



UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

***Ke* in Utterances:**

Uses & functions of the Xhosa Discourse Marker

ke

Abdul-Malik Sibabalwe Oscar Masinyana

Submitted for an MA in Linguistics (2013)
School of African and Gender Studies, Anthropology and Linguistics
Supervised by Assoc. Prof. Ana Deumert

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DECLARATION

I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.

I have used the Harvard convention for citation and referencing. Each contribution to, and quotation in, this dissertation from the work(s) of other people has been attributed, and has been cited and referenced.

This dissertation is my own work and has never been submitted for examination elsewhere.

I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature _____

Date _____

ACKNOWLEDGMENTS

I would like to thank Associate Prof. Ana Deumert for the years of contribution to this work. All that is well with it is because of her invaluable input and what is not is due to my own negligence. I cannot express my gratitude enough. Thank you Ana.

Prof. Rajend Mesthrie for the academic and financial support (SANPAD and NRF) and encouragement; he was the first to see potential in me and continued to support me like one who never lost faith long after I did.

Prof. Mbulu Madiba at the Multilingual Development Unit at the Centre for Higher Education Development (CHED) for providing additional funding for this research.

Prof. Derek Gowlett without whom my thesis would have been a whole lot more ‘folky’, ‘unnecessarily dense’ and many other things besides were it not for his timely and very constructive intervention.

Aditi Humna for being such a generous reader and guide during a very difficult phase of the writing process, as well as Noxolo Tshikila for help with the frequency counts.

Prof. Kay McCormick, Yolandi Klein, the postgraduate research group (UCT linguistics section postgraduate students), Zolani Kupe, Zukile Jama, Dr. Menan du Plessis, Prof. Tessa Dowling, Thabisa Xhalisa, and the colleagues at the Writing Centre – thank you for the discussions, your insights and continuous encouragement.

The staff at UCT library: Mr. Isaac Ntabankulu (Archivist at UCT Library) and Nuroo Ismail Davids (at Knowledge Commons). The former helped me access Xhosa archive material that would have taken me forever to unearth, and even read parts of my work, and the latter for introducing me to Microsoft formatting tools that helped me to organise what was fast turning into an unmanageable mess.

My mother Ntombizandile Aunt’Zaa Griselda Masinyana, my brothers Lungisa and Luvuyo Masinyana, and all my aunts and cousins for the endless encouragement and support. You are the reason I kept going long after I couldn’t.

And my dear friend Tanya Charles – for a lot and for so much more.

Thank you and God bless you all.

In loving memory of

Mongameli Petros Masinyana & Dr. Neville Alexander

– Aluta continua

You are young, my son, and as the years go by, time will change and even reverse many of your present opinions. Refrain therefore awhile from setting yourself up as a judge of the highest matters.

Plato, 'Theatetus'

The speaker is not the biblical Adam, dealing only with virgin and still unnamed objects, giving them names for the first time... In reality... any utterance, in addition to its own theme, always responds (in the broad sense of the word) in one form or another to others' utterances that precede it. The speaker is not Adam, and therefore the subject of his speech itself inevitably becomes the arena where his opinions meet those of his partners (in a conversation or dispute about some everyday event) or other viewpoints, world views, trends, theories, and so forth (in the sphere of cultural communication). World views, trends, viewpoints, and opinions always have verbal expression. All this is others' speech (in personal or impersonal form), and cannot but be reflected in the utterance. The utterance is addressed not only to its object, but also to others' speech about it.

Mikhail Bakhtin, 'The problem of speech genres'

ABSTRACT

Discourse Markers (DM) have been identified in so many languages, utterance contexts, and studied from so many angles and theoretical approaches (Ogoanah, 2011; Jantjies, 2009; Jabeen, et. al, 2011; Dér and Markó, 2010; Verdonik et.al, 2007; Li, 2010; Hernández, 2011; Camiciottoli, 2009), so much that another study hardly seems necessary. Focusing on Xhosa, a linguistic context where hardly any work is being done on DMs, this thesis argues that the Xhosa particle *ke* is a DM that is popular in, but not restricted to, oral utterances and a DM that is present even in 19th century Xhosa utterances. At present, the general agreement between Xhosa grammars and dictionaries is that *ke* is either/and/or a conjunction or conjunctive, an interjective, an adverb, an enclitic, an expletive or a form word with a variety of translation equivalents in English. Using a DM analysis framework provided by Schourup (1999) – which corresponds closely with the pioneering framework by Schiffrin (1987) and also contains elements of Fraser's model (1996, 2009) – this thesis examines these three claims and concludes that *ke* is mainly a DM (over and above being one or more or all of the present classifications) and should be presented as such in future Xhosa dictionaries, grammars and linguistic research.

TABLE OF CONTENTS

DECLARATION	2
ACKNOWLEDGMENTS	3
ABSTRACT	6
TABLE OF CONTENTS	7
INTRODUCTION	9
CHAPTER 1 - TERMS, DEFINITIONS, FUNCTIONS & CHARACTERISTICS: A FRAMEWORK FOR THE STUDY OF DISCOURSE MARKERS	13
1.1 INTRODUCTION	13
1.2 TERMS	13
1.3 DEFINITIONS AND FUNCTIONS	14
1.4 CHARACTERISTICS OF DMS	17
1.4.1 <i>Schiffirin's 1987 Model</i>	17
1.4.2 <i>Schourup's 1999 Model</i>	18
1.4.3 <i>Fraser's Model</i>	25
1.5 CONCLUSION	26
CHAPTER 2 - FORMALISING CURIOSITY: DATA COLLECTION AND METHODOLOGY	28
2.1 INTRODUCTION	28
2.2 RESEARCH QUESTIONS	29
2.3 TYPES OF DATA	30
2.3.1 <i>Spoken corpus 1: Interviews and focus groups</i>	30
2.3.2 <i>Spoken Corpus 2: Oral Narratives</i>	34
2.3.3 <i>Written Corpus 1: Online Material</i>	35
2.3.4 <i>Written Corpus 2: Xhosa Translations of the Bible and Constitution</i>	37
2.4 CONCLUSION	38
CHAPTER 3 - KE IN DICTIONARIES AND GRAMMARS AND KE AS A DISCOURSE MARKER	40
3.1 INTRODUCTION	40
3.2 <i>KE</i> IN DICTIONARIES AND GRAMMARS FROM 1844-2008	40
3.3 <i>KE</i> AS A DM IN CONTEMPORARY TEXTS	46
3.3.1 <i>Connectivity</i>	48
3.3.2 <i>Optionality</i>	49
3.3.3 <i>Non-truth conditionality</i>	50
3.3.4 <i>Weak clause association</i>	51
3.3.5 <i>Initiality</i>	51
3.3.6 <i>Orality</i>	52
3.3.7 <i>Multi-categoriality</i>	53
3.4 CONCLUSION	55
CHAPTER 4 - <i>KE</i> AS A DISCOURSE MARKER IN HISTORICAL TEXTS.....	56
4.2 ANALYSING <i>KE</i> IN MID-19 TH CENTURY PROSE TEXTS	56
4.3 ANALYSING <i>KE</i> IN THE BIBLE	61
4.4 CONCLUSION	63
CONCLUSION AND RECOMMENDATIONS FOR FUTURE RESEARCH	64
BIBLIOGRAPHY	70
GRAMMARS AND DICTIONARIES	70
BIBLE TRANSLATIONS:	71
PUBLIC SOURCES OF DATA.....	71
RESEARCH LITERATURE	72

INTRODUCTION

Perhaps it is the case that, as women and lesbians, we are all too aware that little that 'happens' in a patriarchal culture is 'autonomous'. If language is a matter of convention and males hold the power in a culture, they will determine language as they determine all other conventions capable of sustaining their control. If our position in a patriarchal culture makes us sceptical of facile conclusions that all too often appear to transcend culture, it is because we have that such conclusions are completely enmeshed in the culture of their researchers. So-called native speaker intuitions are particularly suspect when they emanate from the linguist describing a language. In this case, scientific objectivity is a near impossibility.

Susan J. Wolfe and Julia Penelope¹

This thesis argues that *ke* in Xhosa² is a discourse marker (henceforth DM). This is a valid enquiry because (1) most descriptions of the *ke* imbue it with strongly grammatical properties and tend to neglect its DM functions or at least sideline or ignore them, and (2) the two explicit references to *ke* as a DM (Gough, 1987 and De Klerk, 2005) do not sufficiently elaborate on what basis they reached that conclusion. Although *ke* does fulfil grammatical functions in some utterances, these instances tend to be ambiguous and the grammatical functions are always overshadowed by the DM functions of the particle in these utterances. Accompanying this main aim of showing that *ke* is a DM are two other sub-tasks: (1) to show that *ke* is a relatively old DM in Xhosa, one that has been present in the language since it was committed to paper, and (2) that *ke* is also present in written texts of varying levels of formality and also as a DM, but in more common in oral utterances or utterances that have elements of orality in them. All these interlinked aims will ultimately demonstrate the main point

¹ Linguistics: The Impenetrable Paradigm, or Where Were You Two When They Passed Out the Penes? (1997). In *Straight Studies Modified: Lesbian Interventions in the Academy*. Eds. Gabriele Griffin and Sonya Andermahr. London and Washington: Cassell (pp. 228-235)

² The official names of the official South African languages as set out in Article 6 of the Constitution of South Africa (1996) are: Sepedi, Sesotho, Setswana, siSwati, Tshivenda, Xitsonga, Afrikaans, English, isiNdebele, isiXhosa and isiZulu. This thesis, however, uses exonyms for isiXhosa and all other South African and non-South African languages mentioned. The primary language of this work is English and thus all reference to le français or isiXhosa (the languages), for example, and les Français or amaXhosa (the people) – and various other such endonym linguistic forms – are used in their uninflected form in both adjectival and nominal contexts, except in rare, special cases (e.g. quotations).

of this study: that *ke* is a DM and should be clearly indicated as such in Xhosa dictionaries, grammatical texts and future linguistic research.

In the *Greater Xhosa Dictionary*, the *magnus opus* of Xhosa lexicography that was begun in the 1980s and completed in the early 2000s and is currently the best example of its kind, *ke* is classified as a ‘form word’, which is a peculiar grammatical term that serves as an umbrella for all kinds of words including articles, pronouns, conjunctions, adverbs, and some prepositions. As such this term is an insufficient descriptor because it conceals more meaning than it reveals. Also in the same dictionary (this dictionary will be a principal source throughout this study) *ke* is also referred to as a “conjunction” or “conjunctive” as well as an “interjective”. Elsewhere in the dictionary, it is called an “adverb”, an “enclitic” and an “expletive”.

However, redefining *ke* as a DM does not explain its functions and what roles it fulfils in discourse. The DM literature shows such divergent theories and heated debates in linguistics that it is sometimes hard to believe that these two to five letter words are the centre of discussion. Perhaps it is because much more is at stake (language, methods of language study, perceptions, etc) than just words whenever linguists attempt to explain these particles; how we talk about this particular group of words tell us a lot about what philosophy of language we subscribe to, what we believe language to be, and how we think it should be studied.

The field of DMs is so broad that to consider all of its developments within the scope of this work is difficult (if not impossible). In just the last decade alone, there has been research on DMs in Nigerian English (Ogoanah, 2011), in Akan (Amfo, 2010), in Afrikaans (Jantjies, 2009), in British and Pakistani speech (Jabeen, et. al, 2011), in Hungarian (Dér and Markó, 2010), in Slovenian corpora (Verdonik et.al, 2007), in Chinese student’s English writing (Li, 2010), in Spanish medical consultations (Vickers and Goble, 2011), in Spanish language classes (Hernández, 2011), and there has also been research from an endless variety of perspectives including historical ones (Lewis, 2011 and Marcus 2009) and in genres of financial disclosure (Camiciottoli, 2009). This list is not even representative of the totality of what is under study and does not even account for the theoretical and DM-specific debates in the field that have been going on for nearly half a century.

To begin with, the term DM is not a clearly and/or concisely defined term in the DM literature. Moreover, DMs have also been referred to using many other terms (e.g. discourse particles), attributed with numerous sometimes conflicting features, and there is no consensus about many of their functions and whether those functions are universally applicable. There are broad but detailed DM frameworks, some of which I outline in the following chapter, but beyond that it is harder to decide what aspects of the current debate are accepted without contention. I discuss these debates briefly because amidst all these differences of opinion and analyses, some elements do emerge very clearly: the phenomena of DM exists; almost everyone vaguely adheres to a tri-corporal outline of what they perceive them to be; many different languages have them; and across the spectrum of all these languages they perform relatively similar (usually pragmatic or extra-linguistic) functions. Therefore, rather than dwell on the differences between the theorists that have been in the field for several years (the study of DM is over four decades old), I will work with an eclectic and flexible framework within which I can analyse the uses and functions of *ke*. This framework is provided by Schourup (1999) and it corresponds with another popular framework by Schiffrin (1987) and parts of Fraser's (1996, 2009) models on the study of DMs.

I have decided to use both written and spoken (and then transcribed) data from a variety of utterances or genres which is outlined in Chapter 2. I have used interviews; oral narratives; Facebook, internet or blog posts; school or formal learning text books, including dictionaries and grammar books; an extract from the Bible; prose and many other kinds of miscellaneous texts. Some utterances I have analysed both quantitatively and qualitatively but most of them I have examined only qualitatively. Although I have presented most of the data quantitatively, I have mostly confined myself to qualitative analysis except in cases where the argument can only be demonstrated by reference to the presented frequency counts. I am not alone in preferring such an approach to DM analysis; Kroon (1995: 3) has also expressed mild scepticism toward adherence to statistical scientism when studying DMs, stating that her study of the Latin DMs *nam*, *enim*, *autem*, *vero* and *at* is:

not to be characterised as corpus-based research in the sense of aiming to explain all instances in a fixed corpus, a method I regard as not very fruitful for the present undertaking: by studying all instances of a particular particle in a subcorpus of Latin

texts, one runs the risk of registering unnecessarily many identical and unproblematic instances while potentially missing the rare, really interesting ones. [...] Statistical observations will, accordingly, be supplied on an *ad hoc* basis only.

The outline of the thesis is as follows: Chapter 1 gives a brief survey of studies that have been done in the field of DMs. It is a review of the literature and it outlines the various working definitions that have been provided for what is are DMs. It looks at the main features that qualify a lexical item to be considered a DM, and therefore what features *ke* would have to fulfil in order to be considered for such a classification. The second chapter focuses on the methodology of the study and the processes of data collection and selection. It also looks at the complexities of these three processes, and I motivate for the necessity of my approach, and also provide a preliminary analysis of data in its concluding section.

Chapter 3 is the first and main analysis chapter and it is concerned with three tasks. It briefly considers how Xhosa grammars and dictionaries have defined *ke* thus far and examines the relations between these definitions with the category of DMs, and the attendant complexities of that relationship (or lack of). It then looks at the uses of *ke* in the data in order to examine whether it functions as a DM or not. Chapter 4 then looks at the historicity of *ke*, tracing its use in mid-late 19th century prose texts before shifting attention to its changing, unstable functions from that period onward by looking at the same Biblical narrative over various translations of the Bible into Xhosa. Only the Xhosa data examples used in Chapter 3 and in Section 4.2 of Chapter 4 – for the main, fundamental discussions – have been glossed; the Bible extracts used in Section 4.3 and examples taken from the research literature have not been glossed.

Finally, the Conclusion is both a reflection on the process of compiling this study (its findings, its shortcomings) and a detailed outline of possible directions for future research in the field of DMs in general and of Xhosa DMs in particular.

CHAPTER 1

TERMS, DEFINITIONS, FUNCTIONS & CHARACTERISTICS: A FRAMEWORK FOR THE STUDY OF DISCOURSE MARKERS

Eliminate all the other factors, and the one which remains must be the truth.
Arthur Conan Doyle³

1.1 Introduction

This chapter summarises linguistic debates about the meaning and functions of discourse markers (DMs). The chapter has three main sections. Section 1.2 is a summary of the various terms that have been proposed for DMs. Section 1.3 outlines the definitions and functions of DMs found in the literature. Section 1.4 is a detailed look at the characteristics of DMs as proposed by three models: Schiffrin (1987), Schourup (1999) and Fraser (1996, 2009). Taken together, these three models provide a comprehensive framework which spans the last three decades of DM research. I will highlight the similarities and differences between these models as well as some others that have emerged among researchers studying DMs.

1.2 Terms

Schourup's *Discourse Markers* (1999) article will serve as the basis for much of the discussion in this chapter. Although over a decade old, this forty-two page overview of the study of DMs in linguistics is, together with Brinton's (1996) overview, one of the earliest and, in my view, one of the most comprehensive in the field. In this article, Schourup synthesises the many and often divergent views on the topic. Early into his paper, Schourup (1999: 228) notes that despite numerous studies that have attempted to specify the meaning or function of particular DMs or DMs across languages, 'no consensus has emerged regarding the fundamental issues of

³Doyle, A. C. 1890 (2001). *The Sign of the Four*. London: Penguin Classics (p. 9).

terminology and classification.’ Even the term ‘discourse marker’ is in dispute, with researchers using various words and phrases to refer to this phenomenon in a variety of languages.

According to Schourup (1999: 228), terminology ‘presents a particular difficulty. The term DM [...] is merely the most popular of a host of competing terms used with partially overlapping reference.’ He points out that by 1996, there were already more than twenty such terms in circulation. For instance, Brinton (1996: 1-2, 6, 17) lists terms from the literature that included colourful ones such *discourse lubricants*, *conversational greasers*, *discourse glue*, *traffic markers*, *fillers*, *fumble*, *gambit*, *clue words*, *minstrel tags*, and *mystery particles* (italics my own). Maschler (1994: 325) saw DMs as ‘a *subcategory of metalingual expressions*: those used to mark boundaries of continuous discourse’ as a way of framing various parts of the text. Years later, Camiciotolli (2010), who prefers the term *discourse connectives*, also listed a range of available terms such as *pragmatic connectives* (Van Dijk, 1979), *discourse markers* (Schiffrin, 1987; Jucker and Ziv, 1998; Fraser, 2006), *sentence connectives* (Halliday and Hasan, 1976), *pragmatic markers* (Brinton, 1996; Andersen, 2001), *linking adverbials* (Biber et al., 1999), *adverbial connectors* (Altenberg, 2006), *logical connectives* (Hyland, 1998), and *discourse particles* (Fischer, 2006).

Among this wide range of terms, DM is the most commonly used term. According to Schourup (1999: 228-9), the second most common term is discourse particle (DP). However, DM was favoured because the term ‘has a narrower range of reference and has been subject to more precise attempts at definition.’ The term DP, on the other hand, because of the presence of ‘particle’ has the potential of being confused with other compound labels (such as utterance particle, conversational particle, sentence particle, final particle, etc.) which confer tentative grammatical status on a particular set of items as a matter of convenience (Schourup, 1999: 229-230).

1.3 Definitions and Functions

When it comes to defining DMs, researchers are once again as diverse in their suggestions as they are with the terminology. Brinton (1996: 30) argued that the definitions of DMs ‘seem to bear little resemblance to one another’, and then provides

a list of 15 available definitions. De Klerk (2005b: 1184) echoes this assessment when she writes:

Despite all the work to date, in which discourse markers have been analysed from a discourse analytical perspective, from a point of view of conversational analysis, interactional sociolinguistics, relevance theory, and lexical approaches, there is still no clear definition of exactly what is understood by the term ‘discourse marker.’

Without getting into a discussion about how the five schools of thought differ from or are similar to each other, what is significant to note from De Klerk's statement is that none of them have succeeded in presenting us with an indisputable definition of what is a DM. What also complicates the matter is the fact that ‘the referential overlap’ (Schourup, 1999: 242) between the term DM and other suggested terms, can be as great as that between variant definitions of the term DM itself. Furthermore, there is also an overlap between what researchers provide as definitions and what others provide as functions of DMs. This is why in this section I will discuss both the definitions and functions of these markers at the same time. Redeker (1991: 1139), for instance, refers to them as discourse operators, and defines a discourse operator as ‘a word or phrase [...] that is uttered with the primary function of bringing to the listener’s attention a particular kind of linkage of the upcoming utterance with the immediately preceding discourse context.’ In this example, DM definition and DM function are intimately intertwined. Similarly, Schiffrin (1987: 31) defines DMs as ‘*sequentially dependent* elements which bracket units of talk’ (emphasis in the original), and this definition, like Redeker's (1991), also incorporates DM function.

For another illustrative example, Tree and Schrock (2002: 727) begin (and end on a similar note) by saying:

Although *you know* and *I mean* are frequent in spontaneous talk, researchers have not agreed on what purpose they serve. In fact, the indeterminacy of function can be seen as a hallmark of their overarching category, *discourse markers*.

Junker and Smith (1998) have described *you know* as an invitation by a speaker for others to fill out her/his meaning, whilst Stubbe and Holmes (1995) argue that it can be used by the speaker to decrease social distance; Schourup (1985) thinks that it is a particle used to avoid a pause, whilst Östman (1981) shows how it can be used to moderate mood.

I focus on this difficulty in terminology, definition and function of DMs because it is an important one to acknowledge for anyone researching DMs. For example, depending on what term one uses in describing these linguistic units (whether DM or pragmatic marker or discourse operators or discourse connectives, etc.), different definitions would come into play and this would determine which utterances are accepted as DMs and which are excluded from the analysis. This is particularly evident when outlining Fraser's (2009) model in Section 1.4.3 below. As a result of these on-going debates, attempting to construct a definite framework for DM research is a challenging task. Schourup (1999: 228) points out that:

With the profusion of approaches have come inevitable disputes concerning classification and function. These disputes have quickened in recent years as DMs have come to be seen not only as an underexplored facet of language behaviour but as a testing ground for hypothesis concerning the boundary between pragmatics and semantics and for theories of discourse structure and utterance interpretation.

Therefore a lot is at stake. However, instead of delving deeper into this maze of competing camps, I will rather take my cue from many researchers in this field such as Torres (2002: 65) who surmises that despite the many disputes on what constitutes a DM or how DMs should be analysed, 'at a basic level most linguists would agree that discourse markers contribute to the coherence of the discourse by signalling or marking a relationship across utterances.'

The functions of DMs depend strongly on the context in which they occur, but discourse coherence seems to be a central function. For instance, *well* is a DM that often signals the beginning of a narrative after a digression or interruption (Norrick, 2001: 850). Sometimes the function of DMs is to 'orientate listeners' and provide 'organizational functions' in narratives (ibid). Other functions include signalling hesitation, contrast or cancellation of some feature in the previous utterance, to introduce an exposition, to manage transitions in succeeding sections of speech within one speaker's utterances as well as between the utterances of separate speakers. DMs also mark and negotiate speaker roles, fill silences and mark the speaker's right to continue after s/he has organised what s/he wants to say. They signal hedges, are useful for back-channelling and minimal response, serve as end-markers, to repair one's or another's utterance, as face-threat mitigators or for use in other politeness strategies (Schourup, 1985, 1999; Schifffrin, 1987, Brinton, 1996; Heritage, 1998,

Fuller, 2003). This list of functions is by no means exhaustive, and different DMs perform all, some or none of these functions. I will return to these functions in greater detail in the analysis sections of this thesis, Chapter 3 and 4, when I will look at *ke* in particular to compare it to other DMs.

1.4 Characteristics of DMs

In this section I will summarise three often cited frameworks for the study of DMs: Schiffrin (1987), Schourup (1999) and Fraser (1996 and 2009). These frameworks, particularly Schourup's (1999) will guide the data analysis in Chapters 3 and 4.

1.4.1 Schiffrin's 1987 Model

Schiffrin's (1987) pioneering study of 11 DMs in English (*well, now, so, but, oh, because, or, I mean, and, then, and y'know*) provides for many researchers a generally accepted framework for the study of DMs. She sees DMs as linguistic units that help the organisation of discourse. She states that DMs mark 'units of talk' and create coherence within the discourse via the '*integration* among different components of talk' (Schiffrin, 1987: 330; original emphasis). Furthermore, she offers a 'tentative' systematic approach on how researchers could identify expressions as being DMs (Schiffrin, 1987: 328). According to Schiffrin (*ibid*), DMs have the following characteristics:

- (a) DMs are syntactically detachable from a sentence;
- (b) DMs tend to occur at the beginning of an utterance;
- (c) DMs have a range of prosodic contours, such as tonic stress followed by a pause and phonological reduction;
- (d) DMs have either no lexical meaning, a vague meaning, or are reflexive of the language or of the speaker.

Earlier in her argument Schiffrin (1987: 64, 330) also states that DMs are (e) multifunctional, (f) non-obligatory, (g) syntactically diverse and (h) contextual coordinates that add coherence. Schiffrin's outline (and terminology) corresponds very closely to Schourup's (1999), and therefore I will discuss it in greater depth

below and will compare these two models whilst simultaneously illustrating the claims of both models with examples in the next section.

1.4.2 Schourup's 1999 Model

Based on his review of the literature, Schourup (1999: 234) outlines seven core characteristics of DMs. Schourup's (1999: 234) seven characteristics are modelled on Schiffrin's earlier classification, but they take into consideration a greater range of differing perspectives on the subject, including perspectives that disagree with Schiffrin's classification. His categories synthesize the debates in the field even whilst following Schiffrin's proposal, but broaden some of Schiffrin's categories. Schourup also provides terms for the DM characteristics, which is helpful and allows for easy reference. Although some of these features remain contested amongst researchers, they have been generally accepted as a useful starting point in identifying DMs (see Fuller, 2002). The seven DM features proposed by Schourup are (1) connectivity; (2) optionality; (3) non-truth conditionality; (4) weak clause association; (5) initiality; (6) orality; and (7) multi-categoriality. Below I will look in detail at the seven characteristics of DMs and illustrate them with examples from the DM literature, especially examples from Schiffrin's (1987) work used as a way of showing the relation between these two theoretical frameworks.

Connectivity

DMs connect utterances or other discourse units (Schourup, 1999: 230). This characteristic corresponds to Schiffrin's (1987) central claim that DMs integrate different units of talk to create coherence. This characteristic, illustrated in (1) below, is often regarded as the most common, and is seen as a necessary feature an item must have in order to qualify as a DM. In the example below, a/b/c indicate what Schiffrin calls 'units of talk' (see 1987: 31) in the speech of one person, in this case Zelda's in which she is explaining why her daughter-in-law left her previous job.

1. a. *So*, my son wouldn't let her go back there anymore
b. And she tried to get other jobs and she couldn't.
c. *So* what she did, she got a job as uh bookkeeper in an office.

(Schiffrin, 1987: 203)

The DM in this example is *so*. According to Schiffrin (1987: 203), *so* often signals ‘fact-based cause and result relations.’ The event narrated in (a) is the result of prior story events and the event in (c) is the result of the daughter in-law’s inability to find other work as stated in (b). Thus, *so* connects two textual units in one speaker’s utterance by creating textual coherence between the narrative prior to (a) with the unit (a) as well as coherence between units (b) and (c). In example (2) below connectivity is established not within a single-speaker narrative, but between utterances of two separate speakers (A and B). The DM in this example is *well*.

2. B: I never got to go to Rome or Greece or Italy...
A: *Well*, you never chose to go to Rome or Greece or Italy...you could’ve gone, you did not choose to do it. I mean I never had more money than you to do it, I just did it, anyway.

(Fuller, 2003: 24)

Optionality

The characteristic of ‘optionality’ refers to the claim that removing a DM from an utterance does not affect the grammaticality of the sentence. Furthermore, DMs ‘do not enlarge the possibilities for semantic relationship between the elements they associate’ (Schourup, 1999: 231). This means that if a DM expression is omitted from an utterance, the sentence remains not only grammatical, but the coherence relationship it signals and the assumptions which underlie utterances are still maintained. The hearer can still discern them, although they are no longer explicitly cued. Schiffrin similarly states that DMs are non-obligatory (1987: 64). This characteristic is the second most frequently cited for DMs, and is illustrated in examples (3) and (4) below.

3. So you get *like* a worm and then you catch *like* a tiny stupid little dumb fish...
4. I *just* stayed home cause someone was taking care of me. And then I was *just* watching T.V. And I *just* took a nap.

(Tagliamonte and D’Arcy, 2004: 497)

In examples (3) and (4) the DMs *like* and *just* serve specific and important communicative functions, with *just* serving to emphasise the actions narrated after it and *like* allowing the speaker to fill the gap and keep the floor as they think of what next to say, as well as keep the mood of their utterance light and informal. Thus, optionality does not mean DMs are without function; they are dispensable because both these sentences would not be ungrammatical if the DMs were excluded.

Non-truth conditionality

This is the third most common DM characteristic. It is the claim that ‘DMs are generally thought to contribute nothing to the truth conditions of the proposition expressed by an utterance’ (Schourup, 1999: 232). What Schourup means by this is that, unlike ‘content’ words, DMs do not affect the truth or falsity of an utterance. This characteristic is sometimes grouped with the ‘optionality’ feature as the two categories overlap (cf. Fuller, 2003: 24). Schiffrin’s model also refers to DM particles as either having no meaning or a having a vague meaning. Example (5) below, for instance, would be a perfectly grammatical sentence and it would have the same propositional meaning even if *I mean* were absent.

5. she appears to be perfectly happy - *I mean* she can’t be a hundred per cent happy, nobody is, but she appears to be happy.

(Fox Tree and Schrock, 2002: 733)

According to Fox Tree and Schrock (2002), *I mean* fulfils a repair function in this extract. This is an important pragmatic function, but it does not affect the grammaticality or truth condition of the utterance. In addition, *I mean* performs the connectivity function noted above by linking the two statements together. And finally, *I mean* also expresses emphasis in this sentence. Despite this multiplicity of function (which is also typical for a DM as I will discuss below), *I mean* does not affect the truth conditions of (5). Again, non-truth conditionality, like optionality, does not mean that a DM does not have a function in an utterance in which it occurs; as noted, *I mean* does have multiple communicative functions in (5). The meanings that DMs bring to the utterance are not obligatory (content/grammatical) but ‘additional’ (communicative).

The remaining four characteristics (weak clause association, initiality, orality, and multi-categoriality) are not as frequent as the first three in the definitions of DMs or even as widely accepted, but they will nonetheless be discussed below.

Weak clause association

This characteristic highlights the fact that DMs are often clause elements that have ‘a detached role relative to closely interrelated clause elements such as subject, complement, and object’ (Schourup, 1999: 232). In other words they are loosely attached to sentence constituents, or even completely outside the syntactic structure of an utterance. This characteristic has been linked to phonological independence, i.e. DMs are said to constitute ‘independent tone units’; they are frequently separated from the main clause by ‘comma intonation’. However, tonal separation is not a sufficient or necessary feature of DMs (Schourup, 1999: 233). Schiffrin (1987) also argued that DMs must ‘be syntactically detachable from a sentence’ as well as have ‘a range of prosodic contours, such as tonic stress followed by a pause and phonological reduction’. Examples of DMs that are strongly associated with this characteristic are pause fillers such as *yeah*, *um*, *ah*, and *oh*. This is illustrated in (6) where *ah* is coupled with a laugh to express appreciation of a joke, and as indicated by the transcription symbol #, *ah* is an independent tonal unit.

6. C: you’re dirty whore# and she said no I ain’t# I’m a clean whore#
B: ***ah***# (laughs) that’s rather nice# (laughs)

(Adapted from Aijmer, 2002: 146. The prosodic symbol # indicates tone unit boundaries)

Ah should not be confused with *oh*, another, and perhaps more common, DM that can be loosely associated to clauses. For instance, Aijmer (2002: 147) found in her data⁴ that ‘*ah* was a separate unit more often than *oh*... [and] it is not used in more or less lexicalized combinations (*ah look*) or exclamations (*ah God*) where *oh* would be normal.’ However, the two DMs share a number of functions such as expressing surprise and unexpectedness, and they also tend to occur in similar contexts such as responses to utterances and back-channelling.

⁴The data used by Aijmer is known as the London-Lund Corpus of Spoken English (LLC). It comprises of approximately a million words distributed over texts of around 5 000 words. These are recordings from the 1960s and 1970s of face-to-face conversations, telephone conversations, public discussions and interviews, unprepared public commentary, and prepared public oration (Aijmer, 2002: 4).

Two other similarly related DMs are the allomorphs *iya* (yes) and *ya* (yeah) found frequently in Indonesian conversation. According to Wouk (2001: 175), these two ‘are found in virtually all possible positions of an IU [intonation unit],’ thus indicating their loose associations with any clause in a sentence. This issue of positionality of a DM leads directly to the next characteristic, initiality.

Initiality

DMs usually introduce the discourse units they mark, but this refers to ‘the position of DMs in relation to the central clause elements rather than to the position of the first word in an utterance’ (Schourup, 1999: 234). This also matches with Schiffrin’s model above which claims that DMs must ‘be commonly used in the initial position of an utterance’. In other words, a DM could signal the introduction of new information, a shift in topic, a repair within an utterance, or a response. It can do so within clauses and not necessarily at the beginning of a clause or sentence.

7. a. (***After all/Now/However***), corgis are an intelligent breed.
b. Corgis, (***after all/now/however***), are an intelligent breed.

(Schourup, 1999: 234)

8. a. ***Yeah***, I went and I found no one
b. I went, ***yeah***, and I found no one.
c. I went and, ***yeah***, I found no one.
d. I went and I found no one, ***yeah***.

In example (7a) and (8a) the DMs *after all*, *now*, *however* and *yeah* are literally at the start of the utterances. However, in (7b), (8b) and (8c), although the DMs are not at the start of the sentence unit, they nonetheless still occupy initial-positions in the subsequent clause units within an utterance. Example (8d) demonstrates that the initiality position of a DM is not a fixed feature. In this example, we have a DM that appears at the end of an utterance to signal completion of an utterance. Again, in Wouk’s (2001: 175) study of Indonesian conversation, *iya* and *ya* ‘can occur IU-internal, final or initial. The IUs [intonation units] that *ya* and *iya* begin or end may come turn-initially, turn-finally, or turn-medially.’

Orality

This refers to the observation that DMs occur most frequently in spoken utterances. However, according to Schourup (1999: 234), ‘no principled grounds exist on which to deny DM status to similar items that are largely found in written discourse (e.g. *moreover, consequently, contrariwise*)’. It is certainly true that a lot of DM research has focused on spoken utterances, especially conversations. However, the distribution of DMs between speech and writing is hardly strict. It has a lot to do with the formality or informality of the DM in question, as well as the utterance or genre in which it occurs. Studying DM usage by Turkish English-as-a-Second-Language (ESL) learners, for example, Dülger (n.d: 267) concludes by saying that the ‘effective usage of discourse markers is an indispensable part of attaining cohesive, coherent and unified pieces of written texts.’ Another study carried out by Kroon (1995) on five DMs in Latin was based exclusively on written materials. Even in speech, we encounter some contradictions; ‘*you know* and *I mean* are more common in “considered talk” than in spontaneous talk, although spontaneous talk should be more conversational, casual, and fast’ (Tree and Schrock, 2002: 729-730). Nonetheless, there are clear-cut instances of the use of DMs that are typical for casual conversation and unlikely in formal contexts, especially in writing. Example (9), from a focus group discussion at the University of Cape Town in 2007 illustrates this. The speaker is an English-Xhosa bilingual.

9. *Like*, from my understanding, the rural Xhosa is *like* more pure, it’s *like* exact Xhosa that we’re supposed to speak. *And then* in town it’s *like*, it’s *like* more *like*; it’s vague. For example when they’re saying... *Um*... If you’re asking the person ‘Where do you stay?’, *he’s like* ‘kaLanga baba’, *you know*? What’s that, ‘kaLanga’? In Xhosa it’s *like* ‘KwaLanga’ – it’s K-W-A: ‘KwaLanga’, *you know*...

(Thozama, Focus Group 2, 2007)

In this example various DMs are performing different functions. Most of the *like* tokens are multifunctional: repairing false starts, hedging, gap-filling, and one (*he’s like*) is quotative. Such usage would only be permissible in written texts that strive for extreme informality, such as dialogue in a play, film script or novel.

Multi-categoriality

This is the last DM characteristic and it emphasises that ‘DM status is independent of syntactic categorization; an item retains its non-DM syntactic categorization but does “extra duty” as a non-truth-conditional connective loosely associated with clause structure’ (Schourup, 1999: 234). For example, adverbs, coordinating and subordinating conjunctions, interjectives, and verb phrases (such as *you see, I mean, you know*) can all take DM functions without necessarily losing their syntactic functions. Thus, Hlavac (2006: 1890) notes that grammar textbooks define *so* as an adjective, adverb and conjunction, even though *so* has also acquired DM-related functions as seen in example (1) above.

This characteristic of DMs is not the same as the polysemy of functions that DMs can perform, as demonstrated by *like* in the example (9) discussed above. Under ‘multicategoriality’, some DMs might even have ‘recoverable literal meanings’ (Fuller, 2003: 23) that are maintained whilst performing DM functions as in (10) below, where *or* still means *or* even though it’s also performing DM functions; this is not the case in (9) above where the *likes* are DMs (except in the case where it is quotative) with no literal meanings or grammatical functions. In (10) below, Jack is stating how the elites use social diversity to justify economic inequality:

10. a. It’s when there’s an economic situation
- b. then they use race.
- c. ***Or*** they use nationality,
- d. ***or***anti Semitism
- e. ***or*** what have you.

(Schiffrin, 1987: 178)

In this example, *or* connects this litany of discriminatory categories to one another. But as Schiffrin points out, one *or* would have been sufficient for Jack to make his point that elites can misuse diversity for their own economic gain. Therefore, over and above being a coordinating conjunction *or* also allows Jack to ‘mark inclusive options’ which allows him ‘interactional advantages gained through the use of multiple evidence’, including a generalisation: the tag *what have you* (Schiffrin, 1987: 178). Thus, *or* functions both as a grammatical particle and as a DM; syntactic and discourse functions overlap, and therefore the DM *or* can be described as

multicategorical. As I already stated above in reference to Schiffrin's (1987: 64) model, 'multi-functionality' is another term often used in literature to describe DMs, because they can be 'used in several discourse capacities simultaneously'. Another example is the Spanish adverbial *ya* ('already, finally, presently, now') which can be used as an adverbial to convey aspectual information, as well as a DM to convey emotional intensity and to create discourse cohesion (Koike, 1996: 267).

1.4.3 Fraser's Model

Fraser's DM model is useful to outline simply because it demonstrates another attempt at exhaustiveness over an extended period of time to establish a stable definition of DMs. Like the Schourup and Schiffrin definitions above, Fraser's cumulative model (1988, 1990, 1993, 1996, 1999, 2006 and 2009) also includes connectivity and non-truth conditionality in its definition of DMs. In Fraser's model, the third most popular DM feature, optionality, is seemingly collapsed with non-truth conditionality (see also Fuller, 2003: 24), thus making his model a classic DM model in its recognition of these features as key features that lexical items must have in order to be classified as DMs. Nonetheless, the model has been criticised by Blakemore (1996) for its category divisions which lump together items that perform very different functions.

In Fraser's earlier work (1988, 1990 and 1996) DMs are a subclass of *pragmatic markers* which are linguistic constituents that do not affect the propositional meaning of a sentence even though they contribute towards signalling the speaker's possible intension. Within Fraser's framework there were four types of pragmatic markers: basic markers (e.g. *please, admittedly*), commentary markers (e.g. *stupidly, sadly, frankly, mark my words, certainly, allegedly*), parallel markers (e.g. *damned*, as in: 'Get your *damned* shoes off the table), and DMs (*after all, besides, as a result*). According to this framework, DMs are seen as elements that connect discourse units by relating the basic message to the foregoing discourse. Thus, Fraser discusses them as connective elements that differ from the other pragmatic markers in that they do not form part of the representational meaning of the utterance (i.e. non-truth conditionality). Elsewhere, Fraser (1993, 1999) proposed that there are two classes of DMs: those that relate topics and those that relate messages. The last group subdivides

into three classes: contrastive (*in comparison*), elaborative (*furthermore, moreover*) and inferential (*thus, hence*).

In later work, Fraser (2009) reshuffled his categories and introduced another one. By now the category of parallel markers, for instance, was incorporated into commentary markers. In its place Fraser introduced the category of discourse management markers, ‘which signal a metacomment on the structure of the discourse’ (2009: 893). This category in turn has three subcategories: discourse structure markers (*first, then, in summary, I add*); topic orientation markers (*anyway, back to my original point, before I forget, but, by the way, incidentally*); and attention markers (*ah, alright, anyway, anyhow, hey, in any case, in any event, now, now then*).

Fraser’s discourse structure markers (as opposed to his DM category) is worth pointing out in his model because it contains some items that other researchers have classified as DMs, such as *but* in Schiffrin’s model and *ah* as seen in example (6) above from Aijmer’s (2002) classification. Moreover, the discourse structure markers in Fraser’s model are significant because the functions of the items listed under it are similar to DM functions reported elsewhere. For example, the listed functions of the sub-category of topic orientation markers are that they typically signal the speaker’s intention to (i) return to a prior topic, (ii) add to or continue with the present topic, (iii) digress from the present topic and (iv) introduce a new topic (Fraser, 2009: 894). All these functions, as will be demonstrated in Chapter 3, are classic DM functions as well. Therefore, Fraser’s distinction between DMs and discourse structure markers is tenuous and complicated and as such will not be as central to the conceptual framework of this thesis as much as Schiffrin’s (1987) and Schourup’s (1999) models.

1.5 Conclusion

In this chapter, I discussed the debates around the terms, definitions and functions of DMs. I showed how there are many terms in the literature that are used for more or less the same phenomenon, including the fact that by 1996 there were already at least twenty terms in use. Although the definitions of DMs are also as varied, I noted that the basic, generally accepted definition of a DM is that of a linguistic unit that relates

sequentially dependant clauses or utterances thus creating coherence. They also perform the “extra duty” of various pragmatic, interpersonal functions in discourse that distinguish them from grammatical (content) lexical items. Nonetheless, there is still a lot of debate even with what has already been discussed. A good example is the fate of the eleven linguistic items originally accepted by Schiffrin (1987) as DMs in English: *oh*, *well*, *and*, *but*, *or so*, *because*, *now*, *then*, *y’know*, and *I mean*; none of these has been met with unqualified acceptance as a DM in subsequent literature (Schourup, 1999: 242). For example, Fraser (1988: 26-27, 1990: 388, 395) excluded *oh* because it is an interjection, *because* because it is a subordinating conjunction expressing content, and *y’know* and *I mean*, because they are separate utterances signalling ‘a speaker’s attitude of solidarity’; in this regard, Fraser does not seem to consider multicategoriality as an acceptable feature of DMs. The debate regarding these items, therefore, is still on-going.

This reoccurring pattern (the struggle to define terms neatly, consistently and with definite boundaries; what they include and what they exclude) is a standard predicament in linguistics, and the field of DM studies is no exception. Moreover, this challenge is not avoidable: language is complex and attempts at providing neat boundaries on what it is, how it works and how it is used are bound to bring this complexity to the fore. Nonetheless, having looked at these characteristics of what makes a lexical item a DM, one now has a clearer idea of what characteristics to expect from *ke* for it to be considered a DM. These seven features are the basis of the discussion of *ke* as a DM in Chapter 3, where I assess the validity of the claim that *ke* is a DM. Before that however, Chapter 2 will outline the data sets used for the analysis of *ke* in this thesis.

CHAPTER 2

FORMALISING CURIOSITY: DATA COLLECTION & METHODOLOGY

Research is formalized curiosity. It is poking and prying with a purpose.
Zora Neale Hurston⁵

2.1 Introduction

This dissertation is based on spoken and written language data. The *spoken data* is in the form of recorded and transcribed research interviews conducted in 2006, and focus groups conducted in 2007. In addition, seven transcribed Xhosa oral narratives from the early 1970s are used. These were selected from stories narrated by Nongenile Masithathu Zenani of Mboxo (Nkanga), in the Gatyana District, in the former Transkei and compiled by Harold Scheub (2006). This transcribed data will be used mainly to compare how *ke* functions in spoken data as opposed to written data.

Table 1. The two types of data

SPOKEN DATA	WRITTEN DATA
9 Interviews (11 individuals & 2 groups)	7 Xhosa translations of the Bible (Luke 15: 11-32)
7 Oral Narratives	Xhosa translation of South African Constitution
5 Focus Groups	Electronic data (3 blogs and 1 Facebook group)

The *written data* falls under three main categories: seven Xhosa translations of the biblical story of the prodigal son in Luke 15: 11-35; the Xhosa translation of the South African Constitution; and electronic data (blogs and Facebook). The electronic data serves as a bridge between written and spoken genres because although it is written, it retains a lot of the informal aspects of oral language and it is worth seeing how *ke* behaves in that context. The biblical narrative of the prodigal son helps to see whether *ke* also carried DM in functions in nineteenth century Xhosa, or whether this is a new development. The thesis will also use some extracts from nineteenth century literary texts to further illustrate this point but these texts do not form part of the core

⁵ Lewis, J. J. 2003. Women's voices: Quotations by women. What you need to know about women's history. Retrieved January 18, 2010, from <http://womenshistory.about.com/library/qu/blquhurs.htm>

data and will only be referenced when they are eventually cited. The first Bible translation used in this thesis is from 1853, followed by translations and editions from 1864, 1879, 1921, 1931, 1942, and 1962. The Xhosa translation of the Constitution is used to identify the functions of *ke* in a document located within a formal legal genre.

The quantified data is from nine interview recordings (two group interviews and eleven individuals – since four of the nine interviews had two interviewees each); the seven oral narratives; the three blogs and the Facebook group. The *ke* tokens in the five focus groups, the biblical narratives and the Constitution data have also been counted but this data will only be used qualitatively.

As Deborah Cameron (1990: 85) has noted, in variationist linguistic research that is framed within the quantitative paradigm ‘statistical correlations are used to relate frequency scores on linguistic variables to non-linguistic features both demographic (class, ethnicity, gender, age, locality, group structure) and contextual (topic, setting, level of formality).’ In this thesis, the frequency scores of the *ke* variable in the quantitative corpus are related to the contextual features – its function within a sentence and the level of formality of the text – but not to the demographic correlates Cameron (1990) mentions.

2.2 Research Questions

Although – as discussed in the previous chapter – orality is a common characteristic of DMs, DMs also occur in written language. The bi-modal (spoken and written) corpus allows for the investigation of the following questions in more depth:

- (a) whether *ke* is a DM,
- (b) whether, as a DM, *ke* is restricted to orality,
- (c) if there is a difference in its functions in formal/written vs. oral genres, and
- (d) based on the biblical translations and some literary texts, analyse whether *ke* is an old or recent DM particle.

These issues cannot be sufficiently explored using a single data set of spoken utterances. The varied data set allows one to answer these questions sufficiently. Below are the detailed outlines of each data set.

2.3 Types of data

2.3.1 Spoken corpus 1: Interviews and focus groups

The interviews used in this study were conducted in 2006 as part of a research project on South African internal migration and language change.⁶ The focus of that project was on individuals who had experienced migration from the Eastern Cape Province to the Western Cape Province of South Africa, as well as international migrants to South Africa from other African countries such as Angola, Cameroon, Malawi, and Somalia. The international migrant interviews are not used for this thesis (even though some of them were in Xhosa as a second language and *ke* was present). The focus of this thesis is on first language Xhosa speakers, and 9 interviews are used for analysis. The interview corpus was collected over a period of a year in Gugulethu.

According to the 2001 Census (the Census that would have been relevant at the time of the data collection), 94.59 percent of Gugulethu residents spoke Xhosa as their first language, 98.80 percent were black, 52.49 percent female, and 51.22 percent unemployed. In terms of educational levels, 62.54 percent left school before Grade 12 and fewer than 9 percent had no schooling at all; only 22 percent had Grade 12 and only 6 percent had a post-Grade 12 qualification. The area was also dominated by the youth: 26.77 percent of the residents being between the ages 35-64 and the rest younger than that, with the 18-34 group dominating at 37.59 percent.

As the census figures indicate, overall, Gugulethu is an impoverished urban area with high levels of unemployment and low educational performance. It can be described as an urban Xhosa enclave, with over 90% of residents listing Xhosa as their home language. The interviewees were found in the NY1 area in Gugulethu, particularly

⁶ The data was collected as part of a SANPAD/NRF research project on language and rural-urban migration led by Assoc. Prof. Ana Deumert and Prof. Rajend Mesthrie. Gratitude is expressed to the sponsors for the funds to run the project, and another thank you to the UCT Linguistics section for permission to use the data for my thesis. Thanks also go to all those involved in the fieldwork and the transcription of the corpus: Marion Chirwa, Magdalene Mbong, Germain Kouame, Ellen Hurst and Goodwill. Special thanks to Zukile Jama and Zolani Kupe whose Xhosa interviews, transcriptions and translations are used extensively in Chapter 4 of this thesis. Two interviews (Nelson and Nomelikhaya) come from another data set collected by Ana Deumert and Nkululeko Mabandla (Monash Project on Rural-Urban Migration, funded by Monash University, Australia) I am grateful for permission to use the interviews. All the usual disclaimers apply.

around the Tshatshalazeni meat market which is very close to the taxi station and is built in front of the (ex)migrant hostels. The social public space of the Tshatshalazeni meat market is socio-economically of low status, with a large amount of informal housing surrounding it. During the time of the interviews, the ex(migrant) hostels had been recently rescued from a state of decay through government-sponsored renovations, but were engulfed by shacks in serious disrepair.

The basic structure of the fieldwork was as follows. We would ask the participants questions about various aspects of their lives such as where were they born; when they came to Cape Town; how they experienced Cape Town compared to life back at home; who their friends were, and what they did together for fun; what work they did; whether they still had contact with people from their initial or family home, etc. Although the researchers had a prepared set of questions and knew what was essential for each participant to provide answers to, the interviewers adopted an informal approach when talking to the participants in order to elicit conversational data as much as possible. For example, towards the end of the data collection period, as a way of making the interviews less formal and more conversational, we hardly consulted the questionnaire.

The recorded interviews were then transcribed and translated into English. For purposes of this research, the total number of words per individual interviewed was counted and extracted from the transcript (i.e. interviewer's questions were kept separate from the participant's response). The frequency of *ke* tokens and word count were calculated for each individual interviewed using the Simple Concordance program, a free, downloadable concordancer (<http://www.textworld.com/scp/>) and then manually verified. In cases where two individuals were interviewed at the same time, each person's words and tokens have been counted separately. Some interviews had more than two people being interviewed at the same time and some sessions were closer to focus groups than interviews (for example, there are interviews with up to five participants). In this thesis, two interviews of the latter kind will be used, and the tokens produced by all participants in these interviews are treated as one count, i.e. there is no allocation of tokens to each individual present due to the difficulty of distinguishing them in the recordings and the transcripts. Therefore, out of the nine

interview recordings used, there are eleven individuals and two groups (with five and four participants each) for the total word and *ke* token counts seen in Table 2 below.

Table 2. Interviewee Data

PARTICIPANT	SEX	AGE (in 2006)	NO. of WORDS	KE TOKENS	KE per 100 WORDS
Gumede	F	66	1 685	125	7.42
Mbele*	M	49	939	50	5.32
Nontokozi†	F	30-something	413	19	4.6
Ngomane*	M	60	914	41	4.49
Dandala∅	M	Unknown	161	7	4.35
Sandile‡	M	26	1 306	54	4.13
Mandla‡	M	28	722	29	4.02
Group 1 of females	F (5)	20-30	1 820	58	3.19
Mzolisiti†	M	36	1 060	28	2.64
Nomelikhaya	F	26	385	8	2.08
Nceba∅	M	26	975	18	1.85
Nilson	M	49	2381	19	0.8
Group 2 of females	F (4)	21, 22, 23, 25	1 855	14	0.75

N.B: The accompanying symbols indicate which individuals were interviewed together. All participants have been assigned pseudonyms.

The personal histories of all the individuals are, needless to say, vast and complex, and details of some of them will be presented throughout the thesis only when the need arises. The sample includes participants who are working night shifts as security guards or petrol attendants or supermarket employees on leave from work, entrepreneurs and small business owners, unemployed individuals, and new arrivals in Cape Town. Most of them did not have post-Grade 12 qualifications, and had been educated in the Eastern Cape. For purposes of a short argument about the nature of *ke* and its use, particularly in an interview context, I have also extracted the word and token counts of three interviewers (Zukile Jama, Zolani Kupe, and myself Oscar Masinyana) from 5 of the selected interviews. The interviewer tokens are listed in Table 3 below.

Table 3. Selected Interviewer *Ke* Tokens

INTERVIEWER	INTERVIEW WITH	NO of WORDS	<i>KE</i> TOKENS	<i>KE</i> per 100 WORDS
Oscar Masinyana	Mbele and Ngomane	675	25	3.70
Zolani Kupe	Mbele and Ngomane	667	24	3.6
Zolani Kupe	Group 1 of females	956	29	3.03
Zolani Kupe	Mzolisi and Nontokozo	284	8	2.82
Oscar Masinyana	Mzolisi and Nontokozo	800	22	2.75
Oscar Masinyana	Sandile and Mandla	227	5	2.20
Zukile Jama	Gumede	901	14	1.55
Zukile Jama	Sandile and Mandla	636	5	0.79
Zukile Jama	Group 2 of females	801	6	0.75

The five focus groups from which I will draw some insights in the analysis section were carried out by Assoc. Prof. Ana Deumert and me at the University of Cape Town (UCT) in 2007⁷. These focus groups were conducted mainly as control groups for the data that we already had, and we needed non-specialist or external insights about some linguistic observations we had made about the data at that point. The groups had no more than six participants, male and female, between the ages of 20 and 30 (although in one group there was a relatively more mature participant), and the participants were allowed to use either English or Xhosa during the session. All participants in these focus groups, except one, were university students.

Table 4. Focus Groups Data (2007)

FOCUS GROUP	NO. of PARTICIPANTS	SEX	AGE RANGE	LENGTH (hour/mins/secs)
Group 1	6	5 F / 1 M	20-30	1.36.58
Group 2	5	2 F / 3 M	22-44	1.21.38
Group 3	4	5 F	20-29	1.33.04
Group 4	5	3 F / 2 M	18-24	1.37.06
Group 5	5	5 M	20-29	1.26.55

The questions in these focus groups centred on variation in Xhosa, with questions such as whether they thought there was a difference between urban and rural Xhosa

⁷There was a sixth focus group conducted as part of this leg of the research, but because of the poor quality of the recording, I have excluded it as a source of data.

varieties, and if so what these differences were. We asked them what they thought the future of Xhosa was, whether they used Xhosa in their SMSs (text messages), whether they could tell if someone was from the township or from a rural area and if yes, how. For the purposes of this particular study, we probed their perceptions about *ke*: who uses it?, how?, why?; when is it appropriate to use it in a sentence?; what does it mean?. We also provided them with examples from interview data in order to gauge their perceptions about the functions of *ke* in those particular examples. As already mentioned, these focus groups are not used for quantitative analysis.

2.3.2 Spoken Corpus 2: Oral Narratives

Seven transcribed Xhosa oral narratives from early 1970s are used for quantitative and qualitative purposes. The selected tales (*iintsomi* in Xhosa) were all told by Nongenile Masithathu Zenani at her home in Mboxo (Nkanga), Gatyana District, in the old Transkei. They were then compiled into a volume by Harold Scheub (2006). Unfortunately, I did not have access to the audio tapes but relied on the Xhosa transcripts available online (<http://digicoll.library.wisc.edu/SouAfrVc/>). These transcripts have preserved *iintsomi* performed by a skilled artist in the art of storytelling—a highly stylised and traditional literary form in Xhosa discourse. Like the interviews, the blogs and the *Facebook* corpus, these narratives are another discourse genre in which *ke* tokens are quantified.

Table 5. Oral Narratives Data

STORY TITLE	NO. of WORDS	KE TOKENS	KE per 100 WORDS
Origins of the Xhosa	5 871	363	6.18
A Typical Day in the Life of a Xhosa Woman	4 334	209	4.82
Construction of an Instomi	417	17	4.08
A Poor Girl Marries a King	2 519	94	3.73
Seven Men Confront a Strange Fowl	2 537	85	3.35
A Man Hides Food from his Family	4 715	129	2.74
A Child is Born without Limbs	1 522	34	2.23

2.3.3 Written Corpus 1: Online Material

The online/internet corpus was collected from a Facebook (www.facebook.com) group called *Xhosa People Rock!!!!*, and from 3 blogs written in Xhosa (and specifically targeting Xhosa people). These blogs are *Darkie's Place*, *Embo maXhosa*, and *Epozini: the House of Truth*. Data from these groups are quantified as shown in Table 6 below.

Table 6. Online Data

SOURCE	NO. of WORDS	KE TOKENS	KE per 100 WORDS
Embo maXhosa	20757	195	0.94
Epozini: the House of Truth	65535	598	0.91
Xhosa People Rock!!!!	14967	111	0.74
Darkie's Place	13342	36	0.27

Facebook started in February 2004 as a hobby project called *thefacebook* by Harvard student, Mark Zuckerberg. It was an instant success, with over 8 million members in the US alone by August 2006. Two years later, the Middle East-Africa region had 14.9 million (comScore, Inc, 2008), and according to Rabaney (2009) South Africa had 2 318 260 Facebook users by November of 2009. All age groups (under 17 to 61+) were well represented, with the elderly groups (56+) showing similar rapid growth patterns as with the group aged 25 and under.

One of the features of Facebook is that any registered member can start a group to which s/he can then invite other Facebook members who share a similar enthusiasm or interest to join the group. Private groups are only accessible by invitation or 'application', but public groups can be accessed and joined by anyone without invitation or group membership status. *Xhosa People Rock!!!!*, the group from which I use some data, is a public group and it began around June/July 2007. For this study I have limited the data to the topics posted between the 25th of July 2007 (the group's first post) and the post of the 7th of August 2008. In this time period, there were 35 posts and numerous responses to each posting. Group membership was at 7 355 on the 9th of June 2009 – the point at which I started collecting data – and had grown to 10 906 members by the 10th of March 2010. The group uses the slogan *Siluhlang' oluhle! Vumani bo!* ('We are a beautiful people! Concur!'). The group is one of a growing number of Facebook groups that singularly or predominantly concern

themselves with Xhosa-related matters. Some of these groups come across as voices of intervention and calls to cultural revitalisation at the face of a perceived looming threat to Xhosa cultural identity due to contact with Western values.⁸

However, data from online sites such as Facebook presents certain challenges. There are complex issues regarding online data ethics; i.e. just how private is personal data on Facebook? (Jones and Soltren, 2005; Barnes, 2006; and Dwyer and Hiltz, 2007. Cf. also Facebook's privacy policy: <http://www.facebook.com/policy.php?ref=pf>). Although anything posted on the internet or on Facebook automatically becomes public data (besides the Facebook user wallposts, chats and inbox messages), people tend to share very personal details about themselves in forums accessible to anyone. Therefore, I have used material only from the group's discussion boards because these tend to be less personal and more 'issue driven.' I have also changed names where applicable and nothing of a personal nature or identifiable information is reprinted.

Blog data comes from three Xhosa blogs. The first one is the now defunct *Darkie's Place* (<http://darkiesplace.blogspot.com>), which operated from the 17th of October 2003 to the 1st of September 2004. The blog had 34 posts. From its inception until its demise, it was one of the then rare examples of Xhosa web-based electronic communication. Like *Xhosa People Rock!!!!*, its motto was also about asserting cultural pride:

The online home for Xhosa speakers (and those who understand the language!). A place where the world can be seen from a different set of eyes ... a place where we can laugh at ourselves, learn from ourselves and love ourselves ... a place where we can embrace our "loxion culture" ... a place to inspire ... a place where we can be black, loud and proud!⁹

⁸ For illustration: other groups have names such as *how many Xhosas are on facebook* (lack of caps in original; 6 292 members in March 2010), *Xhosanostra* (4 546 members in April 2009; group now defunct), *khanizi thuthe bantu* ('Praise/declare yourselves, people'; 4 205 members in March 2010), *EZASE MTHATHA* [sic] ('UMthatha News', 283 members in March 2010), *I am a Xhosa Virgin and proud* (86 members in March 2010), *Amasiko amaXhosa* ('The Traditions and Customs of the Xhosa', private group, 105 members in March 2010), and *I love Xhosa men* (623 members in March 2010). The list and the members keep growing. These forums provide a unique linguistic ethnic space in the online medium.

⁹"Loxion culture" refers to "location culture", which is another way of saying "township location" culture. There was also a very popular clothing label called Loxion Culture that supposedly drew its inspiration from the townships.

The remaining two sources for online material, the blogs *Embo maXhosa* (<http://blogs.uct.ac.za/blog/embo-maxhosa2>) and *Epozini: the House of Truth* (<http://blogs.dispatch.co.za/epozini/>), have more formal contexts than *Darkie's Place* or *Xhosa People Rock!!!!*. However, they are formal contexts in differing ways. The former was a university-based (UCT) Xhosa blog that deals with issues of cultural pride and history, but the group also had an avid interest in Xhosa poetry and literature. The entries were written in a fairly academic style and often included references. The blog was started in October 2007 but came to a halt by July 2008, after 8 posts and 317 user comments (last checked on 29 March 2009). *Epozini*, on the other hand, is a one-man show run by Xolisa 'Mango' Mgwatyu. The only information I could put together about him was that he worked for the *Daily Dispatch* and was based in East London, in the Eastern Cape. From the way he explained the aims of his blog to his online audience, it is clear they are broad and wide-ranging:

Sizakuthi ke sixoxe izinto gabalala ngokuquka ezopolitiko, ezemfundo, ezolonwabo, ezempilo, ezeqoqosho, okusingqongileyo nezinye-nezinye.

We will broadly discuss issues concerning politics, education, entertainment, health, economic, things that surround/concern us, etcetera, etcetera.

Starting from the 15th of October 2007, I collected its posts and user comments up until the post of 20 March 2009. In total, there were 28 posts or topics of discussion and 826 user comments.

2.3.4 Written Corpus 2: Xhosa Translations of the Bible and Constitution

The biblical narrative of the prodigal son allows us to see whether *ke* also carried DM in functions in nineteenth century Xhosa, or whether this is a new development. The first translation used in this thesis is from 1853 and the last one is from 1962. The story of the prodigal son is useful for analysis, firstly, because it is a narrative and therefore fits well with the oral narratives data set and allows for comparison regarding the functions of *ke* in narratives. Secondly, it is a short, well-contained narrative and thus making possible the analysis of an entire narrative to best illustrate the global as well as local (sentence-level) functions of *ke* in narratives and discourse. Lastly, the various translations can shed light on the evolution of the use and function

of *ke* across time at least in one specific genre and narrative, something that is not possible with the other data sets.

The Xhosa translation of the Constitution is used to identify the frequency of *ke* in a document located within a formal legal genre. The analysis of this contemporary document will be compared with the use of *ke* in another legal document, ‘Umthetho Wokupapasha, 1977’ (Publications Act of 1977) from the Transkei statutes of 1977.

Table 7. Constitution and Bible (Luke 15: 11-32) Data

SOURCE	NO. of WORDS	KE TOKENS	KE per 100 WORDS
Xhosa Constitution	7 522	26	0.35
1853 Bible Translation	293	9	3
1962 Bible Translation/edition	309	16	5
1864 Bible Translation/edition	n/a	9	n/a
1879 Bible Translation/edition	n/a	14	n/a
1921 Bible Translation/edition	n/a	9	n/a
1931 Bible Translation/edition	n/a	15	n/a
1942 Bible Translation/edition	n/a	16	n/a

Additional literary and other sources of data will also be considered for analysis, but these will not be quantified. These texts will be referenced when they are mentioned. For instance, the thesis relies extensively on Xhosa dictionaries and grammar books (especially in Chapter 3) and also uses other 19th century Xhosa texts for illustrative data. Where necessary, these texts will be introduced when they are used.

2.4 Conclusion and Preliminary Analysis

The diverse corpus will be useful for examining *ke* as DM; for its historicity as found in the Bible translations, the oral narratives and some literary texts; and for establishing what are the similarities or differences between the uses and functions of *ke* are in spoken and written genres of varying levels of formality (from the conversational interviews, through the generally informal online material to the very formal legalistic register of the Constitution). From the frequency counts in the tables above, various conclusions can already be drawn.

The first conclusion to reach is that *ke* occurs in oral as well as written utterances and is not limited exclusively to one genre. However, it is on the whole more common in oral utterances than in written utterances, and even less so in written utterances of formal genres. Thus, we have the highest frequency in interview data in Table 2 being 7.42 *kes* per 100 words for Gumede, and the lowest being 0.75 in an interview with two female participants. The oral narratives in Table 5 also have a high frequency count, with the highest being 6.18 *kes* per 100 words in the narrative ‘Origins of the Xhosa’ and 2.23 *kes* per 100 words as the lowest count in ‘A Child is Born Without Limbs’. The popularity of *ke* in oral utterances is made more glaring by the fact that what is the highest frequency count of 0.94 *kes* per 100 words in the online data with the blog *Embo maXhosa* is closer to what is the lowest frequency count of *ke* per 100 words in all of the genres of spoken data (the 0.75 of the interviews). Of the online data in Table 6, the blog *Darkie’s Place* has the lowest count at 0.27 *kes* per 100 words, a figure even lower than the Constitution’s 0.35 *kes* per 100 words in Table 7.

It is hardly surprising that of the *written* data, the online material should have both the highest frequency of *kes* per 100 words (0.94) as well as the lowest (0.27) – with the Constitution falling somewhere in between (0.35). This has to do with the nature of Netspeak, which Crystal (2001: 47) suggests is ‘better seen as written language which has been pulled some way in the direction of speech than as a spoken language which has been written down’. Deumert (2006: 1) also adheres to the premise that electronic utterances are essentially a hybrid between the conventions of speech and writing:

The language used in electronic communication has been described as hybrid, showing both speech-like and writing-like features, as well as features which are unique to the digital medium, and which are, to some extent, the result of its technological restrictions.

Thus, the blogs and the *Xhosa People Rock!!!* Facebook group utterances are able to represent (and average out) the extremity of use of DMs in writing but also be able to keep the process in check by being extremely conscious of the conventions of written utterances. The next chapter will first look at the traditional, grammatical descriptions that are available for *ke* and then compare these descriptions with the functions of *ke* in the Xhosa interview data. Chapter 4 will then compare these with the functions of *ke* in the Xhosa Bible translations and in some nineteenth century literary texts.

CHAPTER 3

KE IN DICTIONARIES & GRAMMARS

& KE AS A DISCOURSE MARKER

Depending on the pragmatic constraints, a culture keeps its meaning-chains short or elaborates them at leisure.

Horst Ruthrof¹⁰

3.1 Introduction

Section 3.2 of this chapter provides an overview of the discussion of *ke* in Xhosa dictionaries and grammars from 1844-2008, and argues that these texts frequently overstate the grammatical properties of *ke*. After this survey of the dictionaries and grammars, Section 3.3 demonstrates that *ke* has features which suggest that it is better described as a DM. Substantiating examples are drawn from the interview data (listed in Table 2 in Chapter 2). These examples are evaluated against Schourup's (1999) seven DM characteristics outlined in Chapter 1.

3.2 *Ke* in Dictionaries and Grammars from 1844-2008

Xhosa grammar books and dictionaries have described *ke* in diverse ways: as an 'adverb', an 'enclitic', an 'expletive', a 'conjunction', a 'coordinative conjunction', an 'interrogative' or as a 'form word'. Table 8 below summarises the grammarians' discussions from 1844 to 2008, providing information not only on the classification suggested, but also listing the proposed translation equivalents. Boyce's (1844) pioneering grammar provided what continues to be some of the enduring translation equivalents of *ke* in English, but was ambiguous in his attempt at placing it under any conventional grammatical classification. Appleyard's grammar (1850: 255-256), on the other hand, classified *ke* (and *ke-*)¹¹, together with particles such *nje* and *bo*, as expletives – a classification also adopted by Stewart's (1921: 133) grammar.

¹⁰Ruthrof, H. 2000. *The Body in Language* (pp. 52).

¹¹Because of orthographic conventions of the time, *ke* was sometimes conjoined with words such as *kaloku* to be *kekaloku*, for instance, hence the *ke-*. Nowadays *ke* is written as an independent unit.

Table 8: Overview of Grammars and Dictionaries on *Ke*

SOURCE	TRANSLATION	FUNCTION	CLASSIFICATION
William B. Boyce; <i>Grammar of the Kaffir Language – 1st Edition</i> (1834)	Then, truly	' <i>Ke, je, bo</i> are sometimes appended to words with the sense of "then", "truly"' (p.54). The only examples given are: <i>Yininake?</i> (wherefore?), <i>Kaloku je</i> (now then) and <i>Ewe bo</i> (yes truly)	Not classified, although its description is listed in Section 8 of the grammar – a section titled 'Adverbs, Conjunctions, Prepositions'
John W. Appleyard; <i>Kafir Language</i> (1850)	And, then, ah [Other expletives listed are <i>nje, bo, and ke-</i>]	' <i>Ke</i> is affirmatory, consentive, precatory, and inferential' (p.256). ' <i>Kekaloku</i> , a derivative of <i>kaloku</i> , now, and <i>ke</i> , appears to be employed, more for the purpose of giving an opportunity for thought, or for affording relief from memory, than for adding any particular force to meaning. Its use is most prevalent in the course of a narrative or an argument.' (p.256)	Expletive
William J. Davis; <i>A Dictionary of the Kaffir Language: including the Xosa and Zulu Dialects</i> (1872)	So, well, and now, then, after what has been said, that is	'Expressive of doubt', used in deprecatory and precatory sense, used in terminating a sentence, affirmatory, consentive, inferential, asserting.	Particle
Albert Kropf; <i>Kafir-English Dictionary</i> (1899/1915)	And, now, but, then, O, and now, therefore, well, and so, so	As a conj.: 'It is used to indicate sequence in time, a progression in the chain of events, sometimes with adversative meaning' (1915: 185). 'It is sometimes used in a deprecatory and precatory sense' (ibid). As an enclitic: 'It is affirmative, consecutive and inferential, referring to what has previously taken place, or been asserted, and often answers to the English "then"' (ibid). 'It makes the verb, adjective or pronoun more emphatic' (ibid).	Conjunction; enclitic
C.J. Crawshaw; <i>A First Kafir Course – 5th Edition</i> (1903)	Not provided	'The particle <i>ke</i> added to the Imperative slightly softens it: thus <i>Hamba, Go; Hamba-ke, Go then</i> (Please go)' (Lesson X, pg. 12).	Particle

SOURCE	TRANSLATION	DESCRIPTION	CLASSIFICATION
J. McLaren; <i>A New Concise Xhosa-English Dictionary</i> (1915/1963)	Now, but, well; then, and so, then	Not provided	Conjunction; enclitic particle
James Stewart; <i>Outlines of Kafir Grammar– 3rd Edition</i> (1921)	‘These words are not easily translatable except by <i>well, then, now</i> , as may suit the expression used. They serve also to give force to a sentence as an <i>affirmation, a request, or an inference.</i> ’ (p.133)	These are considered to belong to ‘a small group of words or syllables which are true particles...which are used as expletives – that is, to fill up a clause or sentence, or introduce one; or to form a rest in the sentence, for thought or utterance; or in some cases for supposed ornament or rhetoric.’ (p.133)	Expletive (together with <i>nje</i> and <i>bo</i>)
J. McLaren; <i>A Concise English-Kafir Dictionary</i> (1923)	And, and so (p.10), now (p.190), then (p.285)	Not provided	Conjunction; adverb
W.G. Bennie; <i>A Grammar of Xhosa for the Xhosa-Speaking</i> (1939)	Not provided	‘Certain conjunctions serve merely to connect sentences, either formally or by implication, with what has gone before; they do not otherwise affect the construction. Such are <i>ke, kaloku, kodwa, kaloku</i> , (with its contraction <i>koko</i>), <i>noko, nokuba</i> ; and to this list some would add <i>gxebe, kambe, and phofu</i> referred to under the adverb’ (p.141).	Conjunction
J. McLaren; <i>A Xhosa Grammar</i> (1944)	Then, so; therefore (<i>ngoko ke</i>); and, but, so, then; now (<i>ke kaloku</i>)	Not provided	Adverb of manner (p.146); adverb of cause (p.147); co-ordinative conjunction (p.148)
Arnold Fischer; <i>English-Xhosa Dictionary</i> (1985/1994/2005)	And, then; now then (<i>ke ngoku</i>), furthermore, also, and then	‘Then’ is <i>ke</i> ‘if used at the beginning of a clause [...] e.g. “You don’t want to become a teacher, ~what do you want to be?: akufuni kubangutitshala, <i>ke</i> , ufun’ukuba yintoni na?”’ (1994: 658). Also: ‘In the continuation of a narrative or in commencing a new sentence “and” is expressed as <i>ke</i> or by verbs <i>-thi, -za</i> , and other words.’	Conjunction; adverb
S.L. Tshabe, B.M. Mini, N.V. Mkonto, E.P. Msutwana, T. Makubalo and R.H. Kaschula; <i>IsiXhosa Sezikhuthali: Ibanga 10</i> (1989)	High school Xhosa language subject grammar therefore none provided	‘U <i>ke</i> ubhalwa yedwa. Umzekelo : Hamba <i>ke, Ke kaloku ke.</i> ’(Ke is written on its own. Example : Go <i>ke, Ke kaloku ke.</i> ’ (p.115)	Enclitic

SOURCE	TRANSLATION	DESCRIPTION	CLASSIFICATION
J.A. Du Plessis and M. Visser; <i>Xhosa Syntax</i> (1992)	Therefore	Grouped with sentential prepositions that denote cause: <i>ngoko (ke)</i> and <i>ngako oko (ke)</i> . 'These two expressions function as sentential adverbs [...] As sential adverbs they have no influence on the tense or mood of the subsequent clause.' (p.204)	Sentential adverb
Via Afrika; <i>Afrikaans-Xhosa, Xhosa-Afrikaans Woordeboek</i> (1994)	Dan, nou, (en), maar, (durf) net	More like a word list than a dictionary so not provided	More like a word list than a dictionary so not provided
Pharos; <i>English-Xhosa, Xhosa-English Dictionary</i> (1998/2002)	Then, now, (and), so, but, just (you dare)	More like a word list than a dictionary so not provided	More like a word list than a dictionary so not provided
B.M. Mini (Director and Editor in Chief) and S.L. Tshabe (Editor); <i>Greater Xhosa Dictionary</i> (2003)	And, and so, to continue, well, now, then, but, on the other hand, on the contrary; and so?; and now?, well?, and so what?; well done!, thank you!, that's it!	'Form word introducing a statement which is the continuation or result of a foregoing statement or action, whether expressed, performed or anticipated' (p.16). 'The interrogative <i>ke?</i> has a falling tone and indicates an inquiry regarding the continuation or outcome of the antecedent statement or action'(ibid). 'May be coupled with other words to form conjunctive or interjective phrases' (p.17). Listing of <i>athi ke mna!</i> as an 'interjective expressing remorse, distress, self-reproach, despair; woe is me alas' (p.17). Separate listing of <i>ngoku (ke)</i> – conjunctive – trans. 'and now' and 'and then' (p.541)	Form word, conjunctive; interrogative; interjective
Philip M. Parker; <i>Webster's Xhosa-English Thesaurus Dictionary</i> (2008)	Then (synonyms) so, accordingly, afterward, again, consequently, later; (conj.) therefore	Not provided	Adverb; conjunction

Third in the chronology was Davis' (1872) dictionary that simply referred to *ke* as a particle, and it was only with Kropf's (1899) dictionary that *ke* was recorded as both a conjunction and an enclitic. Kropf's (1899) classifications were elaborated in the McLaren (1915) version of the dictionary, and again in McLaren's 1944 grammar where *ke* was listed as an adverb of manner and cause as well as a coordinative conjunction. For almost six decades thereafter until the publication of the *Greater Xhosa Dictionary* (2003) (*GXD* henceforth), these four main classifications (expletive, conjunction, enclitic, adverb) remained the dominant grammatical categories mentioned. The *GXD* (2003) introduced two more classifications: form word and interrogative. Moreover, although the *GXD* uses Kropf's (1899/1915) dictionary as a basis, the *GXD* provides a total of seventeen translation equivalents for *ke* where Kropf's dictionary only had eight. Thus, not only has there been a rise in the number of grammatical classifications for *ke* but there has also been a rise in recorded translation equivalents over time. Before delving into the possible reasons for this rise it is important to know what is understood about the terms used by the Xhosa dictionaries and grammars to classify *ke*.

'There is a longstanding tradition which says there are eight parts of speech in English: nouns, pronouns, verbs, adjectives, adverbs, prepositions, conjunctions, and interjections' (Wardhaugh, 1995: 4). If one accepts these parts of speech as being applicable to languages other than English, even if in part, then at least three of the *ke* classifications chronicled in Table 8 above are part of this basic language categorisation: adverbs, conjunctions and interjections. Doke (1976: 16) defines an interjective, as 'an isolated word which has no grammatical or concordial bearing upon the rest of the sentence,' and this definition is interesting because it contains strong leanings towards DM characteristics of optionality and non-truth conditionality.

Xhosa has numerous adverbs, some of which are proper adverbs but many of which are nouns and locative nouns used adverbially. On the other hand, conjunctions are generally not regarded as a separate part of speech in Xhosa although numerous lexical items, usually attached to nouns, can be used for conjunctive purposes as co-ordinative and subordinative conjunctions. Less common than any of the other classifications of *ke* are the terms 'enclitic', 'expletive' and 'form word'. Enclitics are words so unemphatic they are usually pronounced as part of the preceding word, whereas expletives generally serve to fill out a sentence. Again, the latter of these two terms possess strong leanings towards DM characteristics of

optionality and non-truth conditionality. On the other hand, the *GXD*'s working definition of a form word is that it is 'a complete word' that nonetheless has no 'independent meaning on its own. It is more of a functional than meaning conveying word. However it is a complete word structurally' (*GXD Vol 1*, 2006: XXIV). Crystal (2003 – entry under 'function word') defines 'form word' as a term 'sometimes used in word classification for a word whose role is largely or wholly grammatical, e.g. articles, pronouns, conjunctions.' Unlike lexical words, Crystal notes that form words do not carry the main semantic content. *A Dictionary of Sociolinguistics* (Swann, et al., 2004) defines 'form word' as a 'grammatical word', and explains that this term can be applied to articles, conjunctions, some prepositions, and adverbs that have 'a primarily grammatical function, linking other words together.' Therefore, a form word is a term that encompasses articles, pronouns, conjunctions, adverbs, and some prepositions. In short: it is a complete word structurally, usually has grammatical function, but carries no content meaning.

If one looks at the Xhosa dictionary and grammar description of *ke*, it is clear that in the way they are phrased most of them foreground the grammaticality of *ke*. However, most of the dictionaries and grammars also provide descriptions that seem to correspond with DM features and functions. For instance, there are several descriptions that can be read as accounts of the 'connectivity', 'optionality' and 'non-truth conditionality' characteristics of almost all DMs as discussed in Chapter 1. Some of the descriptions are very aware of what could be the DM functions of *ke* even if never explicit with the precise linguistic terminology provided by a term like DM. For example, Bennie (1939: 141) writes that *ke* is part of a certain group of conjunctions that 'serve merely to connect sentences, either formally or by implication, with what has gone before; they do not otherwise affect the construction' of an utterance. The first part of this description encapsulates the connectivity DM feature and the interpretation of the second half of his description accommodates the optionality, non-truth conditionality and weak clause association DM features outlined in Chapter 1. For Stewart (1921), *ke* is part of a group of particles that serve to 'fill up a clause or sentence, or introduce one; or to form a rest in the sentence, for thought or utterance; or in some cases for supposed ornament or rhetoric,' and this description also accommodates at least three of the seven DM characteristics (optionality, initiality, and orality) as well as what could be two DM functions ('rest in the sentence' and 'ornament or rhetoric').

Therefore, according to the various dictionaries and grammars, *ke* is a ‘form word’, an ‘adverb’, ‘enclitic’, an ‘expletive’, a ‘conjunction’, a ‘coordinative conjunction’, an ‘interrogative’ and an ‘interjective’. All these categories are standard, traditional grammatical categories. However, the steady rise of translation equivalents for *ke* from the earliest dictionaries and grammars to the latest *GXD* (2003) highlights the realisation that despite attempts to delimit the word within the well-organised and well-governed system of grammatical descriptions, it keeps expanding, pragmatically, and more precise ways at description are increasingly necessary. The remainder of this chapter will investigate these implicit DM features present in the dictionary and grammar descriptions of *ke*, and will analyse the *ke* tokens in the interview data, oral narratives and online materials against the seven DM characteristics outlined in Chapter 1 to demonstrate the necessity of also classifying *ke* more appropriately as a DM.

3.3 *Ke* as a DM in contemporary texts

There are two references to *ke* as a DM that I am aware of: in Gough (1986) and De Klerk (2005b). The earliest reference to *ke* as a DM is from Gough’s (1986) doctoral thesis, *Xhosa Narrative: an analysis of the production and linguistic properties of discourse with particular reference to iintsomi texts*. Gough does not use the term DM, but uses discourse particle (DP), which is the term Schourup (1999: 228-9) stated was the predominant term for DMs until the mid-1980s. Gough (1986: 73) argues that in *iintsomi* (Xhosa folk/tales) there are ‘certain lexical items and phrases whose function in discourse is not to add content as such, but to indicate a change in conceptual re-orientation’. He cites clause-initial adverbials such as *ngoku* (now) and *ngokwenene* (in truth) when ‘occurring with the enclitic *ke*’ as the most common of these lexical items (ibid). ‘Perhaps the most common discourse particle,’ he continues, ‘is *hayi ke* (‘no then’), from the interjective *hayi*’ (ibid).

Later in the argument and borrowing from Gumperz’s (1977) work, Gough (1986: 203) calls these DPs ‘contextualisation cues’, which he explains as ‘surface features of utterances that are designed to invoke in the mind of the hearer a particular frame of interpretation.’ Upon hearing these cues, the hearer ‘will be in a position to infer the intended implications of the speaker’s utterances without ambiguity’ (ibid). He identifies *hayi ke* (‘no then’) and two instances when *ke* appears with clause-initial adverbials such *ngoku ke* (‘now then’) and

ngokunene ke ('in truth then') as three examples of Xhosa DPs. He claims that these and other Xhosa DPs 'can be seen to have the sole function of indicating breaks of relevance in a stretch of discourse' (ibid).

In De Klerk's (2005b) research on the procedural meanings of *well* in a corpus of what she calls Xhosa English, *ke* is listed as a possible DM. De Klerk points out the apparent absence in Xhosa of any lexical or pragmatic equivalent to DM *well*, thus confirming Fraser's (1990: 395) speculation that *well* and *anyway* are two DMs less likely to occur in all languages. De Klerk (2005b: 1200) then reports that:

Enquiries among several expert MT [mother-tongue] Xhosa speakers confirmed that such a marker is basically absent in Xhosa, and that the nearest equivalents are the words *ngathi* ('like'), *ke* ('what's next') and *eeh* (a signal of contemplation).

In both Gough (1986) and De Klerk (2005b), the claim that *ke* is a DM is stated without sufficient substantiation, more so in the latter's work where her conclusion is based on native speaker informants and not on analysis of any data. Although the informants were obviously not incorrect – five of the dictionaries and grammars listed in Table 8 above provide *well* as one of the English translation equivalents for *ke* – this section of the thesis rectifies this oversight by looking at *ke* in the context of data and analysing the validity of these claims against the seven DM features identified by Schourup (1999) and discussed in Chapter 1. The features are connectivity, optionality, non-truth conditionality, weak clause association, initiality, orality, and multicategoriality.

Using these seven characteristics and illustrated with examples drawn from of the interview data outlined in Chapter 2, this rest of this section shows why it is valid to claim that *ke* is a DM. The section will demonstrate that *ke* has the same characteristics or features that are common to other DMs, and also shares several *functions* with many DMs in other languages found in the DM literature. *Ke* has at least nine DM functions illustrated alongside the discussion of the seven DM characteristics. Some of the ways in which *ke* is used are:

- (1) To create textual coherence
- (2) To mark the staging points or transitions in an unfolding narrative to aid in building up to the resolution of an utterance
- (3) To emphasise a point in an utterance or intensify its force

- (4) To hedge or downplay one's contribution
- (5) To signal hesitation
- (6) To indicate contrast to some aspect of the foregoing utterance
- (7) To introduce modification of an earlier utterance
- (8) To frame an aside or a change in topic
- (9) To indicate conclusion or coda to an utterance

After this section has shown that *ke* is a DM and having touched on its functions in utterances, Chapter 4 will then show that based on these features and functions *ke* is not only a DM, but also an old one, already appearing in Xhosa texts of the mid-nineteenth century.

3.3.1 Connectivity

Connectivity is the feature of DMs to relate utterances or discourse units to one another (Schourup, 1999: 230). This is illustrated in examples (11) and (12) below.

11. Zezonto **ke** ezo zizakusilahlekis'ingqondo silibale ngokugoduka.

Z- ezo- nto **ke** ezo zi- zaku- si-lahlekis'
 COP DEM3 things **ke** DEM2 they IMFUT to us distract
 ingqondo si- libal- e ngo- kugoduka
 brain we forget SUBJUNC about going home

*Those **ke** are the things that will distract our brains, and make us forget about going home.*

(Ngomane, male, 60)

12. Ndada ndayeka **ke**; andisafuni ngoku

Nd- a- da nda- yeka **ke** a- ndi- sa- fun- i ngoku
 I RP eventually I stop **ke** NEG I no longer want NEG now

*I eventually stopped **ke**, I no longer want to now*

(Nilson, male, 49)

In example (11), Ngomane points out that the pleasures of city life are rather distracting to the mind (as opposed to the relative slowness and stability of the rural lifestyle left behind by migrants), and that these distractions cause people to forget about returning to their rural homesteads. *Ke* serves to link the aforementioned types of pleasures listed by the interviewer

(‘those things *ke*’, i.e. going to the tavern and getting drunk) to the interpretation that the possible consequence of following those hedonistic pleasures is a loss of identity. In example (12), Nilson also uses *ke* to link his previous utterances about the challenges he faced in his casual work history with this statement of his eventual decision to stop working. Like other DMs performing a similar function, *ke* links the assumptions that underlie two or more utterances or the cause and effects of particular states in order ‘to create coherence within a speaker’s turn or signal the relationship between one speaker’s utterance and another’s response’ (Fuller, 2003: 24). As already mentioned, connectivity is usually cited as the most common of DM functions.

3.3.2 Optionality

Optionality refers to the fact that the grammaticality of a sentence is not affected by the presence or absence of a DM. This is illustrated in (13) and (14).

13. Ja, nendlela **ke** outie endiziphethe ngayo iyabonisa ukuba ndiyafika...
 Ja, ne- ndlela **ke** outie e- ndi- zi- phethe nga- yo
 Yes, and way **ke** bra DEM2 I myself carry INSTR it
 i- ya- bonisa ukuba ndi- ya- fika.
 it PRES+DISJUNC shows that I PRES+DISJ arrived.
*Yes, and the way **ke** bra that I carry myself shows that I have just arrived...*
 (Nceba, male, 26)

14. Nase-cherrini **ke** iyabonakala.
 Na- se- cherri- ini **ke** i- ya- bonakala.
 And LOC chick LOC **ke** he PRES+DISJUNC is apparent
*Also with a chick **ke** he can be spotted.*
 (Sandile, male, 26)

Examples (13) and (14) are responses about how to spot a newcomer in the township. In example (13), Nceba, who at the time of the interview had been in the township for less than a month, explains that the way he carries himself communicates to others that he has just arrived in the township. In example (14), Sandile is commenting that the way a guy interacts with a female could also be an indicator of whether he is new or not. One of the functions *ke* performs in this example is to highlight the relations between the two parts of the utterance.

Both (13) and (14) would not be ungrammatical if *ke* was omitted. This also holds true for examples (12) and (13), as well as subsequent examples in this chapter. This optionality feature of DMs is their second most frequently cited attribute and *ke* also accords with it.

3.3.3 Non-truth conditionality

Non-truth conditionality is the third most common DM characteristic. It expresses the idea that unlike content words, DMs do not affect the propositional truth or falsity of an utterance as illustrated in examples (15) and (16) below.

15. Apha elokishini kuthethwa kakhulu isiXhosa apha **ke** kule ndawo yethu.
 Apha e- lokish- ini ku- thethwa kakhulu isiXhosa apha **ke** ku- le ndawo
 Here LOC township LOC there-is spoken a great deal Xhosa here **ke** LOC DEM1 place
 ye- thu.
 POSS+our
*Here in the township, Xhosa is mostly spoken **ke** here in our township.*
 (Gumede, female, 66)

16. Bayayenza **ke** abanye lonto
 Ba- ya- y- enza **ke** abanye lo- nto
 They PRES+DISJUNC it do **ke** others DEM2 thing
*Others do **ke** that thing.*
 (Nontokozo, female, 30s)

In example (15), Gumede states that Xhosa is the main language used in Gugulethu, and she uses *ke* to emphasise a sense of locality. This function of *ke* as an emphasis marker is common to DMs in general. Nontokozo in example (16), on the other hand, uses *ke* to hedge the practice of performing traditional rituals in the city instead of going back to the Eastern Cape. She believes the latter should have been maintained as the norm. In a way, her *ke* carries with it a certain sense of disapproval at the fact that there are people who *do* perform Xhosa traditional rituals in the township and not, as she thinks is appropriate, at their homestead of origin in the Eastern Cape. *Ke* performs the important DM function of hedging/softening, but it does not affect the truth or falsity of Nontokozo's utterance. As with all the previous examples, *ke* is optional in examples (15) and (16) and because of this, it affects neither the grammar nor the truth or falsity of the utterances in which it occurs.

3.3.4 Weak clause association

Weak clause association, and subsequent other DM characteristics are not as widely accepted as the above three features. Weak clause association refers to the fact that DMs are clause elements that are loosely attached to, or completely outside, the syntactic structure of an utterance. This characteristic is also often correlated to phonological independence in that DMs are said to constitute ‘independent tone units’, or separated from the main clause by ‘comma intonation’ (Schourup, 1999: 233). Even though *ke* is syntactically very flexible, this characteristic is probably not applicable to *ke* as a DM because of its excessive attachment to or co-occurrence with a certain set of words or types of phrases. This co-occurrence with or dependence on other parts of speech is so prevalent to an extent that in older orthographies, for example, there are even instances where *ke ngoku* (‘so, then, now, so now, then now, so then, therefore’ – depending on the context) is spelt as one word. Phonologically independent DM *ke* is exceedingly rare in spoken utterances.

However, we can have *ke* occurring as an independent tone unit. A hypothetical example of this could be someone saying ‘Ke?’ (said in a specific context to mean *Well? And so? And then? And now? And so what?*), maybe after someone has recounted a seemingly pointless or startling tale or tried justifying something that the questioner deems inexcusable. A says: ‘Sorry I am late’, B responds: ‘Ke?’ (*And so what am I suppose to do about that?*). However, I have no example of a lone standing interrogative *ke* in my data. From personal experience, *Ke ngoku?* or *Pho ke?* are the preferred (less ambiguous) options for use in such instances.

What is important to note is that *ke* is (a) not usually an independent intonational unit, but that (b) it *can* occur as an intonational unit and even stand alone (in which case it is a bit like *well?*; i.e. still a DM). Anyway, regardless of whether the weak clause association feature exists for *ke* as a DM or not, that still does not affect the main argument that *ke* is mainly a DM because it already in accord with the three most important features of DMs listed above.

3.3.5 Initiality

DMs tend to introduce the segments that they mark in an utterance. This refers to initiality in relation to the central clause elements in an utterance and not on whether the DM is the first word of that utterance or not (see discussion of the initiality characteristic in Chapter 1, subsection 1.4.2 for further clarification). This is illustrated in example (17) below.

17. **Ke** ungasixelela ngapha ekhaya, e... emva... Ubukhe wasebenza emaXhoseni?

Ke u- nga- si- xelela nga- pha- a e- khaya, e... e- mva... U- b- u- kh-
Ke you COND us tell LOC there DEM3LOC home, LOC... LOC back... You RT you ever
ewa- sebenza e- maXhos- eni?
PST+DISJ PST+SUBJUNC worked LOC Xhosa place LOC

***Ke** can you tell us: back home at... back there... did you ever work in the Eastern Cape [at the place of the Xhosa people]?*

(Zolani, male, 32)

In example (17) *ke* introduces clauses in utterances. *Ke* prefaces the interviewer's request to probe further into the research participant's work history. The *ke* is at once an indicator of a shift to a new topic as well as a politeness marker to soften the potential offence at digging into the participant's life history. This, however, is a function *ke* rarely fulfills independently since it is often coupled with a conjunctive introducing the new clauses in complex utterances and in those cases *ke* then serves to reinforce that conjunctive. Although in example (17) *ke* is actually at the beginning of the sentence and of the clause, it is a very flexible particle and syntactically it can occur at almost any point in an utterance, including at the end of a clause (as seen in example 18 below, for instance).

18. Yenzek' olohlobo **ke**...

Yenzek' olo- hlobo **ke**...
It happens DEM2 way **ke**...
*It happens that way **ke**...*

(Mandla, male, 28)

3.3.6 Orality

Orality is the assumption that DMs are more frequent in speech than in writing, although, as discussed in Chapter 1, this is contested. The demarcation and distribution of DMs between speech and writing is not rigid, and often it has to do with the formality or informality of the genre in question. *Ke*, however, is definitely more frequent in oral utterances than written ones as already discussed in Section 2.4 of Chapter 2.

19. **M:** Uyabona **ke** leyo –
 U- ya- bona **ke** leyo –
 You PRES+DISJUNC see **ke** DEM2
*You see **ke** that –*

N: Kakhulu!

Much!

M: Banintsi **ke** leyo –

Ba- nintsi **ke** leyo –

There many **ke** DEM2

*There are many **ke** [with] that.*

(Mzolisi, male, 36 and Nontokozo, female, 30s; shopowners)

Example (19) is a brief exchange which occurred after the interviewer asked whether the interviewees can spot if someone is new in the township simply based on what that person buys from the shop. The use of the second *ke* is typical of spoken language. In (formal) written Xhosa it would be a grammatically awkward utterance at best or deemed incorrect at worst (even though it would still make sense). In this instance, *ke* functions as a repairing particle or filler in order to maintain a sense of coherence in an otherwise disjointed construction. Such utterance constructions, characterised by false starts, are typical in conversation and unlikely in utterances written in standard Xhosa. Such utterance constructions and the role *ke* fulfils in them are to be expected in spoken data where a lot of conversational management is necessary and ‘false starts’ constantly have to be repaired or coherence reestablished, and *ke* is a particle well-suited to such a function because of its syntactic flexibility, allowing it to fall almost anywhere in a sentence.

3.3.7 Multi-categoriality

Linked to that last point, is multi-categoriality, which – as the seventh DM characteristic – emphasises that ‘DM status is independent of syntactic categorization; an item retains its non-DM syntactic categorization but does “extra duty” as a non-truth-conditional connective loosely associated with clause structure’ (Schourup, 1999: 234). For example, adverbs, coordinating and subordinating conjunctions, interjectives, verbs and phrases (such as *you see, I mean, you know*) can all take DM functions without necessarily ‘losing’ their syntactic functions. However, I am not convinced that *ke* is ever grammatical (and again, I will illustrate why in the following chapter) except in instances when it is a lone standing

interrogative, as I have already illustrated. However, when *ke* is translated into English (and, I suppose, into other languages as well) it does translate into what are the adverbial and conjunctive equivalents, such as *then, and, therefore*, etc; basically, the list of the translation equivalents listed in the dictionaries and grammars. However, even though *ke* has similarities to these adverbial and conjunctive equivalents (i.e. grammatical particles), they function, in Xhosa, as DMs and not as grammatical particles. For instance, in all the examples provided above, *ke* can be translated into English equivalents such as *as well* or *also* (13) *also* (14), *now* (17), *then* (11, 12 and 18), and even perhaps *yes/yeah* (15 and 16) and *with* (19), although the last two instances are more difficult to translate. But this does not mean that these equivalents function in the same manner as they do in their grammatical functions, but work as pragmatized particles.

20. Yiyo **ke** ekuthatha ikugqithisele kuwo.

Yi- yo **ke** e- ku- thatha i- ku- gqithisel- e ku- wo.

COP it **ke** REL+it you takes it you pass on SUBJUNC LOC them (police)

*It's it **ke** that takes you and passes you on to them [the police].*

(Nilson, male, 49)

21. Andazi **ke** ngeli xesha ndilapha ngoku sendingaka, kodwa ngela xesha lethu, ubunqabil'umsebenzi.

A- nd- az- i **ke** ng- eli xesha ndi- lapha ngoku se- ndi- ngaka

NEG I know NEG **ke** at DEM1 time I+SITU DEM1(here) now I old

Kodwa nge- la xesha le- thu u- bu- nqabil'- umsebenzi.

but REMPST DEM3 time POSS our it REMPST scarce work

*I don't know **ke** about the period I've been here, now at this age, but during our time [when I was younger] work was scarce.*

(Gumede, female, 66)

In examples (20) and (21), *ke* could possibly be translated as 'then' in both instances or as 'then' in (20) and 'however' in (21). In both cases, 'then' and/or 'however' do not fulfill significant grammatical functions (in both the Xhosa original and in the translation).

However, it is clear that these grammatical functions are not necessary. For instance, the *ke* in example (21) provides pragmatic, discourse-structuring functions such as indicating possible contrast between the availability of jobs in the Eastern Cape when Gulwa was a young

woman compared to the situation now. This contrast is a modification of an earlier utterance, and *ke* introduces and emphasises it. The emphasis on the possibility of error is simultaneously a hedging device, thus demonstrated a very rich use of *ke* functions as a DM.

3.4 Conclusion

In short, based on these seven characteristics, it is clear that *ke* is a DM in contemporary Xhosa, even though the prevailing claims for it in Xhosa grammars and dictionaries have emphasised its grammatical nature. This preference for grammatical approaches is not surprising; a language gets its status through the standard variety, and the standard variety is determined by grammatical structures that are consistent. Lexicographers, grammarians and dictionary makers become, by default, the creators, providers or guardians of that standard and as such tend to augment grammaticality rather than ‘fuzzy’, pragmatic or nonstandard or inconsistent uses of words.

The challenge of isolating the ‘core meaning’ of DMs has been discussed in Chapter 1, and in this chapter it is clear that *ke* has more or less the same challenges. To recap, Schourup (1999: 254-5) pointed out that

... the assumption underlying such arguments is that the DM in question has a unitary core meaning, whereas it is not clear what criteria can be used to decide when a unitary core should be forsaken for polysemy.

In other words, DMs that might be carrying a myriad of meanings can be forced into very limiting meaning-categories. In the domain of grammatical meaning-making the inclination to delimit multiple meanings and functions is generally much stronger. As a result of this core-meaning seeking activity, Schourup (1999: 255) (referring to researchers working on DMs) warns that ‘genuinely distinct cores might mistakenly be collapsed; on the other hand.... collapsing uses can result in a core which is so vague that it becomes almost impossible to evaluate in relation to actual data.’ The analysis above has demonstrated the importance of testing the descriptions in grammar books against actual use in data.

CHAPTER 4

KE AS A DISCOURSE MARKER IN HISTORICAL TEXTS

4.1 Introduction

The seven DM features identified by Schourup (1999) in Chapter 1 are connectivity, optionality, non-truth conditionality, weak clause association, initiality, orality, and multicategoriality. In that chapter, I defined the characteristics and illustrated them with examples from the DM literature. Using these seven characteristics as a framework, the previous chapter argued that *ke* is a DM. This chapter then proceeds to show that *ke* is not only a DM, but also an old DM, already appearing in Xhosa texts of the mid-19th century and even then already fulfilling a variety of functions. This chapter has two main sections. Section 4.2 looks at the three important DM characteristics (connectivity, optionality and non-truth conditionality), using fiction and non-fiction Xhosa prose texts published from 1839 to 1914 to illustrate that *ke* is an old DM. Section 4.3 analyses *ke* functioning as a DM in the Xhosa translations of the Biblical narrative of the prodigal son in the Gospel of Luke 15: 11-32. The discussion draws attention to the fact that there has always been an inconsistency not just of its function but also of the interpretation of its meanings even in older texts.

4.2 Analysing *ke* in mid-19th century prose texts

Very early in A.C. Jordan's *Ingqumbo Yeminyanya* (1940), a character utters the following:

22. Nisavukile ke, MaMiya?
Ni- sa- vukil- e ke, MaMiya?
You+PL still awake+PRS+SUBJUNC ke, MaMiya

(A.C. Jordan, *Ingqumbo Yeminyanya*, 1940)

This *ke* is one of the very first three used in the novel as soon as dialogue begins. Prior to this, the third person narrating voice does not use any *ke* (subsequently it does), thus illustrating the quick association of *ke* with conversational utterances and its ample employment in

written narratives as a means to indicate the ‘orality’ that the text aims to convey to the readers. This is corroborated by the higher frequency counts of *ke* in oral as opposed to written quantified data discussed in Section 2.4 in Chapter 2. The novel then proceeds to use *ke* liberally and for every conceivable purpose and function. The first chapter alone (roughly seven and a quarter pages long, depending on edition) already has 31 *ke* tokens – employed by male and female characters as well as the narrative voice. This is in 1940. Yet, in an even earlier classic Xhosa novel, we find:

23. Sibenjalo **ke** isiphelo setyala lamawele.

Si- be- njal- o ke isi- phelo se- tyala la- ma- wele.

It PSTCONT+COP like DEM3 **ke** the end of crime POSS the twins

*That is how **ke** the end of the trial of the twin’s was.*

(S.E.K. Mqhayi, *Ityala lamawele*, 1914)

In the example (23) above from Mqhayi’s *Ityala Lamawele* (1914), *ke* signifies a coda, the end of a narrative, a conclusion to which the story has been building up, and so it helps to emphasise the note of finality. In a way, it is also working as a connecting device by showing (and emphasising) the link between all that has come before and the final consequence of that prior action, the whole novel utterance. As a result, this *ke* is highly significant in terms of its impact in the narrative structure. It fulfils a number of discourse functions in one instance, but also, it can be translatable into a ‘then’, therefore making it multi-categorical. However, despite all these significant functions *ke* is still optional.

Although these two examples already indicate that *ke* is an old DM, this section shows that it is actually even older than 1914 by analysing examples of *ke* already functioning as a DM in prose texts as old as 1839. The use of the DM *ke* in prose should be unsurprising, given prose’s wide-ranging experiments with words and *ke*’s own flexibility to meet the demands that a writer might need to fulfil at any given point in an utterance. So of course writers have used it extensively and for a variety of reasons. Examples cited in this section are taken from W. G Bennie’s *Imibengo* (1935) and Jeff Opland and P T Mtuze’s *Izwi Labantu* (1994), two standard anthologies of Xhosa fiction and non-fiction prose and poetry through the ages. Some of the texts in these anthologies are the same texts that would have been consulted during the process of compiling the dictionary and grammar entries outlined in Chapter 3 (Section 3.2). The selected examples are illustrative and only used qualitatively.

Connectivity

24. Baya **ke**, bafika bangena ebuhlanti ebusuku batya **ke** abanumzana...

Ba- ya **ke**, ba- fika ba- ngena ebu- hlanti ebu- suku ba- ty a **ke** abanumzana...

They went **ke**, they arrived, they entered LOC kraal LOC night they ate **ke** the gentlemen...

*They went **ke**, arrived, [and] went into the pan at night [and] **ke** the gentlemen ate*

(St Peter's Mission, *UDyakalashe nengcuka*, 1875)

In this example of the adventures of two tricksters stealing a farmer's sheep, the two *kes* connect different states of action in order to create narrative coherence and to show continuity in the progression within the narrative. However, compare example (24) with this extract (25) below which occurs a few lines further into the story:

25. Wathi uDyakalashe, 'Wethu hamba siye kukhwela amahashe amaBhulu.'

Bahamba, bafika, bawakhwela.

Wa- thi u- Dyakalashe, 'Wethu hamba si- ye ku- khwela

PST+SUBJUNC say the jackal 'My dear/brother PRS+go we IMMFUT+go LOC ride

ama- hashe ama- Bhulu.' Ba- hamba, ba- fika, ba- wa- khwela.

the horses of the Boers.' They went, they arrived, they PST+SUBJUNC rode.

And the jackal said, 'My dear/brother, hurry let's go ride the Boer's horses.' They went, arrived, [and] rode them.

(St Peter's Mission, *UDyakalashe nengcuka*, 1875)

In example (25), a construction similar to the opening line in example (24) is repeated. In example (24) it is *Baya ke, bafika, bangena* ('they went **ke**, they arrived, [and] they went in') and in (25) it is *Bahamba, bafika, bawakhwela* ('they went, they arrived, [and] they rode them'). But note the difference: *ke*'s connectivity function of linking the previous utterance to what follows is clearly not necessary because in (25) there is no *ke* whereas the coherence and narrative progression are still maintained. Here, then, *ke* is working in the same fashion as it does in the oral narratives where it 'flavours' the utterance to give it an oral, loose feel as well as emphasize the passing or progression of narrative time. This progression of narrative time tied to connectivity functions is also visible in this extract below, where each *ke* indicates different segments of time passing:

26. **Ke** kaloku, Nkosi, u de wa fika **ke** ekaya lo'mtu; ku te **ke** gelinye ixesha, kwe suka abantu ba sem-Gazi, ba ya kwa-Ncapayi, be hambela into zabo.

(Jivashe, *Iindaba zokucholwa komntu oyiMfecane*, 1839)

Non-truth conditionality

The usually interjective *Hayi ke!* has a rather special place in the history of Xhosa, as it is one of the first forms of *ke* utterances to appear in written Xhosa. *Umshumayeli Wendaba* was a pioneer newspaper among the Xhosas and its aim was to ‘explain the ways of the white people to the black converts’ (Opland and Mtuze, 1994: 68). Example (26) above is also an extract from a piece by Jivashe that was meant to be published in *Umshumayeli* but never was, although that piece is officially the first piece of writing by a Xhosa in Xhosa that is known (Opland and Mtuze, 1994: 71). However, in its fourth edition the newspaper featured the first pieces of writing by Xhosa speakers in a form of four letters written by students at Moley mission (cf. Opland, 2003: 9). This was followed a few months later by a transcribed conversation between two boys on the benefits of education, writing and printing. The two conversants in this piece (27) are simply referred to as *Um-Fundi* (‘Learner’) and *Um-Nqeni* (‘Lazy’) and the dialogue begin as follows:

27. Um-Nqeni: Wa, 'M-Fundi!
Hi Learner!
- Um-Fundi: Yo!
Hi!
- Um-Nqeni: Ka u ze apha.
Please you come here.
Please come here.
- Um-Fundi: U di funela ni?
You I want why?
What do you want me for?
- Um-Nqeni: Ke- si ye'ku ghlala e-sibayeni paya.
Let us go to play at kraal there
Let's go play there at the kraal.
- Um-Fundi: Hayi **ke**, 'M-Nqeni di ya e-sikolweni mina.
No **ke**, Lazy I go to-school me
No ke/Oh no, Lazy, me I'm going to school

(Umshumayeli Wendaba, *Isincokolo samakkwenkwe amabini ngako ukufunda*, 1839)

This example is particularly appropriate to illustrate the ‘orality’ of *ke* (a DM feature) because even though this is written material it tries to emulate spoken discourse by using the popular DM expression *hayi ke*. There are four other DM *kes* in the rest of the dialogue that aid to maintain, amongst other things, the dialogue’s ‘orality’ aspect, but none of them contribute to the truth-conditionality of the dialogue. The truth or falsity of example (28) below is also unaffected by the presence of *ke*. Once again, both examples include a *ke* that is

not grammatically necessary (optionality condition), and both utterances retain their propositional meaning even if *ke* were absent:

28. ‘Wenz’ukuba usuke wathi abantwana bakho ndibadlile.’ Wasuka wahamba **ke** wemka.
 We- nz’ukuba u- suke wa thi abantwana ba- kho ndi- ba dlil- e.’
 You do because you RECPSTPST+SUBJUNC say children POSS your I ate SUBJUNC
 Wa- suka wa- hamba **ke** we- mka.
 She moved she walked **ke** she left.
 ‘*You do so because you’ve said I ate your children.*’ *She then walked away ke left.*
 (Elsie Umto, *UKhenkebe noHlakanyana*, 1877)

Optionality

The above examples clearly show that *ke* is also optional in these older texts. In another text, published in 1887, *ke* is also fulfilling an optional role as seen in example (29) below:

29. Oko **ke** kukuthi sekubonakala imisebenzi yakhe yokugaqa ehlalele abanye abantu ngasese ebaxoka, ebahleba, esekeleze inzuzo yakhe kuphela...
 Oko **ke** ku- ku- thi seku- bonakala imisebenzi ya- khe yoku- gaqa ehlalele
 DEM2 **ke** COP LOC say PRSCONT- is apparent works POSS their of always talking
 abanye abantu nga- sese eba- xoka, eba- hleba, e- sekeleze inzuzo
 DEM2+other other people LOC behind lying about gossiping LOC seeking gain
 ya- khe kuphela
 POSS his only
That ke means the revelation of a person’s works of always talking about others behind their backs, lying about them, gossiping about them, only seeking his/her gain...
 (W.W. Gqoba, *Iintsingiselo zamaqhalo esiXhosa*, 1887)

Example (29) above is an explanation of what the idioms ‘Lunyawo lweMfene’ and ‘Lunyawo likaJanya’ mean. Once again, *ke* serves as a connective marker, linking the stated idiom to the explanation that follows it. *Ke* is optional, and the sentence would still be grammatical without it. In example (30) below, *ke* occurs at the start of an utterance or clause to introduce segments that build up on what has been said. Remember, however, that although in these example *ke* appears at the beginning of the sentence, this does not

necessarily need to be the case (as is the case, for instance, in example 29 above, where *ke* is facilitating a similar function even though it is not at the beginning of a sentence):

30. **Ke** kaloku ngelinye ilixa abehamba ngalo endle, ethetha nencwadi yakhe, ngesiko lakhe, ndambona ebhandezelekile kakhulu emxhelweni.

Ke kaloku nge- linye ilixa abe- hamba nga- lo e- ndle, ethetha
Ke now in another time He+PST+SUNJUNC walking in DEM2 LOC wild speaking
 ne- ncwadi ya- khe, nge- siko la- khe, nd- a- m- bona ebhandezelekile
 with book POSS his, about belief POSS his, I REMPST him see deep
 kakhulu em- xhelweni.
 too much LOC thought.

Ke now in other times he wondered in the wild, speaking with his book, about his culture/belief, I saw him deep in thought.

(John Bunyan, Tiyo Soga (trans.), *Uhmambo Lomhambi*, 1867)

4.3 Analysing *ke* in the Bible

The Biblical story of the return of the prodigal son is found in the New Testament in the Gospel of Luke (15: 11-32). It is a parable that Jesus relates to his disciples to demonstrate God's attitude of love and forgiveness, especially toward those who stray furthest from what has been ordained. It is a story of a son who asks his father to divide his estate between him and his brother so that he can take his share of the estate, convert it to cash, and travel to a distant country to spend it there. Over time, he squanders his wealth, and when a famine hits the foreign land he is reduced to being a farmhand, looking after pigs and with no food to eat, 'glad to fill his belly with the pods that the pigs were eating'. He realises that his father's servants eat better he and eventually he decides to go back to his father to apologise and offer himself as one his servants (so unworthy from disgrace is he to remain his son) rather than to continue suffering in the new land. When he arrives home his father is so overwhelmed to see him he robes him with his finest cloak, puts a ring on his finger and slaughters for him the fattened calf to celebrate his return. The brother who had stayed home, however, is not pleased with any of this and says as much, but the father reassures him that celebrating this son's return is a necessary event because he was like one dead who has now come to life, lost and has now been found, whereas he has always been with him and everything the father has also been his.

As seen in Table 7 (in Chapter 2), the frequency counts of the *ke* in the Biblical translations of this narrative fluctuate (there are 9 in 1853 but 16 by 1942; in between these years the various editions had either 14 or 15 *ke* tokens). This already demonstrates the inconsistency of its use. However, even more important than the frequency counts of the *kes* is the shifting positions in the narrative that the different translators use *ke*. For instance, in the 1853 translation, the first use of *ke* is in the 17th verse whereas in the 1864 translation *ke* already appears in the 13th verse. Moreover, although the second translation also has *ke* in the 17th verse, its use is completely different in both.

In the English New International Version (1984) Bible translation, the 17th verse goes: ‘When he came to his senses, he said, “How many of my father’s hired men have food to spare, and here I am starving to death!”’ This verse in the 1853 Xhosa translation (31) is rendered thus:

31. Waza, akuba nokuziqondela, wati: Ukuba baninzi *ke* kwabaqashwa bakabawo abakutya kwaneleyo kwada kwasala, kanti mina ndiyabuba 'kulamba!

In example (31) *ke* is in the second clause of the utterance, but in the 1864 Xhosa translation of the same verse (32), *ke* not only appears at the beginning of the first clause of the utterance, but is now attached to *kaloku* and is absent from the second clause of the utterance:

32. *Kekaloku* kwakubuya ukuqonda kwake, wati: Baninzi kangakanani abaqeshwa bakabawo abanokutya okungapezu kwokubaneleyo, kanti mina ndiyabuba 'kulamba!

The same verse in the 1879 Xhosa translation (33) also starts with *ke* attached to *kaloku* but unlike the 1864 translation in (32), this version has second *ke* but not appearing in the same position as the *ke* in the 1853 translation (31) but in the third and final clause of the utterance:

33. *Kekaloku* akuba nokuziqonda, wati, Kanene bangakanani abaqeshwa bakabawo, abadikwa zizonka, nanku *ke* mna ndiphela yindlala?

These three variations of translating this verse remain basically consistent through the remaining translations, with the slight change of *nanku ke* being hyphenated as *nanku-ke* in the 1931 translation. This inconsistency of positions taken by *ke* within utterances and clauses (and therefore functions) in the various translations of the narrative of the prodigal son also

extends to other verses of this narrative. Basically, what this shows is that the grammaticality and seeming stability/consistency of function of *ke* in Xhosa dictionaries and grammars are perhaps overstated or at least consideration of *ke* primarily as a DM is important for better understanding its role within utterances.

4.4 Conclusion

Using examples from contemporary interview data, Chapter 3 showed that *ke* is a DM in contemporary discourse. It fulfils at least six of the seven DM features as outlined by Schourup (1999), including and especially the three most important ones (connectivity, optionality, and non-truth conditionality). Furthermore, using selected examples taken from mid 19th and early 20th century Xhosa texts, in this chapter I illustrated that *ke* is an old DM, already present in texts such as these. Based on the discussion of the narrative of the prodigal son as it appears in different Xhosa translations of Chapter 15 of the Gospel of Luke in the Bible, this chapter has also argued that the grammaticality of *ke* in Xhosa dictionaries and grammars is overstated and not particularly supported by data even from old written Xhosa texts of varying levels of informality.

CONCLUSION & RECOMMENDATIONS FOR FUTURE RESEARCH

A philosophical statement is the product of a certain personality living at a certain time in a certain place, and not the outcome of a purely logical and impersonal procedure. To that extent it is chiefly subjective; whether it has an objective validity or not depends on whether there are few or many people who argue in the same way.

C. G. Jung¹²

This thesis has shown that the Xhosa *ke* is (1) a discourse marker, (2) an old discourse marker, already present in 19th century Xhosa prose, and (3) a discourse marker that is more common in oral utterances than in written ones, even though it does possess DM features and performs DM functions in both oral and written utterances. The first and third of these findings were already anticipated in the work of Gough (1987) and De Klerk (2005b), but both studies did not provide sufficient elaboration on how they came to their conclusions. This thesis has done so, first by evaluating the characteristics of *ke* against two well accepted models of DM characteristics (Schiffrin, 1987 and Schourup, 1999), and then by analysing its functions in selected extracts from a varied data set to compare its functions to the functions of other DMs studied elsewhere. The thesis's emphasis on this first part was necessitated by the evident lack of consensus about how to classify *ke* in existent Xhosa dictionaries, grammar books and research, and perhaps naming it appropriately can aid in better understanding its characteristics and functions. The third finding was also important to establish because the current lack of clarity about the classification, features and functions of *ke* is in part a result of historical processes already evident in the earliest missionary dictionaries and grammars (as demonstrated by the survey in Table 8 in Section 3.2). It was therefore important to address this specifically, but also to contribute to the body of research that looks at DMs not only as a contemporary development in utterance constructions but one with a very long history (see Lewis, 2011, Marcus 2009 and especially Kroon 1995). In between these findings, however, it was important to also point out the predominance of *ke* (and DMs in general) in oral utterances or utterances that have elements of orality about them (such as prose texts and online data) compared to utterances strictly located within formal registers – even if these too may contain *kes* that fulfil some DM functions within them.

¹²Jung, C.G. 1954 (2000). The Tibetan Book of the Great Liberation: a psychological commentary - The difference between Western and Eastern Thinking (Foreword). In W. Y. Evans-Wentz. *The Tibetan Book of the Great Liberation: The method of realising nirvana through knowing the mind*. New York: Oxford University Press (pp. xxxii).

I think these three findings are significant and valuable contributions not just to the linguistic study of Xhosa but also to DM literature in general, especially that of African languages. Nonetheless, the thesis is nowhere close to being exhaustive, not just about DMs but also about *ke*. The process of researching this thesis brought to the fore many insights about *ke* that had to be excoriated in order to focus on establishing the basic foundation laid down by these three main findings. Throughout the process of analysing the data presented in Chapter 2, many valuable insights about *ke* emerged but expounding upon them necessitated research and qualification of claims to an extent beyond the scope of this thesis, thus lending themselves to further exploration at Phd level or as separate research papers.

Therefore, although the direction and scope of some of the possible areas of research on *ke* can already be gleaned from the way the data sets are presented in Chapter 2, some of them are entirely absent from the presentation of the data sets. Patterns to analyse could be the distribution of *ke* tokens in the utterances of males and females; looking at whether age makes any difference in how *ke* is used; analysing whether there is perhaps an urban vs. rural divide in its use. There is also the classic problem of using interview data instead of natural speech – something that could influence how and how often these ‘conversation greasers’ (see list in Brinton, 1996) are used in such a context. There is also the possibility that a participant’s frequency of *ke* tokens can be influenced by the interviewer or co-participant they are in conversation with, and perhaps their gender and degree of supposed similarity they perceive in and ease they feel with them. Table 3, showing the frequency of interviewer *ke* tokens, shows consistent differences between the three interviewers and in turn mirrors the differences in the frequency of *ke* tokens in the corresponding interviews they conducted. Close analysis of the transcripts looking specifically at interviewer could reveal further insights about *ke*. Lastly, the differences between Tables 5 and 6, showing frequency counts in oral narratives and online data, respectively, (and then compared with Table 2, listing the frequency counts in the interviews) further points to the necessity for a study devoted solely to an analysis of *ke* in different utterance genres.

Another area of focus that deserves further systematic exploration but was eventually left out from this research despite the fascinating results it was showing is *ke* in syntax generally, but more specifically the high frequency of *ke* cliticising with Xhosa adverbials and conjunctive lexical items such as *ngoku* (now/then), *nangona* (although), *kuba/ngoba* (because),

uba/ukuba (if), *okanye* (or), *qha* (but/except) and *kodwa* (but), to sometimes produce utterances in the data such as *Qha ke kodwa ke...* (But [then] but [but]...) and *Kodwa ke uba mhlawumbi ke...* (But [then] if maybe [then]...). A related but separate element in this is *ke* double-coupling with conjunctives and personal pronouns in instances such as *Uyayiphazama ke wena ke ngoba...* (You *ke* are mistaking it *ke* because...), *Kanti ke yena ke lo mntu...* (But *ke* they *ke* this person...), and *Kuba ke thina ke singabobantu...* (Because *ke* we *ke* are people who...). Also of interest for further exploration would be *ke* cliticising with locatives.

Such constructions occur in all genres, but of course with varying frequencies in different utterances and with certain couplings made impermissible by convention in others (such as in the more formal registers). Moreover, in speech and online data (Facebook and blog comments) *ke* is sometimes attached to English adverbs, conjunctions and DMs, resulting in utterances such as *then ke*, *if ke*, *so ke*, *but ke*, *and ke*, *maybe ke*, *you know ke*, and *just because ke*, amongst others. Even the pronoun-*ke*-conjunction troika appears with English, in utterances such as in *But ke wena* (But *ke* you). All these examples are taken from the data assembled for this thesis and such constructions raise many questions about the function of *ke* in utterance construction and management and are worthy of systematic exploration.

Another area of focus that the early stages of data analysis revealed was the existence of what I consider to be other DMs in Xhosa, which another study could explore in depth. These are lexical items such as *uyaqonda* (you understand/you see) and *mhlawumbe/i* (maybe). Needless to say, but there are instances where these (possible) DMs also appear attached to *ke*, which makes for a very interesting study deciding which between the two items in such instances is a DM or whether one is more of a DM in that instance than the other or whether the combination as a whole is one DM. There is also the special case of the popular interjective ‘Hayi *ke!*’ which can be a chapter on its own. What this indicates is that more research into Xhosa DMs in Xhosa (and not only research on Xhosa speakers’ use of English DMs in English as in De Klerk (2005a and 2005b)) is necessary since virtually nothing has been done in this area. Tied to this is the need for a study that also focuses on Xhosa orature specifically in relation to its use of DMs and the Xhosa language as a whole uses them.

It is clear that Xhosa grammars and dictionaries favour more a grammatical account of *ke* which in use is possibly exclusively relegated to formal genres or registers of writing, especially legal texts. This is clear in the radical drop of the frequency of *ke* tokens in a text

such as the Xhosa translation of the Constitution (Table 7), where this low count seems to be as a result of a stronger adherence toward grammaticality and precision necessitated by such a document. The use of *ke* and its frequency become even more interesting when one looks at the legal statutes that were produced in Xhosa by the Transkei government during apartheid and were legally binding statutes, unlike the Xhosa translation of the current South African Constitution, where only the wording in its English version is the binding document in a court of law. A close study on the use of *ke* in extremely formal genres, especially the legal ones, could therefore be useful in reconciling the extreme positions of the grammarians on the one hand and the findings of this study on the other. Although this had started out as being a strand to deal with in this thesis, it also later became too complex to develop sufficiently within the frame of what is the current study.

Furthermore, future studies into *ke* also need to pay attention to questions around grammaticalisation, de-grammaticalisation and pragmaticalisation and these processes relationship with *ke*. Looking at Ziervogel's (1952) Swazi and Ziervogel and Van Schaik's (1959) Northern Transvaal Ndebele grammars which contain brief discussions on a basically identical *ke* in those languages with the *ke* in Xhosa but also with interesting differences, does suggest that it is possible that the Xhosa *ke* was once a dominantly grammatical particle, which then pragmaticalised over time. This pragmaticalisation, however, must have already happened by the time Xhosa began to be committed to paper, because from the earliest writings in Xhosa (as argued in Chapter 4) *ke* seems to have already been in possession of a number of its current DM characteristics and functions. However, one could also speculate that *ke* was a pragmaticalised particle even before it was committed to writing, but with the subsequent standardisation and/or formalisation of the language, particularly in written discourse, attempts by missionary and Xhosa grammarians and lexicographers have been on confining *ke* solely within a grammatical frame.

These ideas about the grammaticalisation, de-grammaticalisation and pragmaticalisation processes possibly surrounding the Xhosa *ke* are fuelled in part by the presence in various Southern African languages of particles that are not only spelt like the Xhosa *ke* but also have functions that seem to be slightly similar to some of the DM functions performed by the Xhosa *ke*, and might therefore tell us something about the historical roots of this *ke*. For instance, within the Nguni language group there is the *ke* in Swazi (Ziervogel, 1952) and the *ke* in Northern Transvaal Ndebele (Ziervogel and Van Schaik, 1959). But even more

interesting is the research by Güldemann and Siegmund (2009) on the !Ui branch of the Tuu and the Khoekhoe branch of the Khoe-Kwadi families in the Cape linguistic area which has been looking at the ‘clause type markers in clause second position after the subject’ (2009: 1). One of these markers is the N|uu declarative *ke* (and its variants *gye*, *dje*, *tje*, and *ken*; but the cognate *tje*, *dje* did not grammaticalise as a declarative marker). In this work, Güldemann and Siegmund (2009) condense the functions of the declarative *ke* into four categories: as an identificational marker, a contrastive term focus marker, as a reason or consequence coordinator, and as a non-focus cleft. Their analysis then shifts to focus on the clause-second non-focus *ke* as a declarative marker (for which, they point out, there are comparable markers in Nama and !Xam). The typical contexts of the declarative *ke* that they identify are clauses of exclamation and surprise, beginning of a narrative introducing participants and setting, the introducing of new participants within a narrative discourse, heightened assertions, weather expressions, as well as in contexts of setting, explanation and other background in narratives.

Güldemann and Siegmund’s (2009) description and analysis of the functions of this *ke* in these clauses are grammatical rather than pragmatic. The particle is not attributed as having any DM functions, although there are some instances in the presented data that seem to have non-declarative contexts and might indicate the pragmaticalisation of this *ke* in N|uu. Furthermore, based on a discussion with Dr. Menan du Plessis (2009, email personal correspondence) on the functions of the ‘declarative marker’ particle *ke*, *ge*, *g’e* (or in some cases *ken*) in some of the Khoekhoe varieties of KHOE (e.g. Nama and !Ora) and also several of the !UI languages, such as /Xam, augments my strong suggestion for a rigorous exploration not necessarily of the Xhosa *ke*’s possible historical areal diffusion across various languages, but a study investigating the Xhosa *ke*’s relationship to the processes of grammaticalisation, de-grammaticalisation and pragmaticalisation in historical terms which might relate it in some way to the *kes* in these other languages.

The last reflection I would like to make is perhaps a philosophical one. One of the key issues that this research has highlighted for me but one also beyond the scope of this thesis to address in any depth is that of the problems surrounding the discipline of linguistics itself. What value is there in knowing that a lexical item (or particle) that is sometimes a conjunction or an adverb is most of the time also a DM? With this new found knowledge are we to now expect the dictionaries of the future to have ‘d.m.’ right there with ‘conj.’ after the *and* entry because we have Schiffrin’s (1987) pioneering study which has identified it as

such? If this is the expectation, is it not risky, seeing how terminology in the discipline is always changing, not least of all when it comes to these discourse particles (as discussed in Chapter 1)? If, on the other hand, the findings of linguistic research are not meant to have a direct bearing on texts such as grammar books and dictionaries, what should be their relationship with these texts, especially considering sociolinguistics' self-imposed distance between its practices (seen as 'descriptive') and the aims of such texts (generally regarded by sociolinguists as 'prescriptivist')? What value should sociolinguistic descriptive work (which is not always without problems) be to prescriptivist approaches to language (which are not always entirely bad)?

It seems to me not exactly enough to generate new, perhaps more precise, terminology for linguistic phenomena if that precision in understanding remains (and sometimes becomes outdated) within the confines of linguistics as an academic discipline, never to be applied. If the argument is that there is no need for application, then the discipline approaches something closer to philosophy than most social sciences, and the implications of that raises more questions than the discipline as a whole has concerned itself with to date. (To start with, the first question to answer—philosophically, that is—would be how can we talk about what is happening in *a language* when we are not yet clear on what *language* is to begin with? From that, other questions would soon follow: the need for a clear working definition of what makes *a language* a language and how do we come to place boundaries between them in order to study them as genetically separate or related entities? What use, if any, can Bakhtin's (1986) definition of an 'utterance' be to dealing with this indeterminacy about 'language'? Was Bakhtin's all-encompassing use of 'utterance' instead of 'language' or 'speech' of any use at all?) I do think that serious engagement with such questions would also bring into question and possibly disrepute many of the methodological approaches thus far accepted as standard in a lot of linguistic research, not least DM research, due to some un-interrogated assumptions that underpin them, such as that this thing called language can be successfully compartmentalised, analysed, quantitatively or qualitatively, the methodology repeatable, etc.

BIBLIOGRAPHY

Grammars and Dictionaries

Appleyard, J.W. 1850. *Kafir language: comprising a sketch of its history which included a general classification of South African dialects, ethnographical and geographical remarks upon its nature and grammar*. Kingwilliam's Town: Wesleyan Missionary Society.

Bennie, W.G. 1939. *A grammar of Xhosa for Xhosa-Speaking*. Alice: Lovedale Press.

Boyce, W.B. 1844. *Grammar of the Kaffir Language*. 2nd Edition. London: Wesleyan Missionary Society.

Davis, W.J. 1872. *A dictionary of the Kaffir language: including the Xosa and Zulu dialects*. London: The Wesleyan Missionary Society.

Du Plessis, J.A & Visser, M. 1992. *Xhosa Syntax*. Pretoria: Via Afrika.

Fisher, A., Weiss, E., Mdala, E., & Tshabe, S. 1985/2005. *E. English – Xhosa dictionary*. Cape Town: Oxford University Press.

Kropf, A. 1899/1915. *Kafir-English Dictionary* 2nd Edition. Rev, Robert Godfrey Ed. Alice: Lovedale Mission Press.

McLaren, J. 1936/1944. *A Xhosa grammar: revised and re-written in the New Orthography*. London, New York, Toronto: Longmans, Green and Co.

McLaren, J. 1915/1899. *A new concise Xhosa-English Dictionary*. London: Maskew Miller.

Mini, B.M. 2003. *The Greater Dictionary of isiXhosa Vol 2*. S.C. Tshabe (Eds.) Alice: UFH IsiXhosa National Lexicography Unit.

Mncube, F.S.M. *Xhosa Manual*. Johannesburg, Wynberg, Cape Town: Juta and Company Limited.

Parker, P.M. 2008. *Webster's Xhosa-English Thesaurus Dictionary*. Imprint: Icon Group International.

Satyo, S.C. 2005. *Igrama noncwadi lwesoXhosa Ibanga 10*. Pretoria: Via Africa.

Stewart, J. 1921. *Outlines of kafir grammar with practical exercises* 3rd Edition. Alice: Lovedale Mission Press.

Tshabe, S.L and others. 1989. *IsiXhosa Sezikhuthali. Ibanga 10*. De Jager-Haum.

Ziervogel, D. 1959. *A Grammar of Northern Transvaal Ndebele*. Pretoria: Van Schaik.

Ziervogel, D. 1952. *A grammar of Swazi*. Johannesburg: Wits University Press.

Bible Translations:

1853. Itestamente Entsha: Okukuti, Innwadi Zonke Zocebano Olutsha, LweNkosi Yetu uYesu Kristu: Ikunyushelwe Kokwamaxosa. Ishicilelwe: Ngesishicilelo Sabafundisi.

1864. Itestamente Entsha: Oko Kukuti, Zonke Izibalo Zocebano Olutsha, LweNkosi Yetu uYesu Krestu: IguQulwe Kulonteto Yanikwa Kuqala Ngayo. Elondone: Kwelamangesi, Ishicilelwe Ngu-w. Clowes Nonyana Bake.

1879. Incwadi Yezibalo Ezingcwele, Zetestamente Entsha, Ziguqu. ELondone, Kwelamangesi. Ishicilelwe Ngu-w. Clowes Nonyana Bake.

1921. Incwadi Yesib'alo Ezingwele e Zetestamente Endaka ne Zetestamente Entsha. Ziguqulwe Kwezonteto Zanikwa Kuqala Ngazo. London: British and Foreign Bible Society.

1942. Incwadi Yezibalo Ezingcwele Izezetestamente Endala Nezetestamente Enta. London: British and Foreign Bible Society.

1962. Ibhayibheli Elingcwele. London: The British and Foreign Bible Society.

2002: Ibhayibheli Engwele. Inhlangano Yebhayibheli Yaseningizimu ne-Africa. Cape Town: Bible Society of South Africa.

Public Sources of Data

Constitutional Court Library and Website. 2005. Updated Constitution of South Africa, 1996, including the Constitution Twelfth Amendment Act, 2005.

Transkeian Governmentt. 1977. *Transkei Statute*.

Harold Scheub oral narratives data: <http://digicoll.library.wisc.edu/SouAfrVc/>).

Epozini – the House of Truth: <http://blogs.dispatch.co.za/epozini/>

Darkie's Place: <http://darkiesplace.blogspot.com>

Facebook Group: *Xhosa People Rock!!!* (<http://www.Facebook.com>)

Embo maXhosa: <http://blogs.uct.ac.za/blog/embo-maxhosa2>

Bennie, W.G. 1935. *Imibengo*. Alice: Lovedale Press.

Opland J & Mtuze, P.T. 1994. *Imizwi Labantu*. Cape Town: Oxford University Press

Research Literature

- Aijmer, Karin. 2002. *English discourse particles: Evidence from a corpus*. Amsterdam: Benjamins.
- Amfo, Appiah Aba Nana. 2010. Lexical signaling of information structure in Akan. *Linguistics* 48: 195–225.
- Andersen, G. 2001. *Pragmatic Markers and Sociolinguistic Variation: A Relevance-Theoretic Approach to the Language of Adolescents*. Amsterdam: J. B. Publishing Company.
- Barnes, B. Susan. 2006. *A privacy paradox. Social nextworking in the United States*. Available at <http://firstmonday.org/article/view/1394/1312>. Date Accessed: 31 October 2009.
- Bakhtin, M.M. (1986). The problem of speech genres. In: *Speech Genres and Other Late Essays*. Austin: University of Texas Press.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (Eds.). (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Blakemore, D. 1996. Are apposition markers discourse markers? *Journal of Linguistics* 32: 325–347.
- Brinton, L.J. 1996. *Pragmatic markers in English: Grammaticalization and discourse functions*. Berlin: Mouton de Gruyter.
- Britain, D. 1992. Linguistic change in information: the use of high rising terminals in New Zealand English. *Language Variation and Change* 4: 77-104.
- Burridge, K and Florey, M. 2002. ‘Yeah-no He’s a Good Kid’: A discourse analysis of ‘yeah-no’ in Australian English. *Australian Journal of Linguistics* 22: 149-171.
- Cameron, D. 1990. Demythologizing sociolinguistics: why language does not reflect society. *Ideologies of language*. J.E Joseph and T.J Taylor (Eds). New York: Routledge.
- Camiciottoli, Belinda Crawford. 2010. Discourse connectives in genres of financial disclosure: Earnings presentations vs. earnings releases. *Journal of Pragmatics* 42: 650–663.
- Crawshaw, C.J. 1903. *A first Kafir course* 5th Edition. Cape Town: J.C. Juta and Co.

- Crystal, D. (2001). *Language and the Internet*. Cambridge: Cambridge University Press.
- Crystal, D. 1995. Speaking of writing and writing of speaking. *Longman Language Review* 1: 5-8.
- Dailey-O’Cain, J. 2000. The sociolinguistic distribution of and attitudes toward focuser ‘like’ and quotative ‘like’. *Journal of Sociolinguistics* 4: 60-80.
- De Klerk, V. 2005a. The use of ‘actually’ in spoken Xhosa English: a corpus study. *World Englishes* 24: 275-288.
- De Klerk, V. 2005b. Procedural meanings of ‘well’ in a corpus of Xhosa English. *Journal of Pragmatics* 37: 183-1205.
- Der, Iona Csilla and Marko, Alexandra. 2010. A pilot study of Hungarian discourse markers. *Language and Speech* 53: 135-180.
- Deumert, A. (2006). Semantic change, the internet, and text messaging. *Pergamon Encyclopedia of Language and Linguistics*, K. Brown et al. (eds). Oxford: Elsevier.
- Doyle, A. C. 1890 (2001). *The Sign of the Four*. London: Penguin Classics.
- Dülger, Osman. (No date). *Discourse markers in writing*. Available at: http://www.sosyalbil.selcuk.edu.tr/sos_mak/makaleler%5COsman%20D%C3%9CLGER%5CD%C3%9CLGER,%20OSMAN.pdf
- Du Plessis, M. 2009. *Email correspondence*. University of Cape Town.
- Fay, W. 2001. Solidarity in Indonesian conversation: the discourse marker ‘ya’. *Journal of Pragmatics* 33: 171-191.
- Finlayson, R., Calteaux, K., & Myers-Scotton, C. 1998. Orderly mixing and accommodation in South African codeswitching. *Journal of Sociolinguistics* 2: 395-420.
- Fischer, Karin (ed.) (2006) *Approaches to discourse particles*. Amsterdam: Elsevier.
- Fox Tree, J.E. 2007. Folk notions of ‘um’ and ‘uh, you know’. *Text and Talk* 27: 297-314.

- Fox Tree, J.E. & Schrock, J.C. 2002. Basic meanings of 'you know' and 'I mean'. *Journal of Pragmatics* 34: 727-747.
- Fraser, Bruce. 2009. Topic Orientation Markers. *Journal of Pragmatics* 41: 892–898.
- Fraser, Bruce, 2006. Towards a theory of discourse markers. In: Fischer, K. (Ed.), *Approaches to Discourse Particles*. Elsevier, Oxford.
- Fraser, Bruce. 1998. Contrastive discourse markers in English. In *Discourse Markers: Descriptions and Theory*, Jucker and Ziv (Eds). John Benjamins.
- Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics* 6: 167-90.
- Fraser Bruce. 1990. An approach to discourse markers. *Journal of Pragmatics* 14: 1-28.
- Fraser, Bruce. 1988. Types of English discourse markers. *Acta Linguista Hungarica* 38: 19-33.
- Fuller, J.M. 2003. The influence of speaker roles on discourse marker use. *Journal of Pragmatics* 35: 23-45.
- Gough, D.H. 1986. *An analysis of the production and linguistic properties of discourse with particular reference to 'iintsomi' texts*. Doctoral Thesis, Rhodes University.
- Güldemann, Tom and Siegmund, Sven. 2009. *The marker 'ke' in N|uu – declarative or not?* Conference presentation at the 6th World Congress of African Linguistics, 17-21 August. Cologne.
- Halliday, M.A.K. and Hasan, R., 1976. *Cohesion in English*. London: Longman.
- Hernández A Todd. 2011. Re-examining the role of explicit instruction and input flood on the acquisition of Spanish discourse markers. *Language Teaching Research* 15: 159–182
- Hlavac, J. 2006. Bilingual discourse markers: evidence from Croatian-English code-switching. *Journal of Pragmatics* 38: 1870-1900.
- Hudson, R. 2004. Why education needs linguistics (and vice versa). *Journal of Linguistics* 40: 105-130.
- Jones, H., & Soltren, J. H. 2005. *Facebook: Threats to privacy*. December 14, 2005. Available at <http://www-swiss.ai.mit.edu/6805/student-papers/fall05-papers/facebook.pdf>

- Jordan, A.C. 1940. *Ingqumbo Yeminyanya*. Alice: Lovedale Press.
- Jucker, A. & Ziv, Y. 1998. Discourse markers: Introduction. A Jucker & Yael Ziv (eds.). *Discourse Markers: Description and Theory*. Amsterdam: John Benjamins.
- Jung, C.G. 1954/2000. The Tibetan book of the Great Liberation: a psychological commentary-the difference between Western and Eastern Thinking. *The Tibetan book of the Great Liberation: the method of realising nirvana through knowing the mind*. W.Y. Evans-Went (ed). New York: Oxford University Press.
- Koike, D.A. 1996. Functions of the adverbial 'ya' in Spanish narrative discourse. *Journal of Pragmatics* 25: 267-279.
- Kroon, C. 1995. *Discourse Particles: a study of nam, enim, autem, vero, and at*. Amsterdam: J.C. Gieben Publisher.
- Lewis, M. Diana. 2011. A discourse-constructional approach to the emergence of discourse markers in English. *Linguistics* 49: 415–443.
- Makoni, S., Brutt-Griffler, J & Mashiri, P. 2007. The use of 'indigenous' and urban vernaculars in Zimbabwe. *Language in Society* 36: 25-49.
- Marcus, Nicole E. 2009. Continuous Semantic Development of the Discourse Marker Well. *English Studies* 90: 214–242.
- Maschler, Y. 1997. Emergent bilingual grammar: The case of contrast. *Journal of Pragmatics* 28: 279–313.
- Mtuze, P.T. 1998. The dilemma facing the African translator in modern-day discourse. *Some Issues in African Linguistics*. Festschrift A.C Nkabinde & J.B Hlongwane (eds). J.L van Schaik.
- Norricks, N. R. 2001. Discourse markers in oral narrative. *Journal of Pragmatics* 33:849-878.
- Ogonah, Felix Nwabeze. 2011. The pragmatic roles of *as in* in Nigerian English usage. *World Englishes* 30: 200–210.

- Opland, J. 2003. A brief survey of nineteenth-century literature. *Proceedings of the Eastern Cape Historical Legacies and New Challenges Conference*. 27-30 August 2003. Fort Hare Institute of Social and Economic Research. Working Paper No.16.
- Östman, J.-O. 1981. *You know: A discourse functional approach*. Amsterdam: Benjamins.
- Penke, M & Rosenbach, A. 2004. What counts as evidence in linguistics? An introduction. *Studies in Language* 28: 480-526.
- Rabaney. 2009. Online: <http://huddlemind.net/profiles/blogs/facebook-south-africa-3-month>
- Redeker, G. 1991. Linguistic markers of discourse structure. *Linguistics* 29: 1139–1172.
- Schiffrin, D. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Schourup, L. 1999. Discourse markers. *Lingua* 107: 227-265.
- Schourup, L. 1985. *Common discourse particles in English conversation*. New York: Garland.
- Statistics South Africa. 2001. Census Data. Available at: <http://www.capetown.gov.za/en/stats/2001census/Documents/Suburb%20Index%20F.htm>
- Swan, J., Deumert, A., Lillis, T., & Mesthrie, R. 2004. *A Dictionary of Sociolinguistics*. Edinburgh University Press.
- Tagliamonte, S & D'Arcy, A. 2004. 'He's like, she's like': The quotative system in Canadian youth. *Journal of Sociolinguistics* 4: 493-514.
- Torres, L. 2009. Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31: 65-83.
- Torres, L. 2006. Bilingual discourse markers in indigenous languages. *The International Journal of Bilingual Education and Bilingualism* 9: 615-710.
- Tree Fox E. Jean and Schrock C. Josef. 2002. Basic meaning of you know and I mean. *Journal of Pragmatics* 34: 727-747.

Vanda, K. H. 2007. Native speaker and non-native speaker discourse marker use: the use of discourse marker 'well'. *Argumentum* 3: 46-53.

Vickers, Caroline and Goble, Ryan. 2011. Well, Now, Okey Dokey: English Discourse Markers in Spanish-Language Medical Consultations. *The Canadian Modern Language Review* 67: 536-567.

Wedekind, Klaus. 1990. Generating narratives: interrelations of knowledge, text variants, and cushitic focus strategies. *Trends in Linguistics: Studies and Monographs* 52. Werner Winter (ed) . Berlin, New York: Mouton de Gruyter.

Wertheim, S. 2007. Cleaning up for company: using participant roles to understand fieldworker effect. *Language in Society* 36: 707-727.

Willis, D. 2007. Syntactic lexicalization as a new type of degrammaticization. *Linguistics* 45: 271-310.

Wouk, Fay. 2001. Solidarity in Indonesian conversation: the discourse marker ya. *Journal of Pragmatics* 33: 171-191.