The Past Perfect in East African English:

Hunting for changes in meaning

Lauri Niskanen
University of Tampere
School of Language, Translation and Literary Studies
English Philology
Pro Gradu Thesis
October 2014

Tässä pro gradu -tutkielmassa käsitellään pluskvamperfektin semantiikan muutosta Itä-Afrikan englannissa verrattuna Iso-Britannian englantiin. Tutkimuksen kohteena ovat had + Ved -muotoa noudattavat pluskvamperfektit, niiden esiintymistiheys Kenian, Tansanian ja Iso-Britannian englanneissa, sekä niiden käyttötavoissa ja tarkoituksissa. Pluskvamperfektin valinta tutkimuskohteeksi perustuu lupaaviin tuloksiin aiemmassa tutkimuksessa Intian englannin vastaavasta muutoksesta, mutta pohjaoletusta samansuuntaisesta muutoksesta Itä-Afrikan englanneissa ei ole.

Kaikkien kolmen englannin muutoksen tutkimiseen käytetään The International Corpus of English -korpuksia, joiden yhtenäinen rakenne mahdollistaa tulosten vertailun. Kyseen ollessa semanttisesta muutoksesta, jonka olemassaolo ei ollut ennen tutkimuksen aloittamista varmaa, on semanttisen analyysin tulkinnanvaraisuus problematisoitu. Tutkielmassa esitellään myös korpustutkimuksen nopeuttamiseen tutkimuksen aikana kehitettyjä digitaalisia keinoja, sekä pohditaan käyttämättä jääneiden tekniikoiden ongelmia.

Tutkimuksessa käy ilmi, että määrällisesti pluskvamperfektin käytössä on eroja kohdemaiden välillä, ja pluskvamperfektin semantiikassa on myös havaittavissa pieni muutos tietyissä rekistereissä. Tansanian englannissa pluskvamperfektiä käytetään puheessa huomattavasti vähemmän kuin brittienglannissa tai Kenian englannissa. Toisaalta molemmissa Itä-Afrikan maiden englanneissa lyhennetyt muodot olivat harvinaisia verrattuna brittienglantiin. Itä-Afrikan englannissa, erityisesti kirjoitetuissa ja oppineissa rekistereissä pluskvamperfekti on saanut uudenlaisen käyttötarkoituksen normaalin rinnalle. Tässä uudessa merkityksessä pluskvamperfektiä käytetään ilman muodolle ominaista kahden tapahtuman väliin jäävää aikaviitettä. Sen sijaan pluskvamperfektiä voidaan käyttää erottamaan mennyt tapahtuma kertomuksen sen hetken aikaviitteestä.

Table of Contents

1 Introduction

The East African subcorpus of the International Corpus of English (ICE-EA) is a relatively new corpus, and though some studies have been made, mostly on style (Skandera 2000 & 2003) and idiom (Schmied, Hudson-Ettle 1995 & 1996) the tense-aspect-modality (TMA) system of East African English has not yet been studied in much detail. This thesis focuses on a very specific part of the TMA system, the semantics of the past perfect construction in East African English. The use of the tense is first studied by comparing the frequencies of contracted and non-contracted past perfects between East-African English and British English. The possible change in the meaning of the past perfect is a less tangible thing, but efforts are made to find and describe it with a theoretical framework. As the variation studied is in meaning, and not in form, searching the corpus is straightforward, but analyzing the data is challenging.

Devyani Sharma's study (2001) of the changes in meaning of the past perfect in Indian English was the inspiration for the current study. It was a comparative corpus study on the past perfect in the Kolhapur corpus of Indian English, compared to the Brown (American English) and LOB (British English) corpora, specifically the press and bureaucratic registers. Sharma's aim was to find out whether the inherent variation of the meaning of the past perfect is different in Indian English as compared to native varieties, where there is some variability towards using the past perfect with simple past meaning. Indeed, differences were found: There were 5% other meanings for past perfects in British English, compared to 22% of other meanings among past perfect forms in Indian English (ibid. 356). These other meanings include present perfect and preterite meanings. It was with a hope to find something similar in East-African English that this study was made.

Examples [1-3] illustrate the point of interest of the current study using Reichenbach's (1947) theory of time points on a timeline, which will be further explained in chapter 3. On the diagrams, sometimes drawn with a line, sometimes without, the past is situated to the left and the future to the

right of the diagram. The time points for (E)vent point, (R)eference point and (S)peech point can be situated on this line to explain temporal relations between events. In [1] and [2] we see the prototypical case of the simple past and past perfect respectively. In the past perfect example [2], there is a distinct point of reference with reference to which the action is situated in time, while in the simple past example [1], this point either does not exist or coincides with the event. In [3], the form of the past perfect carries the meaning of the simple past.

[1] John saw   E,R   S   = simple past with simple past meaning.
(Prototypical usage)
[2] John had seen   E   R   S   = past perfect with past-before-past meaning.
(Prototypical usage)
[3] ?John had seen   E,R   S   = past perfect with simple past meaning.
(Variation.)

The current study aims to find out whether this variance exists in East-African English. My research questions are the following:

1. Is there a change in the way the past perfect is used in East African English as compared to British English?
2. Similarly, is there a change in the meaning of the past perfect?
3. If so, what is the cause for the change?

Further, methodologically the study is divided into a preliminary quantitative part (chapter 5) and a qualitative analysis (chapter 6). In the quantitative part, the aim is to find out how much more or less the past perfect is used in East African English than in British English, and how much contractions are used in either variety. In the qualitative part, the focus is on the research questions: The aim is to find out how much there is variance in the meaning of the past perfect in East African English.

The language history of East Africa is presented in chapter 2. This includes a brief overview of the events that have affected the use of languages in the area – it will be seen that there has been a tendency of top-down language control that has not penetrated the whole social ladder, as many native languages are still spoken in the rural areas of both Kenya and Tanzania. Swahili, the native

lingua franca, has a strong role in the history of both countries, and possible substrate influences are considered in chapter 2.3. Lastly, the features of East African English are briefly described.

Chapter 3 focuses on the usage, meanings and variation of the past perfect in standard British English and in World Englishes. The chapter ends with recap of the semantic features of the past perfect.

The empirical part of the study is presented in chapter 4. Two subcorpora of the International Corpus of English were used in this thesis: The Great Britain (ICE-GB) and East African (ICE-EA) corpora, the latter of which is divided into Kenyan and Tanzanian subcorpora. They are both small for modern corpora, but since ICE-EA is the only available corpus for East African English, ICE-GB was chosen as the point of comparison as the data in both has been collected the same way, guaranteeing comparability. Extending the methods of corpus linguistics with further digital means is also presented in this chapter, followed by a first glimpse in to the variation of the past perfect in East African English in the form of quantitative results in chapter 5.

Chapter 6 presents the meat of the thesis: The methods used to obtain qualitative results about the use and meanings of the past perfect in East African English as compared to British English. The results are then compared to earlier studies in world Englishes. Chapter 7 then offers conclusions and discussion about the findings and some technologies that were considered for this thesis.

2 The linguistic situation in East Africa

This chapter explores the sociolinguistic history of East African English. In this study, the concept of East African English is limited to the English spoken in Kenya and Tanzania. English is also spoken in the neighboring countries of Somalia and Uganda, which can also be considered East African countries, but corpora from these countries are not yet available, and the role of English is not as salient in those countries. Somalia, specifically, has had much more dominant Arabic influence. Zambia, Malawi and Zimbabwe on the other hand are often considered as part of Central Africa for geographical reasons, and because their Englishes can be distinguished from East African English (Schmied 2006, 188).

In both Kenya and Tanzania, English is a lingua franca and a prestige variety, rather than a first language. Its status is comparable to the status of Latin in medieval Europe: It is the language of higher education and law. English has been a force that unifies these countries, perhaps at the cost of indigenous languages (Michieka 2005, 183), but on closer inspection it can be seen that English is not always the oppressor, as it could be said that African socialist language policy of promoting Swahili in turn took this role in the mid-1960s (Blommaert 2005, 398).

The colonial history of Eastern Africa began rather late, in the 19th century (Schmied 2006, 190). In the beginning, only the coastal towns of the Swahilis were of interest as stepping stones to India. Later, missionaries moved inland following the traditional trading routes of the Swahilis, establishing their language, Kiswahili, as an indigenous lingua franca. Indian presence and perhaps linguistic influence in Eastern Africa came originally in the form of indentured workers for railway construction, but later they found employment in administration and trade. The spread of colonial rule inland resulted in a trilingual situation in the countries: Ethnic groups spoke their own indigenous languages, Kiswahili was used in ethnically mixed centers, and English was the language of administration.

2.1 English in Kenya

The national language of Kenya is Kiswahili, with both English and Kiswahili as official languages. Kenya is, however, a multilingual and multicultural country, with some 40 ethnic groups (Michieka, 2005). Exact figures are difficult to obtain because of the multilingual nature of Kenyan society. Blommaert (2005, 391) on the other hand problematizes the assumption that a certain language 'belongs' to a certain group, as people in multilingual societies tend to know more than one language, and because there may be significant differences in linguistic codes even inside a given 'language'. The number of speakers of English in Kenya is estimated between 0.7 million and 18.4 million in 1985, but Michieka (2005, 179) notes that these figures depend on what level of competence is required for a person to be considered a speaker of English. Schmied (2006, 192) further points out that speaker's self-evaluations of language skills are unreliable, because speaking English gives prestige and thus makes it likely to be reported in excess.

Before the colonization of Eastern Africa, Kenya was part of the region that traded with the Arab world and India. Initially held as a colony by Germany, the coastal areas were handed to Britain in 1890. Kenya is a part of the second diaspora of English, that is, part of the colonial effort of the Commonwealth: Kenya was declared a protectorate in 1895. In 1964 the Republic of Kenya was proclaimed with a single party constitution under Jomo Kenyatta. As a legacy of the colonial administration, the laws of Kenya are written in English and English is the official language of the parliament as well as the High Court. According to Michieka (2005, 175-176), there were two early actors in Kenya that promoted English: the colonial administration and missionaries.

The missionaries wanted to translate the Bible into local languages, but because of the magnitude of the task, only Kiswahili was chosen, as it had already established itself as a lingua franca in the region. The missionaries provided most of the facilities for primary education and thus brought the English language among the people. Currently, in rural areas, the language of

instruction is the local language up to the third grade, and from thereon English is used. In urban areas, English is used throughout. According to Michieka (2005, 182), English is so important that failing English disqualifies students from the next level of education. English is also the sole medium of research and instruction in colleges and universities.

Kiswahili is still a lingua franca in rural areas, although Schmied (2006, 191) notes that Kiswahili is losing ground, whereas English is used mostly in towns and only as a lingua franca. Knowledge of English translates to having been to school, and a longer education means higher standard of English because any success in education is based on the learner's knowledge of English. English is a mark of higher socio-economic class and information expressed in English is considered up to date, modern and accurate (ibid. 177). Radio stations generally broadcast in English and Kiswahili and the daily newspapers are mostly written in English.

Attitudes towards English vary: Those who have succeeded can thank their command of the English language, but those who have been marginalized because of their lack of English feel resentment towards the language. According to Schmied (2006, 191), "broken" or "school" English is a source of ridicule in Kenya. English is the language of the colonizer, but it still continues to be used because of the role it plays which is difficult to replace with local languages, which, for example, do not have a tradition of fiction writing, and wider audiences can be reached with English. Kenyan writers do spice up their English with borrowings from local languages (Michieka 2005, 182).

Kenyans typically speak English with other Kenyans, even if only elite families use English at home, at social functions and with friends, because they rarely have a chance to speak with Westerners, except for workers of the tourist industry, for whom knowing English is a necessity. It is possible that the lack of a standard model outside of classrooms might contribute to a development of a Kenyan variety of English. It seems likely that English will continue to hold its

prestigious position in Kenya, especially since English is establishing itself as an international language.

2.2 English in Tanzania

Most of the are of modern Tanzania was a colony of Imperial Germany, before being designated a British Mandate in the post World War I accords. Tanzania became independent in 1961. The post-colonial Tanzania (then Tanganyika) was the first African country to declare an indigenous language, Kiswahili, the national language, making it an official language alongside with English (Blommaert 2005, 398). In the mid-1960s Tanzania, led by Julius Nyerere, embarked on a massive campaign of nation-building to establish a monoglot African socialist state. The spread of Swahili was paired with the spread of socialism across the population: "The better or purer one's Swahili was, the better socialist Tanzanian patriot one would be" (Blommaert 2005, 399). English was a target of language elimination because it was seen as the language of imperialism, capitalism and oppression, whereas local languages and non-standard varieties of Swahili were seen as vehicles for traditional pre-colonial cultures, all in the way of progress. According to Blommaert (2005, 401), in reality people were unwilling to replace their existing language repertoires with the standardized variety of Swahili. The importance of national linguistic unity that was not generally disputed, but only a few people accepted the idea of individual monolingualism. Tanzania is a very poor country, and therefore the instruments that the state had to work with in the spread of a national language were deficient and limited (ibid, 402).

Swahili was standardized and it was spread among the populace through normative literacy produced by the formal educational system modeled after the Western educational system (Blommaert 2005, 400). Likewise, the model for a 'complete' developed and modern language that post-colonial linguists in Tanzania aimed for was English, which had to remain the language of higher education until a similar level would be attained for Swahili. In the 1980s, Swahili was

adopted by the new multiparty, liberal capitalist state for nation-wide communication. English remained the language of higher education, while Kiswahili is used in primary and adult education, as most information and material was in English. Swahili is now commonly used, and local languages as well as non-standard varieties of Swahili are still spoken. According to Schmied (2006, 192), unlike in Kenya, English is not used in nation-wide politics in Tanzania; It could be characterized as a deeply rooted foreign language. Yet the website of the parliament of Tanzania is in English, as well as all the bills and acts issued by the parliament on its website.

According to Blommaert, (2005, 404-405), English in Tanzania is associated with the core values of capitalist ideas of success: entrepreneurship, mobility, luxury and female beauty. Cases in point are Tanzanian beauty pageants, which are very popular events in the urban parts of Tanzania (Billings 2009). The contestants are expected to know English or at the very least Swahili, and to answer some simple questions on stage, but for most of them speaking English on stage is only possible through memorization. According to Blommaert (2005, 408), English is a resource of relocation and appropriation for the urban youngsters, *wahuni* or 'gangstas', of Dar es Salaam. English and standard Swahili words are used as relexified borrowings in their *Kihuni* slang. It is a repertoire that allows them to culturally 'get out' of Dar es Salaam, their blackness, class, status and marginalization.

To conclude, the ability to speak English does not only imply elite status, it is also often necessary for achieving such status (Billings 2009, 582). It can be used to appear educated, to appear internationally worldly, and to appear to be a successful businessman. It has the capacity to suggest the transnational while remaining local, but only to others in the same community – The English of the East African wannabe will fail to impress a native English speaker. Blommaert (2005, 410) notes that African varieties of English are seen as low prestige varieties outside of Africa, and in this way the use of such English is a source of inequality on the international scale. English is a Tanzanian bourgeois-aspiring resource and the only ticket to the elites (Blommaert

2005, 409). According to Blommaert (2005, 404-405), its lasting prestige functions combined with the extremely restricted access to its prestige-bearing, standard varieties, the access to which is completely conditioned by access to post-primary education, is a source of inequality in Eastern Africa.

It can be seen that the situation of English is similar in Kenya and Tanzania. English is the language of law and the urban higher classes whereas Swahili and local languages are spoken in the rural areas and by the lower classes. As a lingua franca, speakers must choose between Swahili and English. Swahili, and related languages and their effects on East African English will be considered in the next two chapters.

2.3 Swahili – the East African native lingua franca

In Bantu languages, noun classes are marked with prefixes. The prefixes work as articles and demonstratives, even objects when infixed in a verb phrase, and can show number as well as the type of the root noun. Some authors use the word *Kiswahili* for the Swahili language, while others leave out the *ki-* prefix used for languages, hand tools and artefacts, and the distinction does not imply two different languages, merely a choice between native versus English orthography.

Here Swahili is taken as indicative rather than representative of the Bantu language family: The Bantu languages naturally differ to a great degree, for example, some have tones while others do not, some prefer CV syllable structure while others allow consonant clusters and final consonants. The agglutinative nature of the languages appears largely universal, but of course there is variation in that respect too. As there is not enough space here for a thorough investigation of all the languages spoken in Kenya and Tanzania, for the purposes of the current study, Swahili is considered the substrate language or close enough due to its undeniably major role in the area, even though not all East Africans have Swahili or even a Bantu language as their first language.

Swahili is the most widespread language in the Bantu language family, spoken by 5-15 million native speakers and perhaps three times that many second language speakers. Swahili is an official or national language in Tanzania and Kenya and also some other East-African countries, and it is also one of the official languages of the African Union. Swahili and most other Bantu languages are agglutinative, and affixes are used to indicate verb tense, subject, object, negation and other functions (Rieger 2011, 116-118). There is a strong single function – single morpheme correlation. Word classes were already alluded to in the introductory paragraph. The 16 word classes of Swahili are marked by prefixes such as *ki-* (tool) and *wa-* (plural people). Used with the root word *swahili*, these would give kiswahili (Swahili language) and Waswahili (Swahili people). Roughly speaking, these prefixes are added to any word that concords with the root word in a sentence: adjectives, numerals and verb phrases, although the latter often use different forms. Ancestral Bantu languages could have up to 23 word classes with a stronger semantic motivation behind the classes.

Swahili is rich in tense markings and also has a rich tradition of disagreement on the definitions of those markings among linguists. Rieger (2011) created a systematic approach where she rejected earlier attempts to classify the Swahili tense markers using the concept of aspect in addition to tense, instead allowing for an essentially tense-based analysis. Aspect, as it pertains to time distinctions, can produce the opposition inner view (imperfective) vs. outer view (perfective) (ibid. 118), that is, the speaker can be positioned inside the events or looking at them from afar. Tense on the other hand works by positioning events on a timeline in relation to the point of speech (absolute tenses) or in relation to some other point in time (relative tenses). Rieger's thesis is that Swahili favors tense and that aspect is a secondary, even optional, quality of the tense-aspect-modality system (ibid. 117). In effect, tense is grammaticalized and aspect mostly lexical. An overview of the tense markers of Swahili is presented in Table 1 below:

aff	neg	Loogman (1965)	Polomé (1967)	Rieger (2011)	Lindfors (2003)
ta	ta	future	future tense	future	future
a	-	simple present	actual present	definite present	present
na	-	progressive present	indefinite present	indefinite present	present progressive
hu			habitual and recurrent	gnomic present	habitual
me	ku	present perfect	perfective/ resultative	proximate past	perfect aspect
ki	ja	participial	imperf./ continuous	simultaneous	
li	ku	simple past	past tense	remote past	past
ka	ku	historical past	subsecutive	consecutive	

(Loogman's categories taken from Lindfors (2003) and Polomé's from Rieger (2011))

Table 1. Swahili tense markers and grammarians' names for them.

What makes the Swahili tense system interesting for the purposes of this study, and as a substrate language for East-African English, are the past tense markers that have no equivalents in English. For example, -*ka* is a consecutive marker or a narrative tense, for situations where the temporal reference has already been established using some other tense, or context. According to Rieger (2011, 126) it is used for continuing a narrative in logical temporal order towards the future, as in example [1], even if the temporal reference point is in the future. English does not have this kind of a tense, instead, past perfect and simple past are often used when the narrative is continued in its temporal order.

[1] ...a**ka**sema bwana maneno haya a**ka**sikia mtumishi a**ka**fahamu maana yake ...
      said   master words these heard   servant   knew   meaning their
"...Then the master said these words, the servant heard them and knew what they meant, ..."
(Riegel 2011, 126)

Comrie (1985, 26) however argues that the -ka marker is in fact probably not a marker of consecutive events, but simply a perfective past tense, describing events that are wholly in the past, which has come to be considered a special tense due to it being commonly used in situations where the linear order of tenses is likely to follow chronological order of events - a feature, quite simply, of narratives in general. Still, the Swahili narrative marker may be a possible case of tense neutralisation, which is a phenomenon where a tense begins a sentence, and the following verbs that

would be expected to have the same tense instead have another tense, but still the same time reference (Comrie 1985, 102). An example is the (reconstructed) Proto-Indo-European injunctive (Comrie 1985, 104). The first verb in a sentence uses past tense, but the rest lose the past tense prefix, still retaining past reference. Whether or not this is what is happening with the -ka marker of Swahili would depend on 1) whether the -ka marker can be used sentence-initially, and 2) whether it is only used for consecutive events, or also possible in contexts where the events are not ordered. In these kinds of cases Comrie (1985, 28) calls for more examples and clearly distinguishing the meaning and implicatures of tenses.

> [2] A-**li**-keti kitanda-ni.
> "He/she sat down on the bed." or "He/she was sitting on the bed."
> (The grandchild was visiting with the infirm grandmother. The visit is over and the grandchild has left.)

> [3] A-**me**-keti kitanda-ni.
> He/she has sat down on the bed." or "He/she is sitting on the bed."
> (The grandchild is visiting with the grandmother. He/she is still there.)
>
> (Riegel 2011, 129)

It is also worth noting that Swahili has separate remote and proximate past tense markers, *-li* [2] and *-me* [3] respectively (ibid. 128). It is unclear whether perfectivity is in play here, as examples [2] and [3] could be considered to differ in regard to whether the speaker is within the sphere of the event or not. Riegel's view seems to be that the apparent perfectivity and current relevance of [3] only arises from the translation, and the English speaker's language intuition. There are other Bantu languages that make finer distinctions of remoteness, but Swahili has only two such tenses. The remote past could possibly be used in similar situations as the English past perfect, and in many accounts is considered such. According to Rieger (2011, 128), the current relevance of *-me* comes from the distance to the present rather than from grammaticalized perfectivity, and in that case it does not have or create an intervening reference point between the past event and the time of utterance.

[4a] present perfect: nimepiga   I have beaten

[4b] past perfect: nilikuwa nimepiga   I had beaten (lit. I had + I have beaten)

In learner grammars, the traditional Western Latin-based tense categories seem to dominate, for example in [4], taken from an automatic Swahili verb conjugator web site. Nowhere in the literature that I have studied are such compound forms as [4b] mentioned and the repetition of redundant morphemes seems an unlikely feature for a natural language. Comrie (1985, 80) even notes that there is no pluperfect (or future perfect) form in Swahili. It seems these forms arise from a translation-based approach to language teaching, and Rieger (2011, 131) notes that it can even feed back to the Swahili language as East African schools use teaching materials with the Western system crudely forced on Swahili. It seems that generally when a grammar is made, even when the very Swahili-specific tenses like the narrative or gnomic present are included in the verb's paradigm, equivalent forms will be created for the English tenses, even if they do not strictly speaking exist in Swahili. Of course, the actual degree of how much this affects native Swahili is difficult for a non-expert to discern, and the effect may be different in cities vs. rural areas due to the availability and quality of grammars. And for this study, it is not the effect of English on Swahili that is important, but vice versa.

As for the substrate effect of Swahili and other Bantu languages on East-African English, the different tense-aspect-modality system provides ample possibility for variation – of course, the effects may be seen in other forms than the past perfect. Yet the consecutive marker -*ka* may be one possible case that can produce variability. What would the Swahili speaker use to replace it in English? It may be difficult to show any changes with the methodology of the current study, unfortunately, as the sample size is limited by the number of concordance lines that can be analyzed semantically.

Another issue is that Swahili appears to be very tense-oriented (Rieger 2011, 133) whereas English seems to be more balanced between tense and aspect in how events are expressed. How

well does the native Swahili speaker then move from expressing aspect with adverbs to expressing it with verbs when he learns English?

2.4 Features of East-African English

Many features of East-African English can be seen to derive from the substrate influence of Swahili and other Bantu languages, especially on phonetic features. According to Schmied (2006, 192), "aping the British" pronunciation is even considered unnatural. Standard English grammar on the other hand is highly valued, although speakers are not always able to produce such language. There are three general ethnic group specific tendencies in pronouncing consonants that serve as subnational identifiers: merging /l/ and /r/, intrusive or hypercorrectly deleted nasals, and difficulties in producing the English fricatives. The English vowel system is typically contracted into a five-vowel system (/a/,/e/,/i/,/o/,/u/) after the vowel systems of indigenous languages (Schmied 2006, 193). Consonant clusters are simplified by inserting vowels in between or deleting consonants (Schmied 2006, 194). Stress is typically on the first syllable in words. Speech typically has a steady "machine-gun" rhythm, where all syllables are allocated the same time and stress (Schmied 2006, 194).

East African English lexis is characterized by loanwords from Swahili and other indigenous languages; notably political terms, words for traditional clothing, local food and African nature. Some loans have penetrated the English language as a whole, like *safari*, although in East African English it is still used to mean 'journey' in general (Schmied 2006, 195). Standard English lexemes are typically expanded in East African English usage, sometimes due to confusion with similar items, sometimes because general terms are favored to more specific technical terms (Schmied 2006, 196). The simple explanation to this is that as speakers do not have access to linguistic resources, their vocabulary is necessarily more limited and thus they have to cope with a more limited vocabulary.

Inflectional markings are not always added to verbs, but such forms are stigmatized and educated East Africans attempt to avoid them (Schmied 2004, 930). Distinctions between the different forms of strong verbs may disappear because of the contracted vowel system of some speakers, and/or because they seem redundant after time adverbials. This may have left a mark in the spoken subcorpora of the ICE-EA. It may not be possible to find all occurrences of intended past perfects, if some of them have an erroneous form. Schmied (2004, 930) also notes that complex tenses tend to be avoided, especially the past perfect and conditionals, which may be due to the lack of a past perfect and future perfect in Swahili (Comrie 1985, 80). Past tense forms are also used less frequently to express modality than in Standard English.

Next, the past perfect of standard British English and its uses and variation is explored in detail, followed by a survey of the past perfect in some world Englishes.

3 Past perfect - at the Intersection of Tense and Aspect

It is conceivable that during the development of the English language around the world, the tense

system has changed and will change, a brief overview of the English tense and aspect system is in

place so that we are able to recognize any claims of territory that the past perfect may have made in

East African English. Various grammarian's views on the past perfect are explored this chapter

including the standard uses of the form, its adverbial collocation, and variability in World Englishes.

An overall view of what the typical or prototypical use of the past perfect in native English is can

then be established, and also how its use varies globally.

The past perfect has two different uses. It can be used to refer to a past event that has

happened before another past event [1]. It can also be used for expressing past unreality, notably in

canonical if-clauses, as in [2].

> [1] He *had* already *left* the office when we finally made it there.
> [2] If we *had seen* anything strange, we would have let you know.

The past perfect is rare even in native English speech (Quirk (1985, 190) says 10% of finite

verb phrases are past perfect according to "a corpus study"). It is also avoided in non-native

varieties of English, perhaps because it is a difficult form to use correctly (Schmied 2004). The uses

and meanings of the past perfect as well as known variation in other world Englishes are discussed

in more detail in this chapter, but first a common terminological ground must be established.

3.1 Tense and Aspect

Palmer (1980, 34) explains the English verb system with three semantic oppositions: tense

(past or present), phase (perfect or non-perfect) and aspect (progressive or non-progressive). The

English tense forms are formed by combining these oppositions. The progressive aspect is not very

relevant to this study, except when it is used with the past perfect to form the past perfect

progressive, and is thus excluded from this analysis. Comrie (1976, 6) uses the term *aspect* for what

he terms the perfective/imperfective opposition, as well as for other aspectual distinctions.

Tenses can be said to be deictic because they are a system which relates entities, events on a

timeline in this case, to a reference point, which can be the time the sentence is uttered or created or

some ither point in time (Comrie 1985, 14). Absolute tenses are those that use the present moment

as the deictic center, whereas the reference point of relative tenses is some point in time given by

the context (Comrie 1985, 56). Comrie (1985, 36) notes that the term *absolute* is somewhat

misleading because absolute time reference is strictly speaking impossible, since every situation is

relative to some other already established time point. The reference point can be supplied by

adverbials main clauses, subordinate clauses, independent clauses or the context more generally.

Comrie (1985, 58) further notes that the available reference points are all those that are compatible

with the given context and the present is usually available as a reference point, potentially leading

to confusion about the distinction between relative and absolute tense. It is also possible for a tense

to be both absolute and relative at the same time, the English past perfect being a case in point; it's

meaning being a combination of a point situated before the time of utterance (absolute) and a point

in time before that point (relative).

Declerck (1992, 86) explains the tenses by establishing a conceptual timeline and then

dividing the timeline into two "time-spheres", the past and the present. Quirk et al. (1985, 176) use

a similar division, but add that the present tense could instead be called the non-past, as it is also

used for expressing future time. Comrie (1985, 38) considers the meaning of the present tense to be

that a situation holds at the present moment, with additional implicature that that situation may or

may not extend in either direction, to the past or the future. Palmer (1980, 43) similarly considers

the past the marked member of the pair, as the present tense can refer to any period of time that

includes the present moment (also Reichenbach 1947, 292), whereas the past always and

specifically excludes the time of utterance, moving the point of reference into the past. The time

periods referred to by tenses are in no way conditioned by absolute extralinguistic time, for example, the present time-sphere can be anything from a couple of seconds to eternal, accounting for both events that are happening at the time of utterance as in [3] and general truths like [4].

[3] He kicks the ball and he scores!
[4] The sun rises in the east.

Aspect on the other hand is non-deictic because aspects are ways of viewing the "temporal constituency" or "Internal temporal contour" of a situation in Comrie's (1976, 2) words. That is to say, the (subclasses of the) imperfective aspect can describe a situation from the inside, with a beginning, middle, and an end, and a duration. The perfective on the other hand describes a situation from the outside, as a single unanalyzable whole. From this it follows that the present moment must be excluded from the perfective, and thus it must describe either a past situation or a future situation (Comrie 1976, 3), but this is simply an implicature, a by-product, not a part of the meaning of aspects (Comrie 1985, 25). Comrie (1976, 25) further divides the imperfective aspect into *habitual* vs. *continuous*, and the continuous aspect into *nonprogressive* vs. *progressive*. The verb forms, or colloquially tenses, of English and other languages can carry a mixture of tense and aspectual components of meaning, which is why the non-technical use of the term *tense* in school grammars often clashes with a more careful analysis.

It is important to distinguish *perfective* from *perfect*, the former being an aspectual component of meaning and the latter a 'tense' in colloquial parlance, which expresses a relation between two time points, a prior situation and the state that results from that situation, or in other words, the perfect indicates the continuous present relevance of the situation (Comrie 1976, 52). In its most general definition, the perfect aspect means anterior time that precedes whatever time orientation is signaled by tense or other elements in a phrase or stretch of discourse (Quirk et al. 1985, 190). As noted above, Palmer (1980, 36) uses the word *phase* for the perfect – non-perfect opposition, stressing the fact that it expresses temporal relations, dedicating *aspect* to the progressive.

According to Palmer (1980, 36) the present non-perfect refers to a period of time in the present including both past and even sometimes possibly future, as long as it also overlaps the time of utterance [4]. The past non-perfect may similarly overlap a point of time in the past, but does not extend to the present. The perfect forms, then, indicate periods of time that specifically began before and (potentially) continued up to a point of time, from which we get the meaning *past-in-the-past* for the past perfect (Quirk et al. 1985, 195). For Palmer (1980, 42), the present perfect is in no sense a past tense. His justification is that it does not collocate with past time adverbials as seen in [5]. Declerck also includes the present perfect in the present time sphere.

[5] *I've seen John yesterday.

The terms *process*, *situation*, and *event* are used somewhat interchangeably in this study, but Comrie (1976, 13&51) makes a difference between them. For him, a situation can be a static state, or an event or process, which requires input of energy for it to continue. An event is a dynamic situation that is viewed perfectively, with no regard to its internal temporal constituency, and a process is in progress, and therefore viewed imperfectively.

3.2 Points in time

Using the idea of separate points in time, the English tense system can be illustrated by placing the points on a timeline. Both perfect forms indicate periods of time that specifically began before and continued up to a point in time. In the case of the present perfect, this point in time is the time of utterance, but in the case of the past perfect, another point in time before the time of utterance is required. Declerck (1992, 117) calls this point in time the central situation in a past domain of temporal reference. The following diagram comes from Reichenbach (1947, 290). It visually represents all the tenses in English using points in time: The (E)vent point indicates the point in time when the action referred to by the verb happened. The (R)eference point indicates a point in time that is indicated by the sentence, that can, but does not have to, differ from the Event

point, and that helps to align the events in relation to itself in the temporal narrative. The (S)peech point refers to the time the sentence is physically uttered or written. Actions with extended time periods are not included, thus the progressive aspect is not considerd here at all. Also, only the prototypical meanings are presented in this diagram – variation will be discussed separately.

(E)vent point, (R)eference point, (S)peech point or time of utterance

	Past		**Present**		**Future**	
Future constructions			S		ER	You will see them.
Future perfect			S	E	R	You will have seen them.
present			ERS			You see.
present perfect	E		RS			I have seen John this morning.
past	ER		S			I saw him this morning.
past perfect	E	R	S			I had seen him this morning.
Future in the past	R	E	S			He said he would help me.

I have seen John this morning indicates that it is still morning, and that John might still be somewhere around. There is no distinct point of reference. *I saw him this morning* indicates that it is no longer morning at the time of utterance, and furthermore that John is no longer available. The point of reference is in the past. In *I had seen him this morning,* the past perfect alone forces an implicit point of reference between morning and the time of utterance, and the time adverb is unnatural at first glance. Note that *He said he would help me* is not necessarily counterfactual even if it is a likely reading of the sentence, as it is also possible to say *He said he would help me, and he did*. The event referred to by the verb *help* is located after the point of reference established by *he said*.

The concept of current relevance is often used to explain the behavior of the perfect, specifically to explain why the present perfect is not really a past tense. Current relevance means that the situation or action referred to by the verb is relevant to something observable at the present moment (Palmer 1980, 50-51). Another way to see this is that the situation envisaged as a whole includes the present moment. *I have seen John this morning* in the table above is reported with the present perfect to imply that John might still be here, and the speaker would only do so if the period

of time includes the time of utterance, i.e. that it is still morning. Similarly an event reported with the past perfect can be seen as having current relevance at the time of reference (Palmer 1980, 53).

One way to explain the use of tenses in a long stretch of discourse is Declerck's (1992, 86) temporal domain model. Absolute tenses, which are the simple past, present perfect, present tense and future tense, can establish a temporal domain which is directly related to the time of utterance. Relative tenses relate a situation to another situation that has been established by an absolute tense. In complex discourse, each new situation in a temporal domain is bound to one of the situations already introduced into it. The past perfect, then, is a relative tense which needs two reference points, the time of utterance and a central situation in a past domain established by one of the absolute tenses. Declerck (1992, 117) calls this temporal subordination. Quirk et al. (1985, 185) refer to continuing in the same temporal domain as anaphoric use of a tense; the past tense then refers to the time that was referenced by a previous verb. It should be noted that the past perfect can in itself be used to implicitly establish the point of reference in the past in addition to the event explicitly expressed by the construction (Quirk et al. 1985, 196).

3.3 Past in the past

The past perfect is then a 'tense' with the meaning of "E before R before S" (Comrie 1985, 125). It expresses a relation between a past situation and an even earlier situation. How far those situations are from each other is not part of the meaning of the past perfect, but an implicature following from any particular context where the form is used (Comrie 1976, 5). The past perfect can create an illusion of having a meaning of objective remoteness when it is compared to its reference point, but similarly narratives can be created where the past perfect is only seconds away from the point of utterance. The unnaturalness of the example given by Comrie (1985, 26) clearly illustrates how the past perfect, even used alone, begs for a reference point: "The Romans had conquered Britain."

As noted above, the past perfect is one possible way to mark the anteriority of an event relative to another past event, but is not always necessary. The past perfect and the simple past are sometimes interchangeable – the past perfect does not necessarily refer to a more remote past in exralinguistic time than the past tense (Quirk 1985, 186). Adverbials, and conjunctions like *after*, *before* and *when* can be used to define the relations of events (Leech 2002, 72), but mere pragmatic knowledge of the world can also be relied on (Declerck 1992, 86). Quirk et al. (1985, 197) notes that the simple past and past perfect are – predictably – not interchangeable when an event is referred to that requires the past-before-past meaning of the past perfect, such as in the indirect speech construction in [6], where the past perfect indicates a backshift into the more remote past. Sentences [7], [8] and [9] below for example can refer to the same real world situation. Sentence [9] is grammatical regardless of when the boy's father died; the order of events is simply not specified.

> [6] I told her the parcel *had* not arrived. (Quirk et al. 1985, 197)
> [7] I spoke to the boy whose father had died a week earlier.
> [8] I spoke to the boy whose father died a week earlier.
> > (Examples from Declerck (1992, 119))
> [9] I spoke to the boy whose father died. (compare: ...had died.)

3.4 Adverbial collocation

It is worth focusing on adverbials in detail because non-standard adverbial collocations may indicate a changed meaning of the past perfect in East African English and will help in deciphering the context of the corpus hits. Palmer lists (1980, 50) adverbials that collocate with the perfect. These include adverbials beginning with *since* (since Tuesday, since we met). They indicate the starting point of the period of time that leads up to the point of reference. Adverbials beginning with *for* (for an hour) are typically used with past perfect, but not exclusively. Comrie (1985, 32) notes that in general the perfect is incompatible with time adverbials that refer to a specific moment or stretch of time wholly past and that this constraint even overrides considerations of present relevance. It is however possible to use such a specification of past time with the perfect if it

includes the time of utterance, such as *today* or *this morning*. Comrie (1985, 79) further notes that the constraint on definite time adverbials with the English perfect does not carry over to the past perfect. The past perfect is thus very much compatible with such adverbials.

According to Reichenbach (1947, 294), adverbials that determine a specific time, such as *now, yesterday, in the year 1906* etc. always refer to the reference point in the sentence. In tenses where the event and reference points coincide, the time adverbial refers to the event point only by chance. Comrie (1976, 56) on the other hand seems to think that the adverbials can refer to either the reference point or the event point, giving "Bill had arrived at six o'clock" as an example of the ambiguity. Six o'clock could be the vantage point from which we are observing the results of earlier events (Bill may have arrived at five o'clock, but still be there at six o'clock). This could be considered a perfect-in-the-past, that is, a past state which results from an even earlier situation and thus at six o'clock it would be said that Bill has arrived. Alternatively, "six o'clock" might refer to the time of Bill's arrival. In that case Bill's arrival preceded some other past situation and could be considered a true past-in-the-past. The hearer's extralinguistic understanding of the context must then be relied on to deconstruct the order of the events, there being ambiguity in the grammar of adverbials co-occurring with the past perfect (Comrie 1985, 66)

Time adverbials which compare time points, such as *when*, *before* and *after,* refer to the reference point, which carries the time position information in the sentence. Thus in sentences like [10], where the two clauses are linked with the adverb *before,* both clauses share a reference point, which also happens to be the same point in time that the event point of the second sentence refers to. Here, and in other cases where the adverb provides enough information about the temporal relations, both sentences could also be in the simple past (Quirk et al. (1985, 196) and Reichenbach (1947, 296)) and the past perfect seems perhaps a bit superfluous.

[10] He had telephoned *before he came*.
          1: telephone   E   R   S
          2: arrival   E,R   S

According to Quirk et al. (1985, 197), on the other hand, there is a difference in meaning depending on where adverbials are placed in a sentence with the past perfect. Adverbials placed initially with the past perfect often identify the time of reference as in [11], whereas an adverbial in the final position typically refers to the event point [12]. The examples below, seen from Reichenbach's strictly logical point of view, are identical. Seen through Declerck's temporal domain, the first clause in each sentence establishes the temporal domain, and the following sentence expands the sequence of events upon it. The slightly different interpretations then arise from the ordering of the clauses – what the speaker thinks more important to mention first – and the understanding of the extralinguistic world.

[11] *When the police arrived*, the thieves had run away.
    1: arrival   E,R   S
    2: escape   E   R   S
[12] The thieves had run away *when the police arrived*. (Quirk et al. 1985, 197)
    1: escape   E   R   S
    2: arrival   E,R   S

3.5 counterfactual / modal past perfect

Counterfactuality is a type of epistemic modality, which is concerned with whether or not a proposition is true (factual) or false (counterfactual), or the possibilities and probabilities in between these extremes (Declerck 1992, 351). When the past perfect is used with modal meaning, it expresses anterior counterfactuality, losing the effect of the perfect aspect in its temporal meaning (ibid, 355). Similarly a past modal expresses simultaneous counterfactuality, the difference can be seen in [1] below. Related to this losing of time reference in conjunction with modal meaning, Palmer (1980, 96) adds that modal auxiliaries in their past forms rarely mark past time. The modal past perfect can be used in various contexts, most notably in conditional clauses. A condition is counterfactual if the speaker makes it clear that the condition was not, is not or could not be fulfilled (Declerck 1992, 425). It should be noted that in formal English, conditions can be expressed without *if*, using zero conjunction and inversion, such as in [2] (Declerck 1992, 425).

[1] I wish I knew / I wish I had known.
[2] Had the body been discovered… = If the body had been discovered…
[3] I wish I had been there this morning.
[4] I wished I had been there this morning.
[5] He talks as if he had seen it himself.

In [3], there are two kinds of modality. *I wish* signals non-epistemic modality (volition), and the past perfect signals that the event did not occur (counterfactuality). Another way to see this is that the wish implies counterfactuality of its object (having been there), and as the time reference requires a construction that can express the past as well as counterfactuality, the past perfect is the best choice here. The timeline is the same as with a simple past construction as no extra reference point is implied by the construction, which also shows in the intuitive ease of accepting the time adverbial *this morning*. In [4], the simple past in conjunction with the past perfect modal creates a time reference similar to a past perfect. [5] is a special case with *as if* that requires a modal use of a tense, not strictly speaking a condition. Other such situations are wishes beginning with *if only* and *I wish* (Palmer 1980,149 and Quirk et al. 1985, 1011). The modal past is also used for tentative or polite questions and requests (Palmer 1980, 147), but the modal past perfect does not seem to be possible here, as it pushes the time of asking into the past and cannot thus be a request at the time of utterance.

[6] If he eats his food, I will be a happy mommy.
[7] If he came I would be happy.
[8] If she had seen me, I would have lost the game.
[9] If you had come tomorrow instead of today, you wouldn't have found me at home.
(Declerck 1992, 431)

Conditional sentences are often taught in schools using three models, which Declerck (1992, 426) calls the canonical types. An important thing to notice about conditionals of the canonical type 3, example [8], is that they always express counterfactuality and that both clauses can refer to the past, present or future, as in [9], depending on the adverbs used (ibid, 431). According to Palmer (1992, 141), there is plenty of ambiguity in the tense reference of even the canonical types, if time adverbials and context are not used to elucidate.

It is at least conceivable that the mixing of modality and tense might be difficult to learners of English, and that the variability it causes in the time reference of the past perfect can pave way for learners, both first language and second language, misanalysing it in the absence of a theoretical framework about tenses, aspects and modality. Unfortunately, proving this is the cause of any possible variability in the use of the past perfect may be difficult.

3.6 The past perfect in World Englishes

Next, a brief survey of literature concerning the variation of the tense/aspect system in World Englishes is presented, focusing mostly on the British Isles and North America. Slight variation exists in the use of the tenses and specifically in the use of the past perfect tense, in meaning as well as in form. It should be noted that only some varieties that are known to have variation in tense usage are presented here, and this short presentation should not be taken as an accurate description of exactly how widespread the phenomenon is, but rather as a short overview of what is possible in world Englishes.

3.6.1 Celtic contact

In Scottish English, tenses in conjunction with adverbs [1] take over functions that in standard English are expressed with tenses only (Miller 2004, 56). The past perfect is absent from some subordinate clauses and is rare in main clauses. Anteriority can be expressed with *once* + simple past, as in [2], where standard English would use the past perfect to signal anteriority (Häcker 1999, 85).

[1] The electrician has just phoned vs. The electrician *just phoned*.
[2] *Once* her children left home, she got a job. (ibid.)

Irish English has a well-known case of variation in its tense/aspect system, the medial-object perfect (Filppula 2004, 74), where the object is inserted between the participle and *have*, and the focus is on the resulting state of an action. According to Miller (2004, 56), the resultative structures

can also be found in Scottish English. The *after* perfect [4] is also reminiscent of the Scottish use of adverbs and simple past for perfective meanings. In addition to these two, four other constructions exist in Irish English that have perfective meanings.

[3] I have it forgot.
[4] You are after ruining me. (Filppula 2004, 74)

3.6.2 The Americas

In early African-American English, the past perfect was used very much like in standard English, but there was a large number of nonstandard participle forms due to consonant cluster reduction, participle switch and double marking (Kautzsch 2002, 346). The perfect could also be formed with *done* + past participle [5], with both present and past reference points, thus being able to replace the past perfect. According to Schneider (2004, 1105), using *had* + past participle for simple past meaning in AAVE is a recent development.

[5] I done told you. I had done quit. (ibid.)

In Chicano English (Latin-American English) similar developments have taken place as in AAVE. Irregular verbs have been regularized due to consonant cluster reduction, which has resulted in unmarked past tense forms, and past tense forms are used instead of past participles (Bailey & Santa Ana 2004, 377). The past perfect is sometimes used with simple past meaning, as in [6], something that also happens in East Caribbean English (Aceto 2004, 448) and could be either AAVE influence or from settler dialects that already had the feature.

[6] I don't know if it was my son or my nephew *that had told me*.
(Bailey&Santa Ana 2004, 377, originally in Fought 2003, 97-98)

3.6.3 Africa

Similarly the past perfect is used with simple past meaning in Fiji English, especially in print (Mugler & Tent 2004, 782).In contrast, in Ghanaian English, there is a tendency to substitute past perfect for the present perfect (Huber & Dako 2004, 855). In Black South African English, the past

perfect is often replaced with the simple past in certain contexts, e.g. reported speech (Meshtrie, 2004, 963). According to Comrie (1985, 26), speakers of West African languages with grammatical distinctions of degrees of remoteness in the past, that is, separate remote past and proximate past tenses, often treat the English past perfect as a translation equivalent of their own remote past tense.

3.6.4 India

Devyani Sharma's study (2001) of the changes in meaning of the past perfect in Indian English was a comparative corpus study on the past perfect in the Kolhapur corpus of Indian English, compared to the Brown (American English) and LOB (British English) corpora, specifically the press and bureaucratic registers. Sharma's aim was to find out whether the inherent variation of the meaning of the past perfect is different in Indian English as compared to native varieties, where there is some variability towards using the past perfect with simple past meaning. Indeed, differences were found: There were 5% other meanings for past perfects in British English, compared to 22% of other meanings among past perfect forms in Indian English (ibid. 356). These other meanings include present perfect and preterite meanings as well as an anterior meaning with no intervening reference point.

According to Sharma (2001, 344), the new use of the past perfect in Indian English appears to derive from nonstandard signaling of pragmatic viewpoint and tense orientation. What has made this development possible is in part the ambiguity inherent in native English(es) (ibid. 344), but first language interference was also considered in the study. Sharma's (2001, 370) speculation for the reason for the variation is that the past perfect has become a generalized marking of remoteness and completion in Indian English, regardless of the reference point of the narrative (see chapter 3.1). There may also be an emergent anterior completive marking using the past perfect form at play, making the distinct reference point unnecessary.

The tense contexts in which the past perfect appears were also mapped in the study in order to see their effect on the usage of the past perfect. In British and American English, a past both follows and precedes the past perfect in about 75 % of the time, while in Indian English, the percentage is closer to 60 %, while the perfect and present tenses show an increase (ibid. 357). Disambiguating adverbials were also studied and found to be least common in Indian English (30 %) and most common in American English (59 %), with British English falling in between (36 %) (ibid. 358). Sharma comments that the absence of time adverbials may facilitate the reinterpretation of tense meanings and may lead to change. The past perfect was found to occur with reported speech with similar frequencies in all varieties studied (ibid. 359).

According to Sharma (2001), the past perfect is used in Indian English with simple past meaning 22% of the time, and also the past perfect is used less than in British or American English. She hypothesizes that the altered semantics of the past perfect is caused by a tendency to shift the reference point easily [7], whereas in British English the reference point is mostly permanent. In addition, the past perfect distances a completed, past event from the narrative focus [8] (ibid. 365). The past perfect may also be becoming an anterior completive marking, making a distinct R-point unnecessary (ibid. 370).

[7] This is the second time that such an object had been sighted here.
[8] In the past, there had been criticism in the J.P.P. Executive Committee over the issue...

From Sharma (2001, 363)

3.6.5 Variation in standard British English

According to Sharma (2001, 356) the past perfect is used with non-standard meaning 5% of the time in the press register of British English, but it would be interesting to see what the figure is for British English in general.

As for adverbial co-occurrence, Milroy (1992, 7) notes that *just* and *already,* adverbs that are often said to only occur with the perfect tense (Palmer 1980, 50), can also occur with the simple past tense in some (Northern) varieties of British English as well as American English, even if the

characterization is accurate for Southern British dialects. It would be interesting to see if the use of time adverbials correlates with the variation in tense-aspect-modality systems cross-linguistically.

3.7 Summary of Meanings and Variation

In conclusion, the main meaning of past-before-past is strong in all the world Englishes, and there is no competing form with the same meaning. The same form however varies to some degree with remote past and simple past meanings. In addition, the standard usage in hypothetical if-clauses gives the modal meaning of past unreality. In narrative, the linear order of tenses is likely to follow chronological order of events. The past perfect can then be used to narrate in another temporal order, and from there on the choice of tense is relatively free, the past perfect no longer necessarily forcing another step backwards.

It can be seen that in the global context, the past perfect has some variability in usage and meaning. As English has come in contact with Celtic languages, new forms such as medial-object perfect and after perfect and other periphrastic forms with adverbs have been added to the tense-aspect-modality system, perhaps displacing the standard forms to some degree. In the Americas, notable variation has happened in form due to phonetics, and to some degree in semantics also, with the past perfect being used with simple past meaning. In Africa, it seems that the past perfect is not used as much as in native English, perhaps in part because its functions have been taken over by other tenses. In Indian English the past perfect has in part become a marker of remote past and completion, due to a shifting point of reference, but the past perfect is still mostly used in its standard meaning.

It has been noted previously that the past perfect has uses in which there is no distinct reference point, such as the narrative usage, for which usage, some languages, like Swahili, have the narrative marker. Likewise speakers of languages where there is no tense with a past-in-the-past meaning, such as Swahili, may find it difficult to grasp the concept; Indeed, as it was noted in

chapter 3, speakers of languages with grammatical distinctions of degrees of remoteness in the past often treat the English pluperfect as a translation equivalent of their own remote past tense. This could be considered one possible source of confusion for second language learners of English. Similarly confusion may arise because the past perfect form loses its perfectness in counterfactual usage. The meaning and context of use are arguably different, yet the form remains the same. Perhaps this has potential to confuse learners, leading to variation.

In addition to these, mere extralinguistic information that is known to the speaker and the listener can be relied on when establishing time reference. For example, this information may include the point of reference. Tenses can then be used fairly freely, if such use does not clash with the situation too much. People who learn the language in a natural manner are likely to pick up English in situations where correct language is not as important as in written discourse. It should be noted that the past perfect can in itself be used to implicitly establish the point of reference in the past in addition to the event explicitly expressed by the construction (Quirk et al. 1985, 196). Because the implicit reference point might go unnoticed by the hearer, this could be interpreted as a remote past.

4 Materials and Methods

In this chapter, the empirical steps of the research are presented. The concept of corpora is first explored and the corpora used are presented, followed by discussion on the methods of searching the corpora and a presentation of methods that can significantly speed up sorting the data for faster handling later on. In chapter 5, preliminary quantitative findings concerning the usage of the past perfect in British English and East-African English are presented from each corpus separately and then compared, forming a starting point for qualitatively comparing the varieties in question and finally delving deeper into the meaning of the past perfect in chapter 6.

4.1 Corpus linguistics

This chapter is a short introduction to corpus linguistics. It begins with the definition and a brief history of corpora, followed by the pros and cons of corpus use and what needs to be taken into account when using corpora for linguistic research.

A linguistic corpus is a "large systematic collection of texts stored on computer" (Biber et al. 2007, 24) "assumed to be representative of a given language, dialect, or other subset of a language, to be used for linguistic analysis" (Francis 1982, 7). The corpus can be searched for a search term, which can be retrieved in its sentence context, as it appeared in the original text. If the search term appears many times, all the hits can be retrieved from the corpus, resulting in a collection of short snippets of text. It is possible to compare the frequencies of different words in the corpus and therefore to see which words are favored by native speakers, assuming the corpus is representative of the language. While the armchair linguist relies on his own understanding and intuition of the language, the corpus linguist relies on real world data as well as his intuition (Partington 1998, 2). Thus corpus linguistics allows the study of other languages and varieties than the ones the linguist is deeply familiar with.

Early corpora were cumbersome as they relied on index cards, which were compiled and searched by hand, in practice limiting the corpus size to a single work of literature and thus the findings to be relevant on one writers use of language. The availability of computers has allowed the discipline of corpus linguistics to grow rapidly since the 1960s (Svartvik 1992, 8). The automation of the collection of digital texts on the internet has allowed corpus sizes to grow up to several billion words, such as the Corpus of Global Web-Based English. Texts that are not available online still make compilation of large and representative corpora difficult.

The advantage of using corpora for studying language is that the corpora have already been compiled by someone else and the researcher therefore does not need to the laboriously gather language data. By using electronic corpora, linguists are able to conduct an unlimited number of queries in a short time. Corpora offer researchers a massive amount of linguistic materials from different dialects, registers and styles, and they also enable anyone in the world to use them easily in the case of freely accessible corpora, an important aspect of the democratization of knowledge and science.

According to Svartvik (1992, 8), some of the most notable advantages of corpora are: Corpus data is objective and verifiable, and it is possible to find out the frequency of occurrence of words and features and their contexts of use. Studies of variation, dialect, register and style are almost impossible without corpora, especially for non-native speakers.

Still, there are problems with the use of corpora. It is difficult to define how big a corpus should be in order for it to be considered representative of a given language of variety. Corpus data can contain mistakes. Frequencies are straightforward, if there are enough hits, and in case two corpora are being compared, the corpora must be similar in terms of text collection. In a small corpus, a single writer's or speaker's style can influence word frequencies. Semantics is difficult as the linguist must use his intuition when interpreting data. Corpus findings may also be trivial, and the weight and relevance of the findings must also be gauged by the linguist (Lindquist 2009, 10).

Different corpora can have different criteria of transcription of spoken materials, which can sometimes make comparing varieties problematic. The corpora used in this study follow the same set of rules of compilation, as the ICE corpora are all made on the same principles in order to guarantee comparability.

Two further problems with corpus-based studies are the problems of *recall* and *precision*. Salton (1973, 262) defines *recall* as the "proportion of relevant matter retrieved" and *precision* as the "proportion of retrieved material actually relevant." Ball (1994, 295) considers the recall problem as the more serious one. Of the two, the problem of precision is much easier to notice by just looking at the search results to see if there are irrelevant hits among the useful ones, while imperfect recall may not even be testable.

To overcome the problem of recall in this study, many searches were made with all the possible forms of the past perfect (had, hadn't, 'd, had had). In addition, Each past participle in the corpora was tagged in a very rudimentary way using regular expressions, and the tags were included in the searches. Thus it is fairly certain that all past perfect forms in the corpora were captured, but also some irrelevant ones, which had to be manually deleted. The method used favors recall at the expense of precision, especially at the early automated phases of searching, but as Ball (1994, 295) notes, sometimes a perfect balance between the two is not possible.

4.1 The corpora

The ICE offers a great way to study world Englishes, as each component is compiled following a common corpus design. Two components of the International Corpus of English are used in the study: The Great Britain and East Africa corpora, each containing roughly one million words of spoken and written English produced after 1989, in text only and without any part-of-speech tagging. The East African corpus contains texts from Kenya and Tanzania. A corpus for the neighbouring Uganda was being compiled but was not yet available when this study was written.

	Written	Spoken	Total
ICE-GB	426912	637632	1064544
ICE-EA	720771	570357	1291128
Kenya	360845	351673	712518
Tanzanya	359926	218684	578610

Table1. Number of words in corpora studied

As can be seen in Table 1, the corpus sizes vary. Most notably the Tanzanian spoken subcorpus is very small for a modern corpus. All frequencies reported in this study have been normalized to N/100,000, that is, tokens per 100,000 words, for making it possible to make meaningful comparisons based on the frequencies, which raw frequencies per subcorpus would not allow. So, for example, if there are 52 hits in the ICE-EA corpus for a given search word, its normalized frequency would be 7.21 by the following formula. If the corpus was 100,000 words long, about 7 hits for the same search word would be expected.

$$\frac{N}{100,000} = \frac{52}{720,771}$$
$$N = \frac{100,000 \cdot 52}{720,771}$$
$$N = 7.21$$

4.2 Data extraction

The format in which the ICE-EA corpus is provided is slightly problematic for the purpose of creating statistics based on it. While ICE-GB comes in either one file containing the whole corpus, two files containing the written and spoken subcorpora, or as individual text categories per file where the category identifiers also serve as the filenames (e.g. S1A-007.txt), ICE-EA is provided only in individual text categories per file, with a naming scheme that does not use the category identifiers (e.g. creative-1K.txt). The final letter, either *K* or *T*, does however identify the country from which the material has been gathered. For the task of creating statistics based on the searches in ICE-EA, an excel sheet was made with the file names and corresponding countries and registers. It was then possible to create statistics based on the file each hit was found in, and to use the excel

sheet to divide the results per country and register. For ICE-GB, the two file version of the corpus

was used, which then made it possible to categorize the corpus hits based on the file and thus

register they were found in.

All past participle form of verbs in the corpora were tagged with the tag *[VVN]* in order to

facilitate finding relevant past perfect forms among all the hits containing *had* in the concordance. A

list of 620 conjugated English irregular verbs was downloaded from

http://www.usingenglish.com/reference/irregular-verbs/ and edited using the *replace* tool with

regular expressions in Notepad++, a powerful raw text editor, to weed out all other forms than past

participles. The final list thus obtained contains 766 forms due to alternative spellings and forms of

some of the verbs. This list was then used to create a regular expression, essentially a powerful

standardized search and replace function available in many programs and programming languages,

for tagging each irregular and regular past participle in both corpora using the program

PowerGREP, which allowed for complex search and replace strings to be used on all the files of the

corpora. The regular expression used is as follows, with the list of past participles abridged: (See

Appendix 1 for the full list)

> Search string: \b(\w*ed|abode|abided|abidden|...|wrung|written|zinced|zincked)\b
> Replacement: [VVN] \1

This matches any regular past participle ending in *-ed* (*\w*ed*, where *\w* matches any

alphanumerical symbol and * allows for many instances of such symbols, allowing to match the

verb body) as well as the list of irregular forms. Such word forms where *ed* is contained elsewhere

in the word, such as *bedstead*, are avoided by using the *\b* expression, which matches word

boundaries, that is, between letters and whitespace or punctuation. The replacement simply tells the

program to repeat the verb form found and append a *[VVN]* tag before it. The list of irregular past

participle forms necessarily excluded *had*, as it would have made analysing the concordance

difficult. *Had had* therefore needed to be handled separately.

Tagging the corpora in this way has some potential adverse side effects because it blindly tags all forms matching the search string, even where the form is not grammatically a part of a past participle. However, for the limited concordancing of the corpus that is necessary for the current study, it is possible to manually remove the forms that were tagged erroneously. Tagging the corpora using part of speech tagger software was deemed impractical, as easy to use programs are costly and the available free software requires programming skills to implement.

Both corpora were tagged and searched in the same way in order to ensure comparable results. In effect, new subcorpora were created containing only instances that looked like past perfects in form, and after manually weeding out non past perfects, only past perfects remained. Next, the process of concordancing the two corpora is explained, taking ICE-EA as an example and then reporting the findings from ICE-GB, after which the quantitative results from both corpora will be compared.

4.3 ICE-EA

Having tagged all past participles, a concordance was made using Wordsmith 6 with the search word *had,* with the context word *[VVN]* within R6, narrowing the search results to 2,832 hits out of 4,146 hits that would have been obtained without the context word. Such instances where there is no past participle after the *had* cannot be past perfects, so this tagging effectively removes some unwanted hits, but not wanted hits, assuming the list of past participles was exhaustive and that there are no more than six words between the *had*s and the main verbs. This is not necessarily always the case, as modal past forms and questions may have a potentially long subject between *had* and the main verb, as in [1]. It is also possible to place a long adverbial between *had* and the main verb, as in [2]. This might reduce the number of hits in the written corpora, making the modal use seem less common in such registers where complex language is used. However, as the

procedure was the same for all corpora compared, the effect in apparent differences between varieties is negligible.

[1] The question is whether **had** the learned trial magistrate **considered** this...
[2] ... long dry spells **had** for the past three years **affected** food production in Rombo..
[3] You had your backbone injured

The remaining unwanted instances were manually deleted. The [VVN] also made manual deletion quick with the use of alphabetical sorting in the *concordance* column in Wordsmith 6, as instances of *had* immediately followed by a past participle can only be past perfects. 2,067 such instances were found. Unwanted uses of *had* still left in the data included causative [3] and obligative instances (*had to*) where a past participle followed within 6 words, but for example in another sentence. After manual deletion, only 2,595 past perfects and modal past forms were left in the concordance. This included some duplicate entries, mainly from written court texts that contain the same text but which, according to the naming scheme of ICE-EA, originate from different countries (judgmt-T.txt and judgmt-K.txt).

A separate concordance was made with the search word *had* with the context word *[VVH] had* within R6, after having tagged the whole corpus again with *\b(had)\b → [VVH] \1* to obtain all past perfects of the verb *have*. This tagging was necessary, because when using Wordsmith, it is not possible to search with a context word that is identical to the search word, because the context always begins from L0 or more, thus always including the search word itself. This search produced 92 additional hits, of which 21 were past perfects of the verb *have*. This concordance was merged with the previous one for all subsequent analysis.

Country	Spoken	Written	Total
Tanzania	198 (18,3%)	882 (81,7%)	1080
Kenya	672 (43,8%)	864 (56,3%)	1536
Avg EAE	435 (33,3%)	873 (66,7%)	1308

Table 1: Distribution of the non-contracted past perfect or modal past in ICE-EA.

1080 non-contracted past perfects or modal past forms were found in the Tanzanian subcorpus

and 1536 in the Kenyan subcorpus (Table 1). One third of the hits were found in the spoken parts of

the corpora. In the Tanzanian corpus, only 18,3 % of the hits came from the spoken corpus and 81,7

% from the written corpus. In the Kenyan corpus, the distribution was more balanced with 43,8 %

hits in the spoken corpus and 56,3 % in the written corpus.

Country	Spoken	Written	Total
Tanzania	2 (12,5%)	14 (87,5%)	16
Kenya	15 (50,%)	15 (50,%)	30
Avg EAE	9 (37,%)	15 (63,%)	23

Table 2: Distribution of the contracted past perfect
in ICE-EA.

Another concordance was made using the search word *'d* with the context word *[VVN]*

within R6 in order to also find all non-negative contracted past perfects (Table 2). This produced 53

hits, out of 158 possible occurrences of *'d*. Searching for *'d* with context word *[VVH] had*

produced no additional hits. A concordance was also made with the contracted negative form *hadn't*,

which produced 15 hits, 13 of which were pertinent. Non past-perfect forms were manually deleted,

leaving 46 past perfects. As can be seen, the distribution is almost identical to the distribution of the

non-contracted forms. 20 of the 46 hits came from the creative writing register.

Country	Spoken	Written	Total
Tanzania	200 (18,2%)	896 (81,8%)	10
Kenya	687 (43,9%)	879 (56,1%)	1566
Avg EAE	444 (33,3%)	888 (66,7%)	1331

Table 3: Distribution of the past perfect and modal
past in ICE-EA.

In total, there were 2,662 instances of the past perfect and modal past in the ICE-EA corpus

(Table 3). This preliminary analysis confirms that the past perfect is more common in written East

African English than in spoken East African English. The percentages would also seem to indicate that the past perfect is more common in Kenyan English than in Tanzanian English. The low percentage of hits in the spoken part of the Tanzanian corpus compared to the written part may perhaps be explained by lower overall command of the English language in Tanzania.

Around 2 % of all past perfect forms in the ICE-EA corpus were contracted forms. There is no remarkable difference between written and spoken or Kenyan and Tanzanian English in the rate of contracted forms to full forms.

4.4 ICE-GB

ICE-GB was tagged and searched in the same way as ICE-EA. Searching with *Had* with the context word *[VVN]* within R6 produced 1664 hits, 700 in the spoken subcorpus and 964 in the written subcorpus. After manual deletion of non-pertinent hits, 1451 hits were left in total, 563 in the spoken subcorpus and 888 in the written subcorpus. To find occurrences of the past perfect of the verb *have*, *had* was used as the search word with context word *[VVH]* *had* within R6, which produced 99 hits,72 in the written subcorpus and 27 in the spoken subcorpus. After deletion 36 hits remained, 18 in each subcorpus. In total, there were 1487 non-contracted past perfects, 581 in the spoken subcorpus, and 906 in the written part of ICE-GB (Table 4).

Country	Spoken	Written	Total
GB	581 (39,1%)	906 (60,9%)	1487

Table 4: Distribution of the non-contracted past
perfect or modal past in ICE-GB.

Concordances were made using the search word *'d* with the context word *[VVN]* within R6 and *'d* with context word *[VVH]* within R6, in order to also find all non-negative contracted past perfects. In total after deletion of non-past perfect cases, these searches produced 437 hits, 323 in

the spoken part and 114 in the written part of ICE-GB. This also found some negated past perfects

of the type "I'd not eaten...".

Country	Spoken	Written	Total
GB	378 (73,8%)	134 (26,2%)	512

Table 5: Distribution of the contracted past perfect
in ICE-GB.

To find the rest of the negated and contracted past perfects, a simple search with *hadn't* was

made, which produced 75 hits, 55 spoken and 20 written. In total, there were 512 contracted past

perfects or modal pasts in the ICE-GB corpus, of which 378 in the spoken register and 134 in the

written register (Table 5). Around 39 % of spoken and 13 % of written past perfect forms in the

ICE-GB corpus were contracted forms.

Country	Spoken	Written	Total
GB	959 (48,%)	1040 (52,%)	1999

Table 6: Distribution of the past perfect and modal
past in ICE-GB.

Duplicates were removed from the concordances, changing the figures somewhat. Removing

duplicates means that two hits that are exactly the same, caused by possible duplicate texts in the

corpora, are eliminated from the concordances, eliminating the bias that such duplicate texts might

create in the results.

5.1 Quantitative results

Next, the two quantitative data sets obtained are compared and illustrated with tables and charts, first with raw frequencies, then with normalized frequencies. This should show how differently the past perfect is used in the two varieties.

5.1 Distribution of the past perfect in the subcorpora

The distribution of non-contracted forms of the past perfect in table 1 already gives a good picture of the differences in past perfect usage between British and Kenyan English on the one hand, and Tanzanian English on the other hand. In Kenya, the past perfect is used, from a quantitative viewpoint, in the same way as in Britain.

| Country	Spoken	Written	Total
Tanzania	198 (18,3%)	882 (81,7%)	1080
Kenya	672 (43,8%)	864 (56,3%)	1536
Avg EAE	435 (33,3%)	873 (66,7%)	1308
GB	581 (39,1%)	906 (60,9%)	1487

Table 1 and Figure 1: Distribution of the non-contracted past perfect or modal past in ICE-EA and ICE-GB.
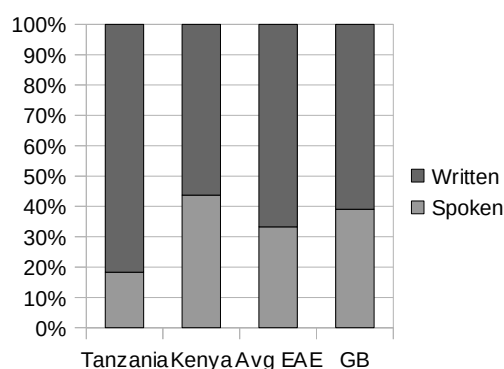
As can be seen in table 2 and figure 2, the contracted form is more frequently used by British speakers of English. The absolute numbers of hits in the corpora are also worth noticing, as the British use significantly more contracted forms than East-Africans. See also table 4 for percentages of contracted forms.

Country	Spoken	Written	Total
Tanzania	2 (12,5%)	14 (87,5%)	16
Kenya	15 (50,%)	15 (50,%)	30
Avg EAE	9 (37,%)	15 (63,%)	23
GB	378 (73,8%)	134 (26,2%)	512

Table 2 and Figure 2: Distribution of the contracted past perfect in ICE-EA and ICE-GB.
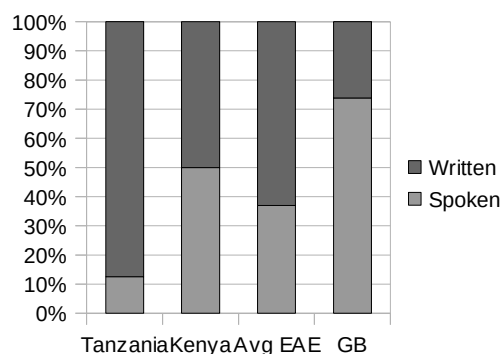
5.2 Differences between written and spoken subcorpora

Variation between written and spoken registers is evident in the data: The past perfect is used more in written discourse, as can be seen in figure 3. There are differences between countries too, Tanzanians using the past perfect the least generally and specifically in spoken language (see table 3). Such a noticeable difference validates considering Kenyan and Tanzanian English separate varieties, at least regarding the tense system.

Country	Spoken	Written	Total
Tanzania	200 (18,2%)	896 (81,8%)	1096
Kenya	687 (43,9%)	879 (56,1%)	1566
Avg EAE	444 (33,3%)	888 (66,7%)	1331
GB	959 (48,%)	1040 (52,%)	1999

Table 3 and Figure 3: Distribution of the past perfect and modal past in ICE-EA and ICE-GB.
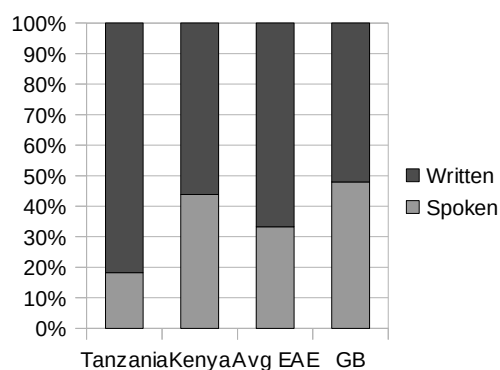
Contracted forms are used most in British English (Table 4), perhaps pointing towards familiarity with the language. Schmied (2006, 192) notes that correct grammar is valued by East-African speakers, which may explain the dislike to using contracted forms. A phonetic reason may be the tendency towards CVCV syllable structure of Bantu languages and aversion of consonant clusters (ibid.).

| Country	Spoken	Written	Total
Tanzania	1,00%	1,56%	1,46%
Kenya	2,18%	1,71%	1,92%
Avg EAE	1,92%	1,63%	1,73%
GB	39,42%	12,88%	25,61%

Table 4: Percentages of contracted forms out of all hits per country/subcorpus

5.3 Normalized frequencies

The numbers of hits obtained were normalized per 100,000 words, as shown in figure 5. It can be seen that the contracted form is more frequently used by native speakers of English than by East African speakers, and there are no differences between the two East African varieties in this respect. As expected, the contracted form was more common in spoken than written British English, yet the lack of difference in the frequencies from the different registers in the East-African corpus is surprising. This may be a caused by hypercorrection by East-African speakers in order to appear more educated, or lower overall fluency in English.

Figure 5. Normalized frequencies (N/100,000)

of contracted and non-contracted past perfects and past modals in the different corpora
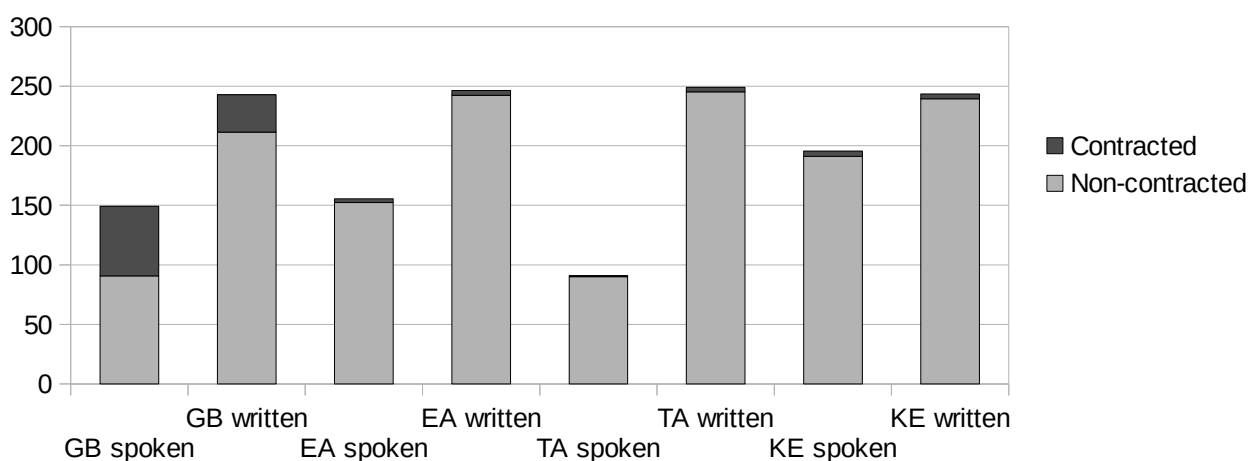
The past perfect is equally common in the different written corpora, yet less common in the Tanzanian spoken corpus than in the British spoken corpus, and surprisingly more common in the Kenyan subcorpus than either of the above (Figure 5). When the East-African Englishes are seen as

a singular variety, the frequency of the past perfect is on par with British English, the only difference being the percentage of contracted forms to non-contracted forms. Yet the fact that the frequency of past perfect use is so different in the spoken East-African varieties seems to point in the direction of considering them as distinct varieties of English. Possible explanations might be the level of proficiency of English in Kenya and Tanzania, or differing traditions of oral narratives. Terblanche (2011) studied markers of narrativity, including the perfect aspect, in East African English, but unfortunately did not distinguish the varieties by country, comparing neither Kenyan and Tanzanian English, nor these two to British English. She did note however that the East-African oral narratives are more interactive and involve the listener more than written fiction. Taken together with the finding that the present tense is used to maintain interpersonal relations, and the past tense as a means to recount experiences, Figure 5 would seem to indicate that the Tanzanian oral storytelling tradition is even more interactive than the Kenyan tradition. More research would of course be needed to verify this hypothesis.

6. Qualitative results

For the final analysis reported in this chapter, randomly thinned samples of the concordances obtained in the preliminary analysis detailed in the previous chapter were coded for various factors including surrounding time adverbials and tenses and the meaning of the past perfect form. First, the coding criteria are explained in detail, followed by presenting the results obtained from comparing the coded data sets. The results are then compared with a previous study of the same type. It will then be possible to hypothesize on the possible causes for any variation of the past perfect in East African English.

6.1 Coding criteria and variables

The coding of the data is constructed in such a way that it will reveal variation in function rather than form, revealing potentially different ways in which East African speakers use the past perfect, if any. The model is taken from Sharma 2001, where the past perfect in the press registers of Indian English and British English was compared using the Kolhapur corpus of Indian English.

The Kenyan and British English concordances were first sorted alphabetically by the whole concordance line, in effect creating a random order for the concordance lines, and then exported from Wordsmith 6 to two separate Excel files, one for each variety, treating Kenyan and Tanzanian English as a monolithic variety. The first 200 lines were chosen for coding in both files, and subsequently expanded to include a few more each due to duplicates or non past perfects that had avoided deletion during the preliminary analysis, so that 200 usable concordance lines were obtained from each file. These included some cases where Wordsmith 6 had lost the link between search word and its location in the corpus file, as well as some contracted instances of *would*. This method of randomization allowed for easy expansion of the data set as needed.

On each concordance line, the semantic function of the past perfect was first coded. While the judgment of the semantics of each instance of the past perfect is admittedly rather subjective, it is necessary because we are looking for variation in meaning rather than form. It is possible for example to be misguided by knowing from which corpus the concordance line originates, perhaps making it more likely to ascribe standard meanings to the British English past perfects – or whatever the researcher's precognition about the situation is. Mixing the data together would also not help, because the text itself very easily reveals the country of origin, if not in mannerism, then in cultural content. Spoken language is difficult to analyze and it is difficult to know for sure what the speaker wanted to say; it often seemed safest to just assume past perfect meaning. Carefully formatted registers would in that respect produce more secure results. In relying on the researcher's language skills, this part of the analysis unfortunately nears armchair linguistics – here only teamwork could aid the research. Sharma (2001, 371) had another person redo a subset of the data to see how accurate the coding was, finding 94% accuracy. In practice, each line was read as a narrative, using the natural way a human reconstructs events in the mind based on a story. E,R and S points were construed, and the apparent lack of R point was a strong indicator of nonstandard meaning. Sometimes more than one possible narrative formed based on the concordance line, and in such cases the past perfect had an ambiguous meaning.

The possible meanings for the past perfect found by Sharma (2001, 352) are the following, yet it is important to let the data suggest the categories as the analysis proceeds:

    a. past perfect – standard meaning
    b. present perfect
    c. preterite – any past meaning

Sharma grouped the non-standard uses together because the difference in meaning is not always clear. In the current study the meanings were coded separately so that they could be analysed separately if needed.

It could be hypothesized that the lack of contextual cues implies strong and uniform semantics for the past perfect, because in such a case no disambiguation is necessary for the meaning to be transmitted clearly, the tense being sufficient by itself. Four contextual cues were coded in order to examine how much contextual marking is used in each variety:

    a. disambiguating adverbials
    b. preceding tense
    c. following tense
    d. reported speech verb

The disambiguating adverbials included only such adverbs or adverbial clauses that clearly indicate a past time before a reference point, such as *earlier*, *afterwards*, *already* and *later*. The preceding and following tenses were marked linearly regardless of the clause structure, as was done by Sharma (2001, 353), but include each tense as a separate category, whereas sharma grouped the present perfect and simple present tenses together. Reported speech verbs include such verbs as *said* and *noted* in the same sentence as the past perfect. Sharma (2001, 354) notes that sometimes the reporting speech verb is the only thing that causes the use of the past perfect.

6.2 Qualitative results

The semantical analysis of the past perfect produced promising results. 198 concordance lines from the ICE-GB were clear cut cases of standard past-before-past meaning, with 24 cases where a simple past or present perfect would have been possible because of normal narrative structure where the past perfect simply continues the narrative in temporal order. The remaining 2 were counterfactuals.
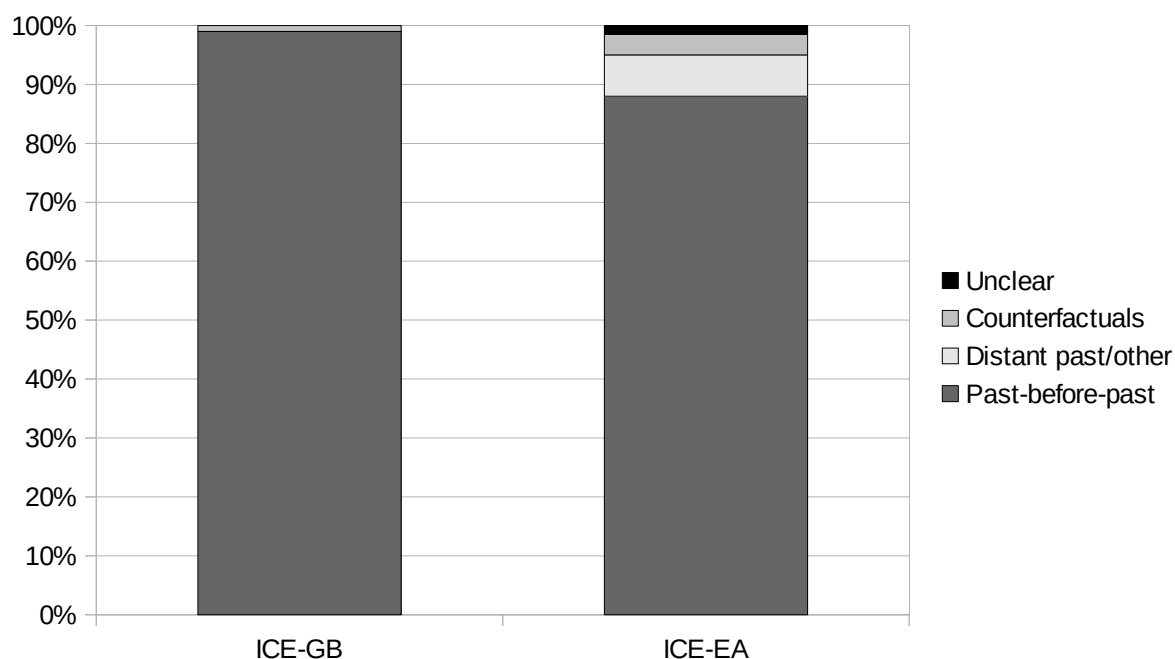
Figure 1. Results of semantic analysis: Meanings of past perfects.

In the ICE-EA, 176 of the 200 past perfects were cases where the past perfect was deemed to have standard past-before-past meaning, and of these, 12 could have been expressed with another tense. There were 7 counterfactuals and 3 unclear cases in the ICE-EA. 14 cases could not sustain a past-before-past reading, examples are provided below in [1-3]. In some cases a simple past would have sufficed [4]. Based on this data set, there seems to be more variation in the meaning of the past perfect in East African English than in British English, but the difference is marginal.

> [1] President [VVN] noted that ... as a result of that peace Western province **had [VVN] developed** tremendously since independence (br-newsK.txt)
> [2] VVN] left for the collection was extremely too short. However I did as you **had [VVN] instructed** that is to give to the members that had [VVN] paid. On (socletK.txt)
> [3] ... diseases in tropical countries, is making a tremendous upsurge. The disease **had [VVN] defied** global effort to eradicate it. Now, Tanzania health exp (ppnats-T.txt)
> [4] ambura to Waceera. Njoroge had [VVN] stormed out of the house when Waceera **had [VVN] started** talking, supposedly to herself. Therefore he was not there... (creative-2K.txt)

According to the corpus data, adverbial disambiguation was more common in East-African English than in British English, with 28 disambiguating adverbials in the ICE-EA data and 14 in the ICE-GB data. This would seem to indicate that the past perfect is not as efficient in carrying meaning in East African English as it is in British English, although adverbials may be favored for

other reasons than reinforcing the temporal relationship of events, e.g. for stylistic reasons or just out of habit.

The tense context of the past perfect was also coded for in the analysis. Linearly searching for preceding and following verbs was not very revealing because of the complexity of the sentences in real speech due to stuttering, or in writing due to subordination and coordination. In spoken language the preceding verb was sometimes the same as the one found by the concordancer due to the speaker repeating a chunk of text as in [5]. In such cases the repeated verb was ignored and the verb that preceded the past perfect in the intended form of the sentence was coded instead. Similarly the disambiguating adverbial and reported speech verb typically needed to be searched from quite a distance before the past perfect.

> [5] ... [VVN] Said that he had no more All he [VVN] **needed** He **had** now [VVN] **got** whatever he [VVN] **needed** So he had no more demand... (sp-lectK.txt)
> [6] It **was** a normal four acre peasant shamba from which we **had [VVN] obtained** all our livelihood when growing up. There **were** four huts... (creative-1K.txt)
> [7] He **collapsed** only five metres from his house when coming from a begging mission where he **had** [VVN] **obtained** just a kilo of flour. He **had not** [VVN] **eaten**...(rep-feat-T.txt)
> [8] The move **had** [VVN] **failed** because the Italians **had** [VVN] **seemed** to be hostile They [VVN] **started** to move their troops (sp-lectK.txt)

The most common tense context was simple past surrounding the past perfect on both sides as in example [6], with 40 % of the instances in both corpora. A simple past before the past perfect token followed by past perfect [7], or vice versa, a past perfect before the token followed by simple past [8] were both as common in both corpora, with close to 10 % of instances for each tense context in both corpora. The findings in this regard are strikingly equal and uneventful.

There is however a difference in the number of reported speech verbs that were found in the corpora. In the ICE-GB 15 % of the tokens were accompanied by a reported speech verb, as opposed to 24 % in the ICE-EA corpus. 14 % of the cases in ICE-GB were accompanied by a disambiguating adverbial as opposed to 23 % in ICE-EA.

6.3 Examples from ICE-EA with non-standard meanings

The following story from column-T.txt in ICE-EA is a fairly convincing example of a past perfect

form with simple past meaning. Simple past tense would be the most natural in [9a]since the

sentence is the beginning of a story in a rant about poor driving conditions in Dar es Salaam, and

there is no time point to which to relate the adverbial 'the previous day' either. [9b] clearly needs the

past perfect since it is something that has happened before Siraj left Arusha. After that, the events

are recounted in the order they happened.

> [9] But, my brother, you should see these people drive in downtown Dar es Salaam, and try to
> get a space to park their cars. They sweat and curse, but that parking space is difficult to get.
> There is this friend of mine, Siraj, who
> [9a] **had** the previous day **driven** from Arusha and
> [9b] **had forgotten** to add more fuel to his <-/guzling> Intercooler Pajero. He **left** his
> residence in Masaki and **headed** for town. That **was** about 7.15 a.m.

In example [10] from the written subcorpus of ICE-EA (ppnats-K.txt) there seems to be no

need for the past perfect. This sentence is from a text describing the challenges of integrating

mental health services to primary health care. A simple past reading seems the most likely in this

case. It is not the assumption made by the observers that should have happened before current or

recent developments, but the inclusion into health care of mental health services. Perhaps the

implicated past-before-past meaning of the noun phrase 'inclusion...' has somehow bled into the

verb phrase.

> [10] However, when it comes to mental health, things seem to have been slow and most
> observers **had assumed** that its inclusion into mainstream of the primary health delivery
> system may have been merely for prestige.

Example [11] makes one think that excessive past-before-past reference may even be a feature

of East-African English. Could the intention and meaning behing the use of the past perfect in fact

be to situate an event before a reference point, situated before or possibly sometimes at the time of

utterance? Perhaps they do not trust the capability of the simple past to break the expected temporal

order of narration, and the past perfect has taken its place. Sharma (2001, 365) found a similar shift

in meaning in Indian English. Sharma's explanation for what is happening is that the past perfect distances a completed (perfective) past event from the narrative focus.

> [11] The babaangu and mamaangus here were rumoured to have been planning a coup de' etat(sic) over a certain regime that makes golden-coloured waters near Ilala Mchikichini, because they feel their mbege brew has come of age and should be made a national drink. Of course it **had been** one of those many rumours that TZ boys and girls love to circulate - for pleasure.

Here the past perfect is to first establish a past-before-past reference and then to continue the narrative in that time sphere, as evidenced by the causal relationship between [12a] and [12b]. As was noted in chapter 3.5 this usage is in line with standard usage, even if the past perfect could be replaced by the simple past in [12b] without making the order of events unclear.

> [12] Now the explosion occurred particularly in the coffee growing areas uh particularly in the northern part of Angola Coffee prices
> [12a] **had been** high in nineteen fifties And this
> [12b] **had led** to coffee speculators coming into Angola uh particularly for <O/> And because of the high prices that uh coffee was catching the Portuguese were anxious to settle in northern part of Angola so that they could share in the piece of the cake (sp-lectK.txt)

[13] appears to be a case of events being recounted in an unspecified order, the only pair of tenses with a causal and temporal order being [5c] and [5d]. As [5c] is in the past perfect, but still apparently lacking a past-before-past meaning, its function seems to be to signal the beginning of a match between linear order of tenses and temporal order of events.

> [13] Otherwise we got those knee-jerk reactions: Some members of the Opposition and clergy
> [13a] were fanning the clashes, the Opposition
> [13b] had a penchant for violence and oathing, some Opposition figures
> [13c] **had called** for civil disobedience, and
> [13d] vowed to create a Somalia- type situation, churchmen
> [13e] favoured the Opposition anyway.

In [14] from ldsoc-T.txt, the present perfect would be expected, the reference point forced by the past perfect being superfluous. In [15] from parlia-K.txt likewise the past perfect seems unnecessary for the same reason, but simple past would perhaps be expected.

[14] This being so, it seems entirely reasonable to set a "standard of mastery" of the previous phase of training which is necessary before the student can move on to the next phase. Only those students who **had mastered** such "minimum essentials" would then be promoted to the next grade or be awarded a certificate, diploma or a degree.
[15] At the moment, there is a lot of robbery on those who have taken out insurance by the insurance companies. Many insurance companies, during the time of paying compensation, under-value whatever **had been insured**.

[16] is the very first sentence in a written report about problems with malaria control from

ppnats-T.txt. [17], a report about beef shortage, comes from pptech-K.txt. In both examples, no

intervening reference point is evident, the most likely reading being once again simple anteriority to

the narrative focus.

[16] Malaria, one of the killer diseases in tropical countries, is making a tremendous upsurge. The disease **had defied** global effort to eradicate it. Now, Tanzania health experts are taking a number of measures to combat both mosquitoes and malaria...
[17] They revolutionary changes proposed by the ministry will go a long way to ensure that farmers get the services whenever they desired. But it is lamentable that from past experience, dips handed to the village committees **had not performed** well.

Most if not all of the previous examples, at least [9]-[12] and [15]-[17], can satisfactorily be

explained by the past perfect having an additional meaning of anteriority from the current time

point of the narrative, without a necessary reference point between the events. It seems that the

development of the East-African English past perfect has mirrored the development of the Indian

English past perfect as reported by Sharma (2001, 365) to some degree. Next, the qualitative

findings from this study and that of Sharma's will be further compared.

6.4 Comparison to earlier studies

Having used the same methods of analysis as Sharma (2001), it is now possible to compare

the results with the said study. Here it must be noted that Sharma studied the press register of the

Kolhapur corpus of Indian English, whereas the current study focuses on a broader spectrum of

English in East Africa. In British English, Sharma (2001, 356) found 5 % of non-standard

meanings, whereas in the current study, 0 % of non-standard meanings were found for British

English (see Figure 1. below). It would appear that the non-standard meaning of the past perfect is a feature of the press register, yet this can also be an illusion caused by sample size and subjective decisions in the interpretation of the past perfects.
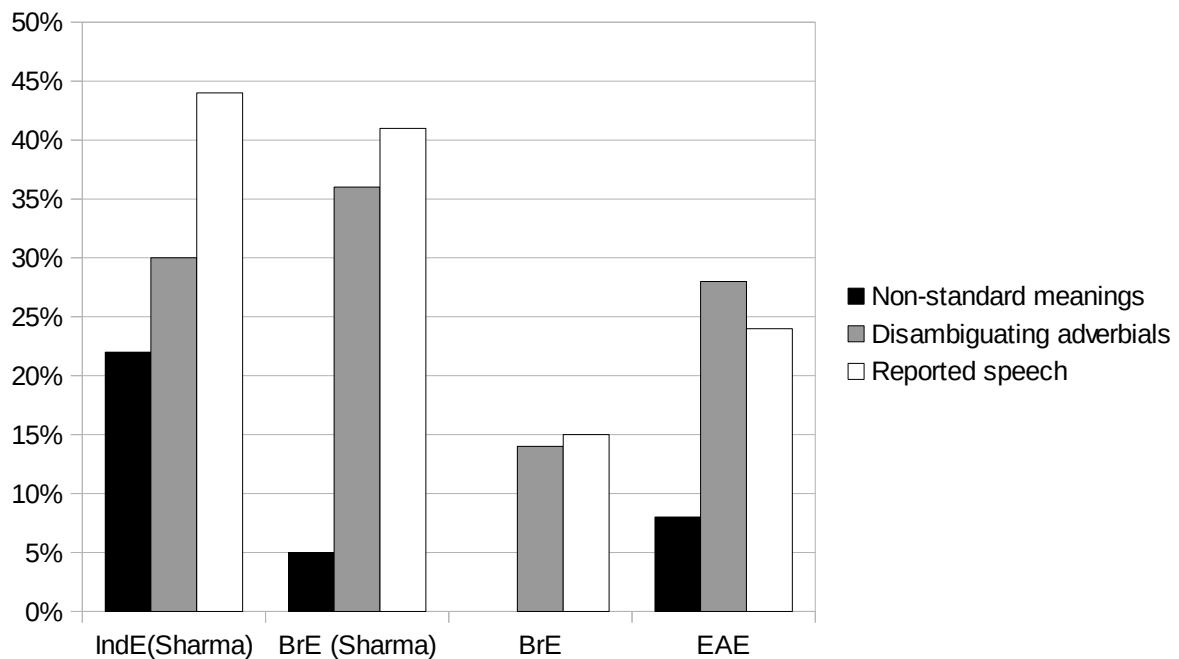

Figure 1. Comparison to Sharma's (2001) results.

While Sharma found Indian English to have a different usage of the past perfect, the same can only tentatively be said about East African English, because the percentage of non-standard meanings for the past perfect in East African English is only 8 %, which is very close to Sharma's figure for British English. Yet the few cases that were found were quite convincing examples of non-standard meaning. Interestingly, most of the non-standard cases come from written or learned contexts, such as lectures, in the ICE-EA, to which the press register might also be included, but more study would be required to determine whether the non-standard use is actually a feature of only written and learned registers.

Disambiguating adverbials were more frequently used with the past perfect in the press registers of British English and Indian English, 36 % and 30 % of cases respectively (Sharma 2001, 358), than in British English and East-African English in general. The difference between the latter

is more striking, however, with East African English having a percentage of 28 %, which is closer to Sharma's figures than the 14 % of British English.

The difference in the figures for reported speech should come as no surprise due to the register specific nature of Sharma's study: 41 % in British English and 44 % in Indian English press registers (2001, 359) as opposed to 15 % in the British English and 24 % in the East African English corpora. The uneventfulness of the other figures obtained from the current study may also be explained away with the fact that all registers were considered here, it may well be that concisely situating events in their temporal order is more important in newspaper text than in general.

7 Conclusion

The findings are next summarized. Some technological possibilities for future methodologies are considered and some suggestions are given for further study.

7.1 Summary of findings

The research questions are repeated below for convenience.

1. Is there a change in the way the past perfect is used in East African English as compared to British English?
2. Similarly, is there a change in the meaning of the past perfect?
3. If so, what is the cause for the change?

To answer the first research question, there were clear differences between the two varieties in the usage of the past perfect. Contracted forms are used more in British English than in East-African English. Surprisingly the past perfect is equally common in all written language in both varieties, with the aforementioned difference in contractions. In speech, the past perfect is perhaps surprisingly even more common in Kenyan English than in British English.

As for the second research question, the variation is more subtle in meaning, but it certainly exists. It was seen that most of the examples with non-standard meaning in ICE-EA can be explained with anteriority to the current time reference of the narrative, without an intervening reference point. The difference in speech of the past perfect in British English versus East-African English is thus mostly a matter of frequency and fluency, and in written texts and learned registers, such as lectures, there is a change in meaning too.

Thirdly, the cause for these changes is difficult to locate. The reason for the change in meaning is probably not embedded in native English, as was claimed for the press register by Sharma (2001,), since the variation was found not to exist in the British English corpus. In usage, sociological reasons, such as wanting to appear learned and civilized, seem likely to be the cause for favoring non-contracted past perfects. As for the change in meaning, first language interference

from the tense-aspect-modality systems of local Bantu languages, especially Swahili, may

contribute. A potential cause for change, as was pointed out in chapter 2.3, is the fact that Swahili

does not have a past perfect tense, but in stead it has a remote past tense, and speakers may use the

past perfect of English as a direct translation equivalent. The lack of the past perfect in Swahili may

make the distinction between past-before-past and simple anteriority difficult to grasp; Yet the fact

that most of the non-standard examples came from written and learned registers, where the

language was otherwise fluent, would seem to not favor this interpretation.

7.2 Remarks on digital data handling

Outside of corpora, the following single sentence example from the header of a Tanzanian

website would seem to show an example of the distant past usage of the past perfect form,

especially since there is no context given from where to infer any kind of intervening reference

point. It seems that the beginning of the current state of affairs, Nyanduga's being advocate of the

High Court, is considered the intervening reference point that merits the use of the past perfect, or

the past perfect might be used with simpe or remote past meaning, or this snippet may have simply

been copied from some other, longer stretch of text as is and left unedited. As East African websites

become available in greater numbers, the prospect of forming a corpus based on online texts

becomes more and more tempting.

> Bahame Tom Nyanduga* is Advocate of the High Court of Tanzania, and **had been** the
> President of the East Africa Law Society between October 2004 - October 2006.

Automation was a great help in the part of this study that focused on frequencies and usage,

but there is precious little that can be done to help with the analysis of meaning. The ICE corpora

readily lend themselves to automated quantitative analysis, because their form is standardised,

removing the need for manual intervention in the early stages of making the data compatible with

the programmatic data extraction. This study could fairly easily be replicated on all the ICE

corpora, with the proviso that they could be obtained in similar formats - as was noted in chapter 4,

the two corpora were in fact in slightly different file hierarchies, which somewhat complicated matters.

Language parsing technologies were considered for helping in the qualitative analysis of the study, but found to be very complicated and to require programming skills to implement in even a basic way. While language parsing modules are freely available as open source software, the corpus hits would have to be fed to such a module, in effect a program would have to be created. So far, such technologies can fairly reliably create tree diagrams for sentences and tag sentences with part-of-speech tags, but such fine-grained semantic analysis as finding the simple past meaning among past perfects was deemed too complex. Tree diagrams reveal the grammatical structure of a sentence, but as the structure of a past perfect and a past perfect with simple past meaning are identical, it would be of no use. In the finding of past perfects in a corpus, a language parser could have been of use, although it would still have required more set-up time than using regular expressions and elbow grease, as was done in the current study. It appears likely that an actual artificial intelligence beoynd the scope of our current abilities would be required in order to facilitate the qualitative semantic analysis of this kind of a study.

7.3 Discussion

As there were no nonstandard meanings for the past perfect and less disambiguating adverbials in the British corpus, the past perfect appears to have a uniform meaning in British English. It is possible that East Africans do not trust the past perfect to convey the intended meaning, and that they therefore use more disambiguating adverbials. It was noted in chapter 2.3 that the native Bantu languages are poorer in aspectual verb markings, instead relying on lexical means to express aspect where tenses fail to convey the wanted meaning. This might then add to the character of their English via first language interference, resulting in English that is richer in

adverbials. Indeed, there were 7 % more disambiguating adverbials in the admittedly small sample, lending moderate support for this hypothesis.

Either East-Africans generally report more speech, or the past perfect is more tightly connected to reported speech in East-African English than in British English. A search of the corpora with common reported speech verbs would be needed to clear this matter. Further study would be needed in Indian English in general and perhaps the press register of East-African English in particular for a better comparison between the varieties. Currently it is not possible to compare the results in any meaningful way, but at least it was possible to compare British English and East African English.

Perhaps it is surprising that there is so little change in the meaning of the past perfect of East-African English, as it could be expected when a new language is appropriated in a country with a rich linguistic background, such as Kenya or Tanzania. Milroy (1992, 10) notes that usually only change is investigated and "maintenance" is ignored, which can lead to one-sidedness of theories of change. The question "Why do some features change while others do not?" is an interesting one in historical linguistics as well as contact linguistics. Why has the meaning of the past perfect not changed even though the frequency of use has changed? In discourse, shared information about the events may make very complex tense structures unnecessary, which may cause them to be used less in face-to-face communication. Perhaps in natural language acquisition, this can contribute to difficulties in picking up the different nuances of meaning and substituting with first language interference.

As noted in chapter 2, English is a language of higher education and business that has been somewhat forcibly made the official language through language policies in these countries. It appears that since the language is so strongly tied with the school system which has taken as its model the Western school and its language, the users of the language can easily be made to follow native standards in their use of English. The prestige status the language gives to its speakers may

act as a strong deterrent to allowing first language influence, perhaps contributing to the small amount of variation in the past perfect of East African English.

The current study focused on just one tense, but a systemic study of the whole tense-aspect-modality system would certainly be more revealing. Are other tenses used with similar frequencies compared to British English? Are contracted forms equally rare with all tenses and thus a feature of East-African English in general? A diachronic corpus would reveal the history of the change. Has East-African English developed towards British English during the time of its use in Kenya and Tanzanya as the nations have learned the language? Was the root of the change already present in the first seeds of English sown in East Africa, or was the change something that happened later on?

Works Cited

Primary sources

International Corpus of English; East African Subcorpus (ICE-EA)

International Corpus of English; Great Britain Subcorpus (ICE-GB)

Secondary sources

Aceto, Michael. 2004. "Eastern Caribbean English-derived language varieties: morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 439-453. Berlin, Mouton de Gruyter.

Ball, C.N. 1994. "Automated Text Analysis: Cautionary Tales." *Literary and Linguistic Computing* 9: 295-302.

Bayley, Robert and Otto Santa Ana. 2004. "Chicano English: Morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 374-390. Berlin, Mouton de Gruyter.

Biber, et al. 1998. *Corpus Linguistics*. Cambridge: Cambridge University Press.

Billings, Sabrina. 2009. "Speaking beauties: Linguistic posturing, language inequality, and the construction of a Tanzanian beauty queen." *Language in Society* 38: 581-606.

Blommaert, Jan. 2005. "Situating language rights: English and Swahili in Tanzania revisited" *Journal of Sociolinguistics*, 9/3: 390-417.

Comrie, Bernard. 1976. *Aspect*. Cambridge, Cambridge University Press.

Comrie, Bernard. 1985. *Tense*. Cambridge, Cambridge University Press.

Declerck, Renaat. 1992. *A Comprehensive Descriptive Grammar of English*. Tokyo, Kaitakusha.

Filppula, Markku. 2004. "Irish English: morphology and syntax" *A Handbook of Varieties of English*: 73-101. Berlin: Mouton de Gruyter.

Francis, Nelson W. 1982. "Problems of assembling and computerizing large corpora" in *Computer corpora in English language research* ed. S. Johansson, 7-24. Bergen, Norwegian Computing Centre for the Humanities.

Huber, Magnus and Kari Dako. 2004. "Ghanaian English: morphology and syntax" in A *Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 855-865. Berlin: Mouton de Gruyter.

Häcker, Martina. 1999. *Adverbial Clauses in Scots: A Semantic-Syntactic Study*. Berlin, Mouton de Gruyter.

Kautzsch, Alexander. 2004. "Earlier African American English: morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 343-347. Berlin: Mouton de Gruyter.

Leech, Geoffrey and Jan Svartvik. 2002. *A Communicative Grammar of English*. Essex, Pearson Education Limited.

Lindfors, Anna-Lena. 2003. "Tense, Aspect and Mood in Swahili" [Accessed 2014-04-29 http://www2.lingfil.uu.se/ling/semfiler/Swa_TAM.pdf]

Lindquist, Hans. 2009. *Corpus Linguistics and the Description of English*. Edinburgh: Edinburgh University Press.

Meshtrie, Rajend. 2004. "Black South African English: morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 962-973. Berlin: Mouton de Gruyter.

Michieka, Martha Moraa. 2005. "English in Kenya: a sociolinguistic profile" *World Englishes,* 24/2: 173-186.

Miller, Jim. 2004. "Scottish English: morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 47-72. Berlin: Mouton de Gruyter.

Milroy, James. 1992. *Linguistic Variation and Change*. Oxford, Blackwell Publishing.

Mugler, France & Tent, Jan. 2004. "Fiji English: morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 770-788. Berlin: Mouton de Gruyter.

Palmer, F.R. 1980. *The English Verb*. London, Longman Group Ltd.

Partington, Alan. 1998. *Patterns and Meanings*. Amsterdam: John Benjamins.

Quirk, Randolph, S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A comprehensive grammar of the English language.* London: Longman.

Reichenbach, Hans. 1947. *Elements of Symbolic Logic*. New York, MacMillan.

Rieger, Dorothee. 2011. "Swahili as a Tense Prominent Language. Proposal for a Systematic Grammar of Tense, Aspect and Mood in Swahili" in *Swahili Forum 18,* 114 – 134. [Accessed 2014-04-29 through http://www.qucosa.de/fileadmin/data/qucosa/documents/9058/18_08_Rieger.pdf]

Ross, H., Oetting, J.B., Stapleton, B. 2004. "Preterite had + V-ed: A developmental narrative structure of African American English" *American Speech,* 79/2: 167-193.

Salton, Gerard. 1973. "Recent Studies in Automatic Text Analysis and Document Retrieval" *Journal of Association for Computing Machinery* 20: 258-278.

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

Schmied, Josef. 2004. "East African English (Kenya, Uganda, Tanzania): morphology and syntax" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 929-947. Berlin: Mouton de Gruyter.

Schmied, Josef. 2006. "East African Englishes" in *The Handbook of World Englishes*, ed. Braj B. Kachru, Yamuna Kachru, Cecil L. Nelson, 188-202. Oxford: Blackwell Publishing Ltd.

Schneider, Edgar W. 2004. "Synopsis: Morphological and syntactic variation in the Americas and the Caribbean" in *A Handbook of Varieties of English*, ed. Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, Clive Upton, 1104-1115. Berlin: Mouton de Gruyter.

Sharma, Devyani. 2001. "The pluperfect in native and non-native English: A comparative corpus study" *Language Variation and Change* 13: 343-373.

Skandera, Paul 2002. "A categorization of African Englishes" in *Perspectives on English as a World Language* ed. Allerton, D. J., Paul Skandera, Cornelia Tschichold, 93-103. Basel: Schwabe.

Svartvik, Jan. 1992. "Corpus linguistics comes of age" in *Directions in Corpus Linguistics. Proceeding of Noble Symposium 82 Stockholm 4- 8 August 1991*, ed. Jan Svartvik. New York, Mouton De Gruyter.

Terblanche, L. 2011. *Modelling narrativity in East African English*. MA dissertation. Vanderbijlpark: North-West University, Vaal Triangle Campus.

Appendix 1: Full regular expression search term

\b(had|hadn't|'d)(?:\W+\w+){0,6}?\W+(\w*ed\b|had|abode|abided|abidden|abought|alit|alighted|
arisen|awoken|backbitten|backfit|backlit|backslid|backslidden|been|born|borne|beaten|become|
bedight|bedighted|befallen|begotten|begun|begirt|behight|behoten|beheld|belaid|belayed|bent|bereft|
bereaved|besought|beseeched|beset|bespoken|bestrewed|bestrewn|bestridden|bet|betaken|bethought|
bewept|bid|bidden|bided|bound|bitten|blawn|bled|blended|blent|blessed|blest|blown|bottle-fed|
bowstrung|broken|breast-fed|bred|brought|broadcast|broadcasted|browbeat|browbeaten|built|burnt|
burned|burst|bust|bought|cast|caught|cheerled|chid|chidden|chosen|clapped|clapt|clear-cut|cleft|
cleaved|cloven|cleeked|cleped|ycleped|yclept|clung|clad|clothed|colorbred|colorcast|come|cost|cost-
cut|counterdrawn|counterlit|cowritten|crash-dived|crept|creep-fed|crib-bitten|cross-bitten|cross-
strung|crossbred|crosscut|crosslit|crowed|cut|dared|daydreamt|daydreamed|dealt|deep-frozen|dug|
dight|dighted|dispread|disproven|disproved|dived|done|dogfought|dought|dowed|drawn|dreamt|
dreamed|drunk|drip-fed|driven|dwelt|eaten|end-run|engirt|engraved|engraven|enwound|fallen|
farebeaten|fast-cut|fed|felt|fought|found|fit|fitted|flash-frozen|fled|flung|floodlit|flown|flyblown|
forborne|forborn|forbidden|force-fed|fordone|forecast|forecasted|forefelt|foregone|foreknown|
forerun|foreseen|forshowed|foreshown|forespoken|foretold|forgotten|forgiven|forlorn|forsaken|
forsworn|fraught|fraughted|free-fallen|frozen|frostbitten|gainsaid|gaslit|gaslighted|gelded|gelt|got|
gotten|ghostwritten|gilt|gilded|gan|girded|girt|given|gnawed|gnawn|gone|been|graven|graved|ground|
grown|hagridden|halterbroken|hamstrung|hamstringed|hand-fed|hand-ridden|handsewn|handsewed|
handwritten|hung|hanged|heard|heaved|hove|hewn|hidden|hoten|hit|hoised|hoist|hoist|hoisted|held|
house-sat|housebroken|hurt|inbred|indwelt|inheld|inlaid|input|inputted|inset|interblended|interblent|
interbred|intercut|intergraved|intergraven|interlaid|interset|interwoven|interwound|inwoven|
inweaved|inwound|jerry-built|kept|kent|kenned|knelt|kneeled|knit|knitted|known|laden|laded|
landslid|landslided|laid|led|leant|leaned|leapt|leaped|learnt|learned|left|lent|lept|let|lain|lit|linebred|lip-
read|lost|made|meant|met|molten|melted|misbecome|miscast|mischosen|miscut|misdealt|misdone|
misfallen|misfed|misgiven|misheard|mishit|misknown|mislaid|misled|mislearnt|mislearned|misread|
missaid|missent|misset|misspoken|misspelt|misspelled|misspent|missworn|mistaken|mistaught|
mistold|misthought|misunderstood|misworn|miswed|miswedded|miswritten|moonlit|mown|naysaid|
nose-dived|nose-dove|offset|outbid|outbred|outdone|outdrawn|outdrunk|outdriven|outfought|
outflown|outgrown|outlaid|outleapt|outleaped|output|outputted|outridden|outrun|outseen|outsold|
outshone|outshot|outsung|outsat|outslept|outsmelt|outsmelled|outspoken|outsped|outspent|outspun|
outsprung|outstood|outsworn|outswum|outtold|outthought|outthrown|outworn|outwound|outworked|
outwrought|outwritten|overborne|overbid|overblown|overbred|overbuilt|overbought|overcast|
overcome|overcut|overdone|overdrawn|overdrunk|overdriven|overeaten|overfed|overflown|
overgrown|overhung|overheard|overlaid|overleapt|overleaped|overlearnt|overlearned|overlain|
overpast|overpassed|overpaid|overridden|overrun|overseen|oversold|overset|oversewn|oversewed|
overshot|overslept|oversown|oversowed|overspoken|overspent|overspilt|overspilled|overspun|
overspread|oversprung|overstood|overstrewed|overstrewn|overstridden|overstruck|overstrung|
overtaken|overthought|overthrown|overworn|overwound|overwithheld|overwritten|partaken|paid|
pent|penned|pinch-hit|pled|pleaded|potshot|practice-taught|prebound|prebuilt|precut|predone|
premade|prepaid|presold|preset|preshrunk|presplit|price-cut|proofread|proven|proved|put|quick-
frozen|quit|rapped|rapt|re-proven|re-proved|read|reft|reaved|reawoken|reawaken|rebid|rebound|
rebroadcast|rebuilt|recast|recut|redd|redded|redealt|redone|redrawn|rove|reeved|refit|reground|
regrown|rehung|reheard|reknit|reknitted|relaid|relearnt|relearned|relit|relighted|remade|rent|repaid|
reread|rerun|resold|resent|reset|resewn|resewed|reshot|resat|resown|resowed|respelled|respelt|
restrung|retaken|retaught|retorn|retold|rethought|retrodden|retrofit|retrofitted|rewoken|rewaked|
reworn|rewoven|reweaved|rewed|rewedded|rewet|rewetted|rewon|rewound|rewritten|rid|ridded|

ridden|rung|risen|riven|rived|rough-hewn|roughcast|run|sand-cast|sawn|sawed|said|seen|sought|self-fed|self-sown|self-sowed|sold|sent|set|sewn|sewed|shaken|shaven|shaved|shorn|sheared|shed|shent|shewn|shewed|shone|shat|shit|shod|shot|shortcut|shown|shrunk|shriven|shut|sidewound|sight-read|sightseen|sung|sunk|sat|skywritten|slain|slept|slid|slidden|slung|slunk|slit|smelt|smelled|smitten|snapshot|sneaked|snuck|soothsaid|sown|spoken|sped|speeded|spelt|spelled|spellbound|spent|spilt|spilled|spun|spat|spit|split|spoilt|spoiled|spoonfed|spread|sprung|stall-fed|stood|stove|staved|stoved|stolen|stuck|stung|stunk|straphung|strewn|strewed|stridden|struck|stricken|strung|stript|stripped|striven|sublet|sunburned|sunburnt|sworn|sweat|sweated|swept|sweeped|swollen|swum|swung|swonken|switch-hit|taken|taught|team-taught|torn|telecast|telecasted|told|test-driven|test-flown|thought|thriven|thrived|thrown|thrust|thunderstruck|thunderstricken|tint|tined|tossed|tost|trodden|troubleshot|typecast|typeset|typewritten|unborn|unborne|unbent|unbound|unbuilt|unclad|unclothed|underbid|underbidden|underbought|undercut|underdolven|underdelved|underdone|underfed|undergirt|undergirded|undergone|underlaid|underlet|underlain|underpaid|underrun|undersold|undershot|underspent|understood|undertaken|underthrust|underwritten|undone|undrawn|undrawn|unfrozen|unhung|unhidden|unheld|unknit|unknitted|unladen|unladed|unlaid|unlearnt|unlearned|unmade|unrove|unreeved|unsaid|unsold|unsewn|unsewed|unslung|unspoken|unspun|unstuck|unstrung|unsworn|untaught|unthought|untrodden|unwoven|unwound|unwritten|upbuilt|upheld|uppercut|uprisen|upset|upsprung|upswept|uptorn|vext|vexed|woken|waylaid|worn|woven|wed|wedded|wept|wended|went|wet|wetted|whipsawn|whipsawed|won|wound|winterfed|wiredrawn|wist|withdrawn|withheld|withstood|wont|wonted|worked|wrought|wrapped|wrapt|wreaked|wroken|wrung|written|zinced|zincked)\b