, it can be pointed out that Miller and Weinert (1998: 135) use the term 'simple noun phrase' in a slightly different fashion in that they say that 'a relatively simple noun phrase consists of a noun modified by one or two adjectives, or a numeral/quantifier, or a prepositional phrase, or some combination of these modifiers.' Miller and Weinert (1998) regard complex noun phrases as involving (post)modification by means of clausal elements, 'possibly in combination with adjectives, quantifiers, and prepositional phrases' (1998: 135). In this way, a noun phrase as illustrated in (6) is seen as 'relatively simple' by Miller and Weinert, although it contains an adjectival premodifier and a postmodifying prepositional phrase.

(6)   ... that *the early dose of deflation* will deepen … (UPCA02NP26)

For the purposes of the present study, the noun phrases have been coded and analysed in the following way. All top-level noun phrases and noun phrases that are themselves modifiers of other noun phrases have been coded for the presence or absence of adjectival premodification, along with the number of adjectival premodifiers. The quantitative analysis takes into consideration raw frequencies, the relative distribution over the text types investigated, and normalized scores (to occurrences per 1,000 words). Tables 2 and 3 show first the absolute numbers, then the normalised figures.

3.2 Results: The Kenyan data

The results of the noun phrase analysis of the Kenyan data do show parallels with the orality discussion above. The two more 'oral' genres, namely the social letters and short stories, have fewer noun phrases, whether with or without adjectival modification, than the other two genres, the more informal social letters having the lowest figures (cf Table 2). With adjectival premodification these are 20.74 and 26.09 at the 'oral' end, and 44.63 and 48.95 for feature articles and editorials respectively.

Table 2: Distribution of noun phrase premodification by text type: Kenyan data

	Short stories		Social letters		Feature articles		Editorials	
	N	Norm	N	Norm	N	Norm	N	Norm
Total number of words	20,086		22,079		20,098		20,041	
Total number of noun phrases	3,038	*151.25*	2,587	*117.2*	3,893	*193.7*	3,772	*188.2*
Noun phrases with NO adjectival premodification	2,514	*105.99*	2,129	*96.43*	2.996	*149.1*	2,794	*139.4*
Noun phrases WITH adjectival premodification	524	*26.09*	458	*20.74*	897	*44.6*	978	*48.8*

Chi-square: 103.94 with 3 df; $p<0.001$

The distribution as to the number of adjectival premodifiers in each noun phrase is shown in Table 3:

Table 3: Distribution of the number of adjectival premodifiers by text type: Kenyan data

	Short stories		Social letters		Feature articles		Editorials	
	N	Norm	N	Norm	N	Norm	N	Norm
ONE adjective	483	*0.05*	422	*19.11*	813	*40.45*	861	*42.96*
TWO adjectives	37	*1.84*	31	*1.40*	78	*3.88*	107	*5.34*
MORE THAN TWO adjectives	4	*0.20*	5	*0.23*	6	*0.30*	10	*0.50*

The grouping of the relative distribution is evident in Table 4. Short stories and social letters have similar percentages (82.75 % and 82.73 % are without adjectival premodification); the figures for feature articles are similar to those for editorials (76.96 % and 73.99 %).

Table 4: Relative distribution of noun phrases with and without adjectival pre-
modification: Kenyan data

	Short stories		Social letters		Feature articles		Editorials	
	N	%	N	%	N	%	N	%
Noun phrases WITH adjectival premodification	524	*17.25*	458	*17.7*	897	*23.04*	978	*26.01*
Noun phrases WITHOUT adjectival premodification	2,514	*82.75*	2,129	*82.3*	2,996	*76.96*	2,794	*73.99*
Total	3,038	*100.00*	2,587	*100.00*	3,983	*100.00*	3,772	*100.00*

We will now look at the distribution of adjectival premodification on the sepa-
rate levels, top and embedded, the figures referred to being normalized counts.
On the top level, there is more premodification in the feature articles (35.43) and
editorials (42.66) than in the social letters (19.25) and short stories (24.49), as
displayed in Table 5:

Table 5: Distribution of adjectival premodification in top-level noun phrases:
Kenyan data

	Short stories		Social letters		Feature articles		Editorials	
	N	Norm	N	Norm	N	Norm	N	Norm
Noun phrases WITH adjectival premodification	492	***24.49***	425	***19.25***	712	***35.43***	855	***42.66***
Noun phrases WITHOUT adjectival premodification	2,294	*114.21*	1,950	*88.32*	2,467	*122.75*	2,348	*117.16*
Total	2,786	*138.70*	2,375	*107.57*	3,179	*158.17*	3.203	*159.82*

Chi-square: 95.28 with 3 df; p<0.001

Embedded noun phrases are preferred in the former text types and are more fre-
quently premodified. The 'oral' written texts have only 9.6 and 12.55 instances
of embedding per 1,000 words, compared with 35.53 and 28.39, as seen in Table
6:

Table 6: Distribution of adjectival premodification in embedded noun phrases: Kenyan data

	Short stories		Social letters		Feature articles		Editorials	
	N	Norm	N	Norm	N	Norm	N	Norm
Noun phrases WITH premodification	32	*1.59*	33	*1.49*	185	*9.20*	126	*6.29*
Noun phrases WITHOUT premodification	220	*10.95*	179	*8.11*	529	*26.32*	443	*22.10*
Total	252	***12.55***	212	***9.6***	714	***35.53***	569	***28.39***

Chi-square: 24.34 with 3 df; p<0.001

This preference is made clearer in Table 7, where we see from the percentage figures that embedded premodified noun phrases are more typical of the less 'oral' data (4.75 % and 3.34 %, compared to 1.05 % and 1.28 %).

Table 7: Relative distribution of top-level and embedded noun phrases with and without adjectival premodification: Kenyan data

	Short stories		Social letters		Feature articles		Editorials	
	N	%	N	%	N	%	N	%
Top level noun phrases with adjectival premodification	492	*16.19*	425	*16.43*	712	*18.29*	855	*22.67*
Top level noun phrases without adjectival premodification	2,294	*75.51*	1,950	*75.38*	2.467	*63.37*	2,348	*62.25*
Embedded noun phrases with adjectival premodification	32	*1.05*	33	*1.28*	185	*4.75*	126	*3.34*
Embedded noun phrases without adjectival premodification	220	*7.24*	179	*6.92*	529	*13.59*	443	*11.74*
Total	3,038	*100.00*	2,587	*100.00*	3,893	*100.00*	3,772	*100.00*

3.3 Feature articles: A comparison of Kenyan and British data

In the following, we take a closer look at the three corpus components containing feature articles. The discussion focuses initially on the distribution of the orality features in the three corpus components. The same orality features are used as above. After this, the distribution of noun phrases with and without adjectival premodification is investigated.

3.3.1 Orality features in three selections of feature articles

A comparison of the results obtained for the three collections of feature articles regarding orality features shows very clearly that the general Kenyan and UPC feature articles have similar scores, whereas the figures for the British travel articles are, in several cases, radically different. The results are summarized in Table 8.

It is above all in the use of indefinite and personal pronouns that the differences between the general feature articles and the travel articles are most clearly seen. As regards the indefinite pronouns investigated, the travel articles have scores that are up to seven times as high as those attested for the general feature articles. Furthermore, the adversative conjunction *but* is roughly six times as common in travel articles compared to the general feature articles. With the added support lent by the figures for the intensifiers *just* and *really*, where the figures for the travel articles are nearly three times as high as those attested for the general feature articles, it is possible to draw the conclusion that travel articles have several very clear indicators of a more 'oral' style than is the case for the general feature articles studied.

Table 8: Orality features in selections of feature articles

Indicator	Kenyan features (ICE-EA)	British general features (UPC)	British travel features (UCTT)
Total words	20,098	20,366	20,511
Sentence length (words)	21.42	22.43	23.14
Type/token (standardized)	43.12	47.57	49.09
Clausal supordinators (per 1,000 words)			
if	1.14	1.41	1.83
because	1.04	0.6	0.5
much/many	1.84	1.5	8.8
all	4.13	1.9	13.0
some	2.74	0.8	6.7
but	4.83	4.5	23.2
just/really	1.24	1.6	8.9

Personal pronouns (per 1,000 words)			
I	1.49	3.4	14.1
me	0.25	0.7	2.6
we	2.44	1.5	18.1
us	0.99	0.2	4.5
you	0.85	1.0	33.9
he/him/his	9.45	15.8	10.8
she/her/hers	6.12	4.1	3.7
they/them/their	11.05	7.2	25.1
+ contractions			
I	0.05	0.6	1.6
you	0	0.05	2.4
we	0	0	0.9
he/she/they	0	0.3	0.8
Total 1st/2nd person	**6.07**	**7.45**	**78.1**
Total 3rd person	**26.62**	**27.4**	**40.4**

The frequency of personal pronouns is also distinctly different in travel articles compared to the other general feature articles. All the pronouns investigated have higher scores for travel articles compared to the general Kenyan and British feature articles. The figures for the second person pronoun *you* are especially striking. The pronoun *you* occurs nearly 36 times more often in the travel articles compared to the general feature articles. The use of this pronoun is an obvious indication of reader involvement, and can be accounted for as a linguistic reflection of a personal, rather informal tone. This explanation is further supported by the very high figures for the first person pronouns in the travel articles. Together, the prevalence of first and second person pronouns is indicative of a very personal tone where the writer makes his/her presence very much felt, at the same time as the reader is directly addressed.

3.3.2 Adjectival premodification in the feature articles

We now turn to a discussion of the other main issue of this paper, namely to what extent the scores of the analysis of orality features tally with the extent to which adjectival premodification is used. The question we ask is whether the

scores of the orality analysis pattern with those of the analysis of adjectival pre-modification. We expect to find a relatively large proportion of noun phrases with adjectival premodification in texts with generally low orality scores.

The analyses in this section are based on a total of 11,076 noun phrases. The feature articles in the Kenyan part of the ICE-EA are contrasted with feature articles representing British English taken from UPC and UCTT. Our hypothesis is that Kenyan English displays more indicators of orality, and that this will be reflected in a smaller proportion of noun phrases with adjectival premodification in these feature articles compared to the British general and tourism/travel feature articles.

The distribution of noun phrases with and without adjectival premodification is given in Table 9, which shows raw frequencies and the relative distribution over text types. The majority of noun phrases, 74 per cent, do not have adjectival premodification. Adjectival premodification is found in 26 per cent of cases. In relation to these figures, the travel feature articles represented in UCTT have a distribution of noun phrases with adjectival premodification that is different from that found in the other two text types. In UCTT, adjectival premodification is found in 33 per cent of the noun phrases, compared to 23 per cent in the Kenyan feature articles, and 22 per cent in the British feature articles.

Table 9: Distribution of noun phrase premodification by text type: feature articles

	ICE-EA		UPC		UCTT		Total	
	N	%	N	%	N	%	N	%
Noun phrasess without adjectival premodification	2,996	*77*	2,779	*78*	2,438	*67*	8,213	*74*
Noun phrases with adjectival pre-modification	897	*23*	789	*22*	1,177	*33*	2,863	*26*
Total	3,893	*100*	3,568	*100*	3,615	*100*	11,076	*100*

Chi-square: 126.90 with 2 df; $p<0.001$

There are, in other words, only minor variations in the relative distribution of noun phrases with adjectival premodification between the two corpus components that represent general feature articles, whereas the travel feature articles in UCTT have a notably larger proportion of noun phrases with adjectival premodification. These differences are statistically significant. The results can be interpreted tentatively in the light of the focus on detailed descriptions of sights of interest in UCTT travel articles. The content of these articles is specialized in

the sense that they deal with a limited subject matter. Furthermore, the travel articles are personal accounts of an experience where the reporter's opinions of what is described are reflected in a higher proportion of adjectival premodifiers. In contrast to this, the general Kenyan and British feature articles deal with various topics where the focus is more on providing background information and comments. Examples (7)–(9) will illustrate the differences:

(7) The *jerry-built* original, along with much of the tower, fell vertically into the nave at lunchtime on February 21, 1861, and was replaced by an uncharacteristically *modest* (and hence uncharacteristically *successful*) *new* one by the *Victorian* meddler and *gothic* revivalist Sir George Gilbert Scott. (UCTT01NP98–104)

(8) Customers wanting to use cash machines at several banks in London have faced signs claiming that the card swipe mechanism that opens the doors is out of order. (UPCA02NP371–374)

(9) For, although a scholar in those ancient times was expected to be thoroughly well-versed in the arts, sciences and *fine* arts – and although this is impossible today – the idea still should be to produce an individual who has some *passable* acquaintance with all the disciplines and an *intimate* relationship with one or two. (ICE-EA W2C022K)

Another way of viewing the distribution of noun phrases with adjectival premodification is to express its occurrence as a normalized score (to occurrences per 1,000 words). Figure 1 shows the normalized scores for top-level and embedded noun phrases, noun phrases with no adjectival premodification, and noun phrases with adjectival premodification. The Kenyan features have a higher overall incidence of noun phrases, 194 noun phrases per 1,000 words, compared to the two British components, where the normalized scores are 173 for UPC feature articles and 176 for UCTT. As these figures include noun phrases with common noun heads only, one possible explanation for the differences can be that the British feature articles utilize pronominal reference to a higher degree.[7] The distribution of certain pronominal items accounted for in Table 8 clearly show that the normalized scores for first and second person pronouns are decidedly higher for the travel articles compared to the general feature articles.
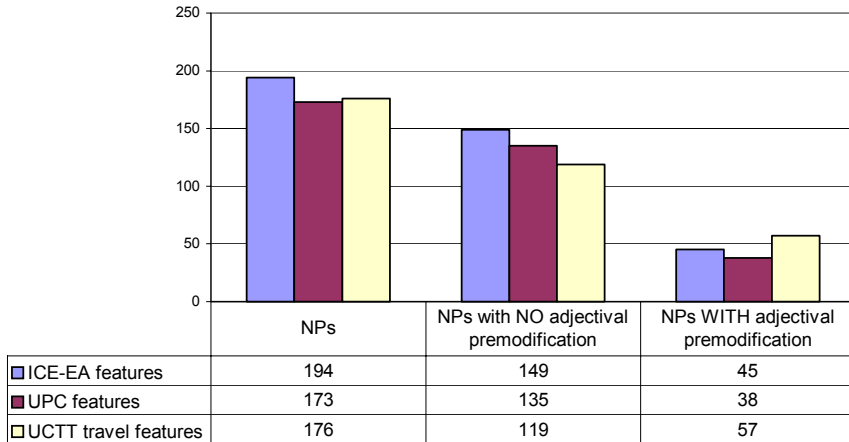
	NPs	NPs with NO adjectival premodification	NPs WITH adjectival premodification
ICE-EA features	194	149	45
UPC features	173	135	38
UCTT travel features	176	119	57

Figure 1: Distribution of NPs and adjectival premodification (scores normalized to occurrences per 1,000 words)

Figure 1 also indicates that the travel articles represented in UCTT not only have the highest number of noun phrases with adjectival premodification, 57 noun phrases, but also the lowest score for noun phrases without adjectival premodification, 119 noun phrases. In comparison, the figures for the British and Kenyan general feature articles suggest that the proportion of noun phrases with adjectival premodification is smaller in these texts. This is particularly so for the Kenyan feature articles, which have the largest number of noun phrases per 1,000 words, 194, but not the largest number of noun phrases with adjectival premodification.

The figures for the general British feature articles, finally, pattern in a way that is similar to that of ICE-EA. The score for noun phrases without adjectival premodification is 135, which means that the British UPC feature articles are in-between UCTT and ICE-EA. On the other hand, the score for noun phrases with adjectival premodification is the lowest, 38.

These results suggest that adjectival premodification is not a feature that separates Kenyan feature articles from British ones. Rather, it is the subject matter, seen in the scores for UCTT, which has an influence on the use of noun phrases with adjectival premodification. When seen in conjunction with the scores obtained for the orality features discussed above, it seems that the British travel articles in UCTT represent a text type with more orality features than the

general feature articles in ICE-EA and UPC. Another important conclusion is that the level of generality or spread of topics is a crucial factor in accounting for the patterns found. Consequently, it is not possible to compare the results obtained from the three corpora included in this study without taking into consideration the subject matter of the texts.

A closer look at the number of adjectival premodifiers attested in the three text types reveals some notable differences. The relative distribution of noun phrases with one, two, and more than two adjectival premodifiers is shown in Table 10. On average, 85 per cent of the noun phrases with adjectival premodification have one premodifier, 14 per cent have two premodifiers and three or more premodifiers are found in one per cent of the noun phrases only. Again, the figures attested for UCTT stand in sharp contrast to those for the other two corpus components. In UCTT, the proportion of noun phrases with two adjectival premodifiers is considerably larger, 24 per cent, compared to the other two text types, where the proportions of noun phrases with two premodifiers are nine and seven per cent, respectively. The proportion of noun phrases with three or more adjectival premodifiers is small in all three text types.

Table 10: Distribution of adjectival premodifiers by text type: feature articles

Number of adjectival premodifiers	ICE-EA		UPC		UCTT		Total	
	N	%	N	%	N	%	N	%
One premodifier	813	*90*	732	*93*	883	*75*	2,418	*85*
Two premodifiers	78	*9*	57	*7*	280	*24*	415	*14*
More than two premodifiers	6	*1*	0	*0*	14	*1*	20	*1*
Total	897	*100*	789	*100*	1,177	*100*	2,853	*100*

Chi-square: 150.57 with 4 df; p<0.001

On a general level, the results in Table 10 agree with those reported in several other studies on noun phrase modification, namely that more than two (adjectival) premodifiers are rare (see eg Varantola 1984; Jucker 1992; Nilsson, 2001). Miller and Weinert (1998: 139) note that, in spontaneous spoken English, 'noun phrases with more than one adjective are rare'. Biber et al (1999: 578) found that the proportion of noun phrases with premodification is lower in conversation compared to fiction, news, and academic writing. The share of premodified noun phrases was highest in news discourse. However, as the figures in Table 10 clearly demonstrate, there are considerable differences within this broad text type. Jucker (1992), in his comprehensive investigation of noun phrase structure in British newspapers, demonstrates that several aspects relating to the complex-

ity of the noun phrase vary both across types newspaper and across sections. However, as regards the occurrence of noun phrase premodification, Jucker finds no statistically significant variation across either newspaper type or section (Jucker 1992: 109f). The results regarding the relative frequency of adjectival premodification obtained in the present study, namely a significantly large proportion of noun phrases with more than one adjectival premodifier attested for the travel articles in UCTT, strongly indicate that the internal composition of text types is very important. This further supports the conclusion advanced in the previous section, namely that the subject matter in rather specific terms has a notable influence on noun phrase premodification.

We will now examine in more detail the distribution of adjectival premodification in the two layers of noun phrase structure outlined above. We shall look at the distribution of adjectival premodification in top-level noun phrases and in noun phrases that are themselves modifiers of other noun phrases, referred to as embedded noun phrases. This will be done by first investigating the relative distribution of adjectival premodification in top-level noun phrases, followed by the same analysis regarding embedded noun phrases. Finally, the results of the two levels will be compared.

The distribution of top-level noun phrases with and without adjectival premodification is given in Table 11. On average, 76 per cent of the top-level noun phrases do not have adjectival premodification, against 24 per cent with adjectival premodification. The same pattern as found in the analysis of adjectival premodification as such (see Table 9) is present here, too. The top-level noun phrases in UCTT have adjectival premodification to a much higher extent, 32 per cent, compared to the other two text types, where premodification is found in 22 and 19 per cent of cases, respectively.

Table 11: Distribution of adjectival premodification in top-level noun phrases: feature articles

	ICE-EA		UPC		UCTT			Total
	N	%	N	%	N	%	N	%
Top-level noun phrases without adjectival premodification	2,467	78	2,005	81	1,177	68	5,649	76
Top-level noun phrases with adjectival premodification	712	22	484	19	546	32	1,742	24
Total	3,179	100	2,489	100	1,723	100	7,391	100

Chi-square: 88.98 with 2 df; p<0.001

The distribution of adjectival premodification in embedded noun phrases is shown in Table 12. Overall, the proportion of embedded noun phrases with adjectival premodification is 30 per cent, which suggests that embedded noun phrases more frequently take adjectival premodification compared to top-level noun phrases (cf Table 11). Furthermore, the embedded noun phrases in UCTT have a relatively large share of adjectival premodification, 33 per cent, compared to the other two text types. In ICE-EA, 26 per cent of embedded noun phrases have adjectival premodification, and for UPC, the share of noun phrases with adjectival premodification is 28 per cent. The scores are statistically significant.

Table 12: Distribution of adjectival premodification in embedded noun phrases: feature articles

	ICE-EA		UPC		UCTT		Total	
	N	%	N	%	N	%	N	%
Embedded noun phrases without adjectival premodification	529	*74*	773	*72*	1,261	*67*	2,563	*70*
Embedded noun phrases with adjectival premodification	185	*26*	304	*28*	630	*33*	1,119	*30*
Total	714	*100*	1,077	*100*	1,891	*100*	3,682	*100*

Chi-square: 16.81 with 2 df; p<0.001

A comparison of the normalized scores for adjectival premodification in all noun phrases, top-level noun phrases, and embedded noun phrases is presented in Figure 2. The scores show clearly that adjectival premodification is most common in the travel articles, irrespective of noun phrase level. However, adjectival premodification is more frequent in top-level noun phrases than in embedded noun phrases in the general feature articles (both British and Kenyan), whereas in the travel articles, more adjectival premodifiers are found in embedded noun phrases compared to top-level noun phrases.
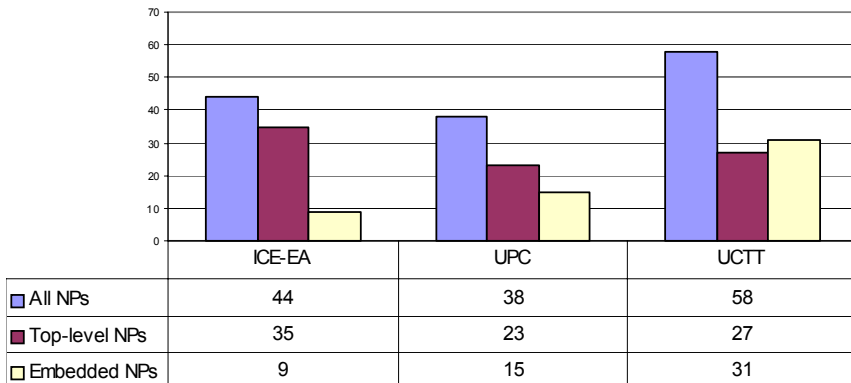
	ICE-EA	UPC	UCTT
☐ All NPs	44	38	58
☐ Top-level NPs	35	23	27
☐ Embedded NPs	9	15	31

Figure 2: Adjectival premodification in different layers of NP modification (scores normalized to 1,000 words)

One explanation for this surprising result can be found in the fact that a comparatively large share of the noun phrase heads in the travel articles are pronominals that are postmodified by prepositional phrases where the noun phrase heads are common nouns. An example of this is shown in (10):

(10)   … and some of the town's *best Georgian* houses… (UCTTB01NP173)

Another explanation is found in the narrative purpose of the travel articles. In contrast to nominal premodifiers, which are basically classifying, adjectival premodifiers most frequently have a descriptive function (Nilsson 2001). Consequently, adjectival premodifiers should be expected to be relatively frequent in texts where the focus is on narrating a personal experience as is the case with the travel articles included in UCTT. The fact that these texts are characterized by a high degree of personal involvement (high scores for first and second person pronouns, for instance) further supports this conclusion, in that personal accounts are more likely to contain a fair amount of (subjective) description.

4 Conclusion

This study has pursued issues related to the orality and formality/informality debate by addressing the possible connection between indicators of orality and the occurrence of adjectival premodification of noun phrases. The material used

consisted of written English texts produced in Kenya and in Great Britain, thus representing English as a second language (lingua franca) and English as a first language. Six components from three existing corpora were utilized: Short stories, Social letters, Feature articles, and Editorials from the Kenyan part of ICE-EA; Feature articles from UPC; and feature travel articles from UCTT.

The study was carried out in two steps. The first step consisted of a comparison of the four components of the ICE-EA. The aim of this step was to investigate the connection between degrees of orality and the occurrence of adjectival noun phrase premodification. A number of linguistic features indicating orality (cf Biber 1988) were used to see whether the four Kenyan components patterned in a way that could be interpreted in terms of an orality cline. It was found that the short stories had shorter sentences and less lexical variety than the features and the editorials, as well as more finite causal (*because*) and conditional (*if*) subordination – findings which had previously been found to be characteristic of spoken language. Features indicating vagueness or indefiniteness (*much*, *many*, *some*, *sort of*, *just*, *really*) were not more frequent in the creative writing but certainly in the social letters. This leads to the conclusion that they are a feature on the formality cline, the social letters being far more informal than the short stories. This is also indicated by such colloquial expressions as *I think*, *I know*, *I mean*, *you know,* which do occur in the short stories but are far more common in the letters. Personal involvement and interaction are major characteristics of orality and the results show social letters and the short stories to differ immensely in their use of first and second person pronouns.

Having established that two text types (short stories and social letters) could be positioned towards one end of the orality cline and the other two (feature articles and editorials) towards the other end, we compared the distribution of noun phrases in each text type. The total number of noun phrases, but especially premodification and embeddedness, were distinguishing features between these two groups.

The second step involved a comparison of three collections of feature articles. For this part, the Kenyan feature articles, the British feature articles from UPC, and the British travel articles were used. The main hypothesis to be tested was that, since Kenyan English is an ELS variety used as a lingua franca, the Kenyan feature articles should contain more oral features and, consequently, fewer noun phrases with adjectival premodification, compared to feature articles written in British English. The two British components were tested for the same general orality features as the Kenyan material, and comparisons across the three text types were made. The results indicated clear differences in the scores for several of the general orality features. However, these differences were not in accordance with our expectations.

The scores for indefinite pronouns (*much*/*many* and *some*) including *all* were considerably higher in UCTT than in the other two texts. First and second personal pronouns, indicating personal involvement, also had decidedly higher scores in UCTT, as were the figures for pronominal reference in general (third person pronouns). Furthermore, as an indicator of informality, the scores for contracted forms (verbal contractions involving *I*, *you*, *we*, *he*, *she*, *they*), were all consistently higher in UCTT than in the general feature articles although the differences in this respect were not as considerable as for the first sets of indicators. The main conclusion drawn on the basis of these results was that, in several respects, travel articles have linguistic patterns that are clearly different from general feature articles. The differences between the components compared were interpreted as arising from the fact that the travel articles focus on a more limited set of topics compared to the general feature articles. This is an important observation regarding the comparability of corpora, as such a specialized focus leads to a concentration of certain linguistic features.

As regards the analysis of adjectival premodification, the results confirmed the findings of the analysis of the general orality features. Both as regards the occurrence of adjectival premodification and the number of adjectives used, it was found that the general feature articles from ICE-EA and UPC formed a group with similar proportions, and that the articles from UCTT had scores that were different. For all the aspects investigated, the results clearly indicated that adjectival premodification is more common in travel articles compared to the general feature articles. Consequently, our hypothesis that Kenyan feature articles, because of their status as an ESL variety, should have more indicators of orality has not been confirmed. Instead, it was suggested that the subject matter, and indirectly, the communicative purpose, of the feature articles exert an influence on linguistic structure, adjectival premodification included.

Whereas one would expect text types with the same 'labels', also referred to as 'genres' (Biber 1989: 6), to be comparable, even if the corpus as a whole contains different components, our study has shown that their composition is far from a straightforward matter. Editorials seem to be an established genre, consisting as they should of institutional (ie representing the respective newspaper) rather than personal comments on topics of political or social interest (cf Hudson-Ettle, forthcoming, for a contradiction). Stories and novels – which are grouped together as creative writing in ICE – may be found to differ in their linguistic preferences. Analyses of the Kenyan short stories have shown that features of orality are considerably more frequent in these than in the other published Kenyan text types. This may also be true for the novels, which will lead to the conclusion that creative writing in Kenya is influenced by the oral

tradition of story-telling, in that it includes considerable stretches of dialogue as direct speech. If the analysis of the novels presents a different picture, this will be evidence that the two cannot be grouped together – they are different text types of one genre. The results we have obtained for the three subcorpora of feature articles and the conclusions we have drawn point to the need to take the content of the genres into consideration.

Notes

1. We are grateful to Margareta Westergren Axelsson for kindly making the feature articles in UPC available to us for analysis.
2. This is the same sub-section that forms part of the basis for analyses of noun phrase structure in Nilsson (2001).
3. The number of conjunctions per thousand words in samples of written and spoken language was evidence of the more fragmented nature of speech (Chafe 1982: 38).
4. Chafe's analysis did not include so-called zero *that*-clauses. Beaman (1984) and Greenbaum, Nelson and Weitzmann (1996), however, found more *+that* clauses in their spoken data than in their written data. In Hudson-Ettle (forthcoming), *+that* clauses are more frequent in the written Kenyan data. Consideration of text type rather than medium alone is seen to be essential.
5. The corpus findings of *all* in Biber et al (1999: 278) show this quantifier to be most frequent in conversation and fiction (the more 'oral' text types), less frequent in news and academic prose (the four text types analysed in their study). But there are more instances of *much*/*many* and *some* in their subcorpus of academic prose, which they account for as 'expressing guarded generalization' (1999: 277).
6. There was in fact a distinct cline from 55.5 per cent of all pronouns analysed occurring in the spoken data, 26.5 per cent in the creative, 12.3 per cent in the columns and 5.7 per cent in the editorials.
7. This is the case for the travel articles represented in UCTT, where the proportion of noun phrases with pronominal heads is by far the highest for the three text types investigated in Nilsson (2001). The share of pronominal noun phrase heads is 28 per cent in the travel articles, compared to 15 per cent for tourist brochures and 11 per cent for travel guides.

References

Beaman, Karen. 1984. Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In D. Tannen (ed). *Coherence in Spoken and Written Discourse*. Norwood, NJ: Ablex, 45–80.

Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas. 1989. A typology of English texts. *Linguistics* 27: 3–43.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Chafe, Wallace L. 1982. Integration and involvement in speaking, writing and oral literature. In D. Tannen (ed). *Coherence in Spoken and Written Discourse*. Norwood, NJ: Ablex, 35–53.

DeVito, Joseph A. 1967. Levels of abstraction in spoken and written language. *Journal of Communication* 17: 354–61.

Greenbaum, Sidney, Gerald Nelson and Michael Weitzman. 1996. Complement clauses in English. In M. Short and J. Thomas (eds). *Using Corpora for Language Research: Essays in Honour of Geoffrey Leech*. London: Longman, 76–91.

Gruner, Charles R., Robert J. Kibler and James W. Gibson. 1967. A quantitative analysis of selected characteristics of written and oral vocabulary. *Journal of Communications* 17: (1), 152–158.

Halliday, M.A.K. 1979. Differences between spoken and written language: Some implications for literacy teaching. In G. Page, J. Elkins and B. O'Connor (eds). *Communication Through Reading: Proceedings of the Fourth Australian Reading Conference*. Adelaide, S.A.: Australian Reading Association, Vol. 2, 37–52.

Halliday, M.A.K. 1989. *Spoken and Written Language*. Second edition. Oxford: Oxford University Press.

Halliday, M.A.K. 1994. *An Introduction to Functional Grammar*. Second edition. London: Arnold.

Hudson-Ettle, Diana. 1990. Linguistic features of style in Kenyan English novels of the late 1980s. Unpublished M.A. thesis, University of Bayreuth.

Hudson-Ettle, Diana. 1998. Grammatical Subordination Strategies in Differing Text Types in the English Spoken and Written in Kenya. PhD dissertation published on the Web. www.tu-chemnitz.de/~dihe.

Hudson-Ettle, Diana. Forthcoming. Nominal *that*-clauses: A study of the relevance of text type, medium and syntactic function. *Journal of English Linguistics* (2002).

Jucker, Andreas H. 1992. *Social Stylistics*. Berlin: Mouton de Gruyter.

Kachru, Braj B. (ed). 1982. *The Other Tongue: English Across Cultures*. Urbana: University of Illinois Press.

Miller, Jim and Regina Weinert. 1998. *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Clarendon Press.

Nilsson, Tore. 2001. Noun Phrases in British Travel Texts: A Corpus-Based Study. Unpublished. PhD Thesis. Uppsala University.

Platt, John, Heidi Weber and Mian Lian Ho. 1984. *The New Englishes*. London: Routledge and Kegan Paul.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Schmied, Josef. 1991. *English in Africa: An Introduction*. London: Longman.

Schmied, Josef and Diana Hudson-Ettle. 1996. Analysing the style of East African newspapers in English. *World Englishes* 15:(1), 103–113.

Varantola, Krista. 1984. *On Noun Phrase Structures in Engineering English*. Turun Yliopiston Julkaisuja, Sarja B, 168. Turku: Turun Yliopisto.

Westergren Axelsson, Margareta. 1998. *Contraction in British Newspapers in the Late 20th Century*. Studia Anglistica Upsaliensia 102. Uppsala: Acta Universitatis Upsaliensis.

Zuengler, Jane E. 1982. Kenyan English. In B. Kachru (ed). 112–124.