

## The pragmatic markers *anyway*, *okay*, and *shame*: A South African English corpus study<sup>1</sup>

---

Kate Huddlestone and Melanie Fairhurst

Department of General Linguistics, Stellenbosch University, South Africa  
E-mail: [katevg@sun.ac.za](mailto:katevg@sun.ac.za)

### Abstract

Pragmatic markers are “a class of short, recurrent linguistic items that generally have little lexical import but serve significant pragmatic functions in conversation” (Andersen 2001:39). While pragmatic markers are receiving growing consideration in the literature, pragmatic markers in South African English have been given little attention compared to other varieties of English. This paper provides a description of the distribution and functions of the pragmatic markers *okay*, *anyway* and *shame* as they occur in the spoken component of the South African version of the International Corpus of English (ICE). Using the commercially available Concordance program, WordSmith Tools, all instances of *okay*, *anyway* and *shame* were identified in the corpus and all non-pragmatic marker instances were then excluded. The remaining instances of *okay*, *anyway* and *shame* were then hand-coded to determine the primary functions that these elements exhibit. The classification of the functions of the pragmatic markers was carried out according to Fraser’s (1996, 1999, 2006) framework for identification of pragmatic markers. The findings of the corpus investigation included identifying the functions of *okay* as both a conversation-management marker and a basic marker, as well as its role in turn-taking. *Anyway* was found to function as an interjection, a mitigation marker, a conversation-management marker and a discourse marker. *Shame*, as a uniquely South African pragmatic marker, was found to function both as an interjection and as a solidarity marker, as an expression of sympathy or sentiment.

**Keywords:** discourse markers, pragmatic markers, South African English, corpus linguistics

### 1. Introduction

Africa offers many opportunities to study both New Englishes and World Englishes. While first-language varieties of South African English (SAE) are not considered to be New Englishes, South Africa’s many other languages have had a profound effect on the variety of the English language that is spoken in the country today, making it quite unique, as Crystal (2008:143) concurs:

---

<sup>1</sup> This paper is partially based on work done for the second author’s Master’s thesis (Fairhurst 2013).

I had studied the evolution of South African English over the years. There is nothing quite like it in the English-speaking world. The vocabulary is the really striking thing. It is hugely distinctive and diverse, thanks to the number of languages which feed it. There are eleven official languages in South Africa. Each one borrows wildly from the others. And English borrows most of them all.

The starting point for the study from which this article developed was the desire to delve into some of what makes SAE so unique. One aspect of a language that is strongly influenced by culture is that of pragmatics, how language is used and interpreted in context. The decision was made therefore to focus on pragmatic markers, as part of the vocabulary of SAE, and in light of their important role in contributing to pragmatic meaning. Such markers add little, if anything, to the semantic content of an utterance. Rather, they provide information on the speaker and on the speaker's attitude, among other aspects of the linguistic situation. Due to the nature of pragmatic markers, they are thought to reflect a speaker's cultural and linguistic background, and so to be ideal for contributing to an examination of what makes a particular first-language variety of English unique.

Aijmer and Simon-Vandenberg (2009) note that most studies of pragmatic markers place the emphasis on (spoken) corpus data, as corpora "make it possible to investigate the distribution of pragmatic markers in speech and writing and in different registers". For this reason, we elected to work with the International Corpus of English (ICE) for South Africa, ICE-SA, as this was the only spoken language corpus of SAE we were able to gain access to at the time.

Various researchers have examined different aspects of SAE, including non-native varieties such as Black SAE, with some even making use of corpora (cf. Mesthrie 1992, 2002; Gough 1996; Lass 2002; De Klerk and Gough 2002; Jeffery 2003; Jeffery and Van Rooy 2004; De Klerk 2005; De Klerk, Adendorff, de Vos, Hunt, Simango, Todd and Niesler 2006; Da Silva 2007 and Bekker 2009). However, there is little research to be found on pragmatic aspects of SAE.

This article will give a brief historical description of the variety of English examined in the study, SAE, followed by a general characterisation of pragmatic markers. A brief sketch of the field of corpus linguistics will then be provided, including a description of the corpus and the methodology used in the study. Finally, the data analysis and discussion will conclude the article.

## **2. South African English**

The English language holds a very interesting place in the South African linguistic landscape which goes back to when the British took over the government of the Cape Colony from the Dutch in 1795. The early years of British rule in South Africa centred on the Cape as a stopover for ships travelling to and from the East. Most of the English speakers living in the Cape at the time were military and government officials. In the late 1810s, Britain decided to expand their hold on South Africa and to start settling in some other areas of the country. The main goal at the time was to create a buffer between the Xhosa-occupied Eastern Cape and the British-settled Western Cape. For this purpose, the British government started providing assisted passage and land grants in the Eastern Cape, around the Fish River (Mesthrie 2002:108). In 1820, a group of about British 5000 settlers arrived in the Eastern Cape. Although the English speakers were,

at the time, outnumbered by the Dutch speakers, Lord Charles Somerset declared English to be the official language of the Cape Colony in 1822 (Mesthrie 2002:108). Even in the Boer Republics, which were established in the Free State and Transvaal, English was considered to be the language of the well-educated (Mesthrie 2002:109). In the 1840s and 1850s, a second large wave of settlers arrived in the Natal region. The third and most diverse wave of settlers, however, arrived from around 1875–1904, when gold was discovered and first came to be mined in the Witwatersrand. Although the settlers from the different waves mentioned would have brought with them different dialects and varieties of English, it would seem that “standard” SAE was mostly influenced by the first English-speaking settlers from the 1820s (Mesthrie 2002:109).

English has a fair distribution throughout South Africa, as both a first and second language, although it is more prominent in the metropolitan and urban areas. English in South Africa is not monolithic; it has a wide range of varieties. Clear distinctions can be made between White SAE, Coloured SAE, Indian SAE and Black SAE, with the latter being a predominantly second-language variety of English. Many people speak an African mother tongue at home, but go through their school careers in English; because of this, “South Africa’s second-language varieties of English are heavily marked at every level of linguistic structure by the primary language of their speakers” (Kamwangamalu 2006:162). This is of particular interest to the current study because, although all the data collected for the study were from people who received their schooling in English to matriculation level or beyond, their English might be marked by specific features if they are fluent bilinguals or multilinguals, or if English is not their mother tongue.

### **3. Pragmatic markers**

Pragmatic markers (PMs) serve several purposes in discourse. One of their primary functions is to point to features of the context indexically (Schiffrin 1987). Aijmer and Simon-Vandenberg (2009) further characterise PMs as reflexive, because they comment on the utterance, and thus assist in the interpretation thereof. Östman (1995, cited in Aijmer and Simon-Vandenberg 2009) refers to PMs as the “windows” that hearers use to make deductions and assumptions about the speaker’s attitude and opinion. Holker (1991, cited in Aijmer and Simon-Vandenberg 2009) lists four key features which can be used to characterise PMs:

- (i) PMs do not affect the truth conditions of an utterance;
- (ii) PMs add nothing to the propositional content of an utterance;
- (iii) PMs are related to the speech context or situation, rather than to the situation under discussion; and
- (iv) the function of the PM is emotive and expressive, rather than referential, denotative or cognitive.

PMs have been studied in various fields in linguistics, and the definition of a PM depends greatly on the linguistic approach that is taken in a particular study, which also influences whether or not an element is considered to be a PM. For this reason, the same element has also been referred to, variously, as “discourse particle”, “pragmatic marker”, “segmentation marker”, “modal particle” and “pragmatic particle”. In this paper, we use the term “pragmatic marker” and focus on the uses of PMs as outlined by Fraser (1996, 1999, 2006).

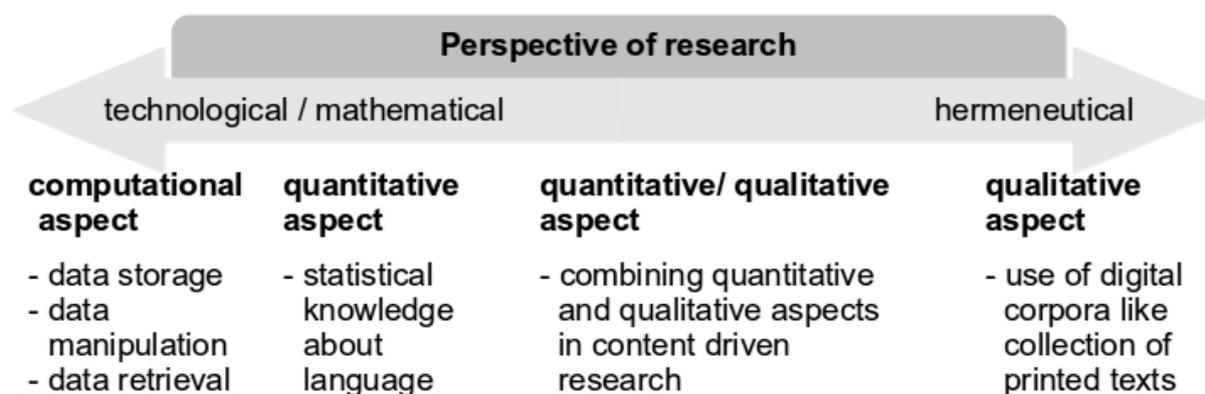
The first type of PM is the basic PM, with such markers conveying the illocutionary force of the speaker. The second type of PM is the commentary marker, which is used to indicate the fact that the following segment of discourse is connected to the previous segment. There are several types of commentary markers laid out by Fraser (1996, 1999, 2006). The third type of PM identified by Fraser is the parallel marker, which, in contrast to a commentary marker, is used to indicate that the following segment of discourse is separate from the previous segment. One of the subtypes of the parallel-marker type of PM is the conversation-management marker. The fourth and final type of PM is the discourse marker (DM).

#### 4. Corpus linguistics

In the language sciences a corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description. Over the last three decades the compilation and analysis of corpora started in computerized databases has led to a new scholarly enterprise known as corpus linguistics.

(Kennedy 1998:1)

The compiling of corpora for linguistic purposes has been performed since the 1950s; however, the field expanded significantly with the rise in computer technology. According to Baker (2007:1), corpus linguistics involves using “large bodies of naturally occurring language data stored on computers”, as well as “computational procedures which manipulate this data in various ways”, in order to find linguistic patterns. Stegmeier (2012) provides a summary of the different research perspectives that can be adopted for corpus linguistics, as is illustrated in Figure 1.



**Figure 1.** Research perspectives in corpus linguistics (Stegmeier 2012:96)

The present study falls under the quantitative/qualitative aspect of corpus linguistic research, as both small-scale statistical and context-based data are presented and analysed.

The corpus used in the current study originated as part of the ICE project, which aimed to compile parallel corpora of varieties of contemporary English (Nelson 2006). The ICE corpora have a common corpus design and a common methodology (Greenbaum 1996), and data are collected for the project only in countries where English is either the first language of, or is used as a second official language by, adult speakers of the language.

#### 4.1 International Corpus of English

The ICE corpora consist of 200 samples of written texts and 300 samples of spoken texts, all 2000 words in length, making a total of one million words for each corpus. The samples are drawn from several specified aspects of day-to-day life (see Table 1 as an illustration of how an ICE corpus is compiled).

**Table 1.** Design of ICE corpora

<b>SPOKEN (300)</b>	<b>Dialogues (180)</b>	<b>Private (100)</b>	Face-to-face conversations (90) Phone calls (10)
		<b>Public (80)</b>	Classroom lessons (20) Broadcast discussions (20) Broadcast interviews (10) Parliamentary debates (10) Legal cross-examinations (10) Business transactions (10)
		<b>Unscripted (70)</b>	Spontaneous commentaries (20) Unscripted speeches (30) Demonstrations (10) Legal presentations (10)
	<b>Monologues (120)</b>	<b>Scripted (50)</b>	Broadcast news (20) Broadcast talks (20) Non-broadcast talks (10)
<b>WRITTEN (200)</b>	<b>Non- printed (50)</b>	<b>Student writing (20)</b>	Student essays (10) Exam scripts (10)
		<b>Letters (30)</b>	Social letters (15) Business letters (15)
	<b>Printed (150)</b>	<b>Academic writing (40)</b>	Humanities (10) Social sciences (10) Natural sciences (10) Technology (10)
		<b>Popular writing (40)</b>	Humanities (10) Social sciences (10) Natural sciences (10) Technology (10)
		<b>Reportage (20)</b>	Press news reports (20)
		<b>Instructional writing (20)</b>	Administrative writing (10) Skills/Hobbies (10)
		<b>Persuasive writing (10)</b>	Press editorials (10)
		<b>Creative writing (20)</b>	Novels and short stories (20)

Although the ICE corpora can stand alone as a useful tool for research, their true value comes from the fact that they are exactly comparable, and therefore indispensable to today's study of World Englishes.

## 4.2 ICE-SA

SAE was originally not going to be included in the ICE corpora, due to political reasons; however, this ban was eventually lifted and research began in June 1992 (Jeffery 2003:341). Chris Jeffery of the University of Port Elizabeth was the lead researcher from the start, but worked with teams collecting data from all over the country. The initial plan was that all the data used would be collected between 1990 and 1996. The set time frame, however, proved to be too restrictive and so was left open-ended. The population to be sampled had to be 18 years of age or older, and they had to have completed their education in English up to matriculation level (Jeffrey 2003). This corpus has yet to be released via the ICE website and was made available to the researchers by Bertus van Rooy (NWU), who, through his collaboration with Jeffery and in his role as director of the South African component of the International Corpus of Learner English (ICLE) project, now has control of the ICE-SA project.

Table 2 provides a statistical characterisation of the make-up of ICE-SA's spoken component. As can be seen from the number of tokens (running words) in the text, ICE-SA is not complete, falling approximately 200 000 words short of the 600 000 word target for ICE corpora. As Jeffery (2003:343) notes, certain categories, specifically the Spoken Monologue section, are difficult to fill, while access to private telephone calls is also problematic. It is worthwhile noting that about half of the words in the corpus are contained in what can be characterised as private conversations/dialogues, which one could argue are the most authentic types of spoken discourse. In this respect, then, one can consider the ICE-SA corpus to be sufficiently representative of SAE, taking into account its current size.

**Table 2.** Statistical composition of ICE-SA

	ICE-SA	Overall
tokens (running words) in text		407 254
tokens used for word list		404 285
types (distinct words)		19 240
type/token ratio (TTR)		5

Unfortunately, where the pre-final state of ICE-SA is somewhat of a hindrance to comprehensive corpus analysis is in the lack of mark-up in a portion of the transcriptions that comprise the corpus, specifically, in certain transcriptions of face-to-face conversations. Furthermore, some might see the fact that the corpus is not tagged as a drawback, however, as Hunston (2002:93) points out, "the categories used to annotate a corpus are typically determined before any corpus analysis is carried out, which in turn tends to limit, not the kind of question that can be asked, but the kind of question that usually is asked". As the present study is corpus-driven, pre-tagged text is not required; rather, the raw text is examined directly and, as Sinclair (2004:191) notes, "patterns of this uncontaminated text are able to be observed".

One final problematic aspect of the spoken component of ICE-SA is the apparent lack of comprehensive metadata for all the texts included in the corpus. While Jeffery (2003:343) states

that, for example, “each speaker’s population group is identified in the header”, identifying metadata – including speakers’ sociological and linguistic background – is not consistently indicated across all the texts included in the corpus.

### 5. Methodology

Statistics on the composition of the corpus were determined using the Concordance program WordSmith Tools (Scott 2012). As can be seen in Table 2, the total number of words for the spoken component of ICE-SA is approximately 400 000 words, with an overall type/token ratio<sup>2</sup> (TTR) of 4.75. An initial search was undertaken to determine the prevalence of various pragmatic markers, specifically, *anyway*, *but*, *I mean*, *ja*, *just*, *like*, *no*, *now*, *oh*, *okay/ok*, *right*, *shame*, *so*, *well* and *you know*. Figure 2 presents a screenshot of the WordSmith concordances of *shame* in the ICE-SA corpus as an illustration of the results of such a search.

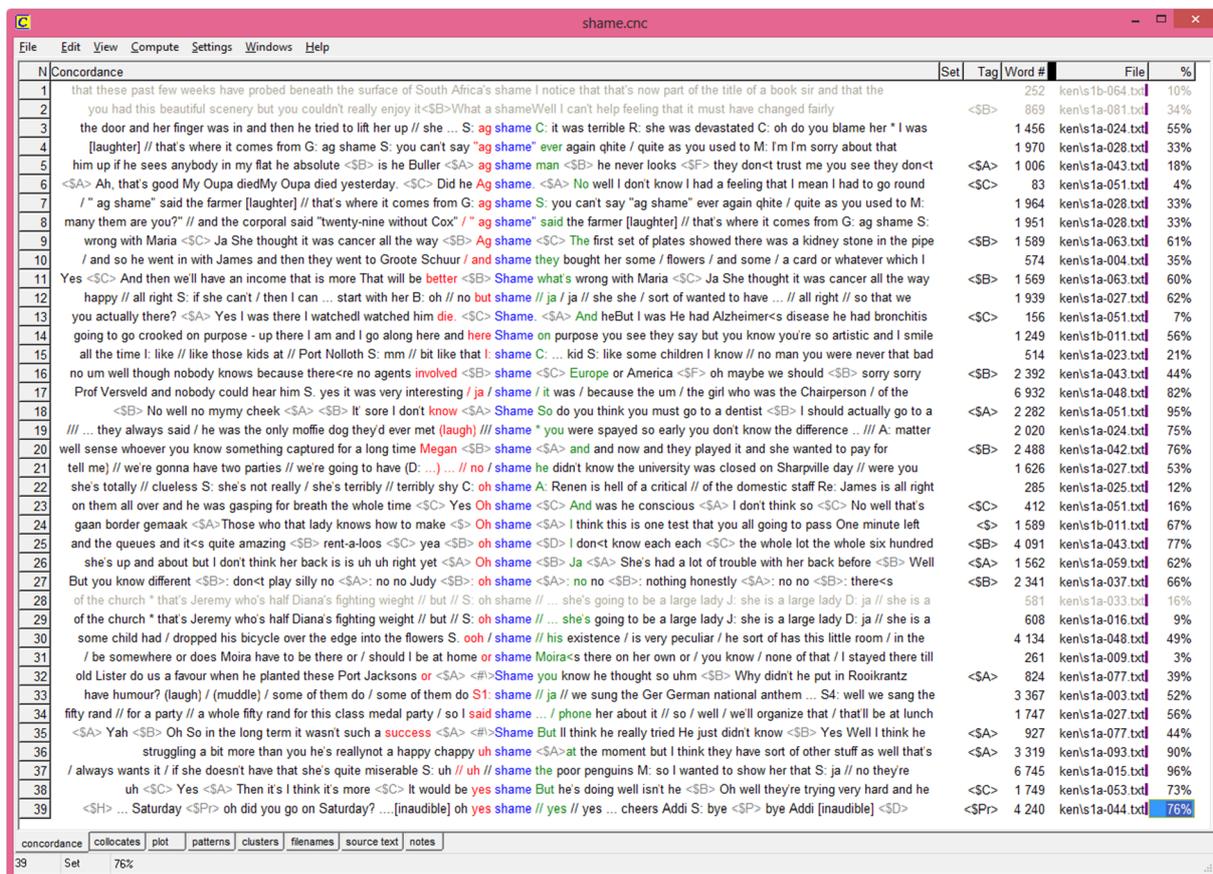


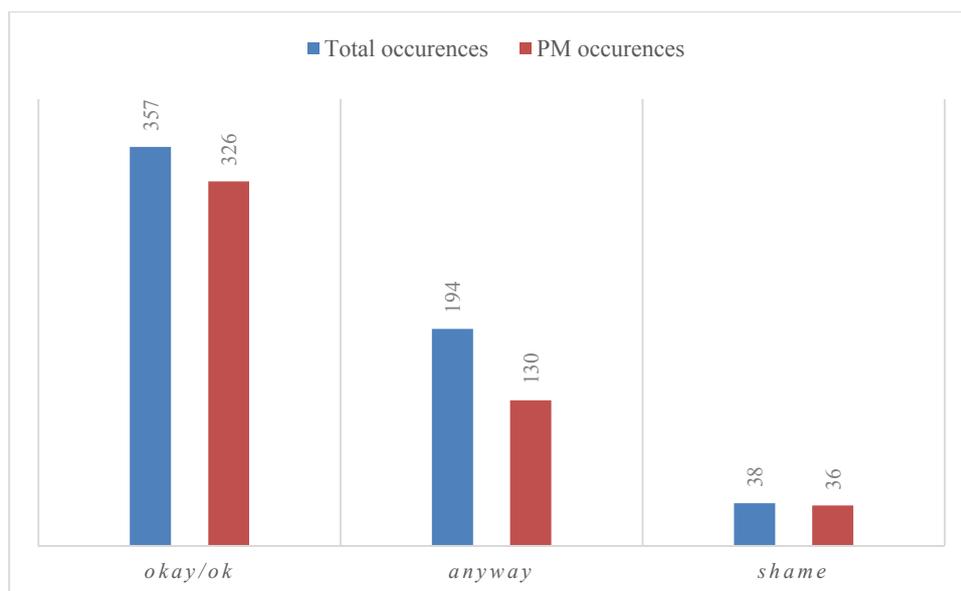
Figure 2. Wordsmith Tools screenshot

The choice of these specific markers was determined by various factors. Firstly, we considered the literature to determine which specific PMs had been examined as particularly representative of culture or group. Secondly, we considered Fraser’s (1996) categorisation of PMs when looking at representative PMs of different categories. Thirdly, we used our own intuitions about which PMs are likely to be unique to SAE. As one of the characteristics of PMs is “multi-categoriality” (Schourup 1999:234), it was essential to determine which of the instances in the search results were non-PMs, and exclude them from the analysis. Given that the scope of the study from which this article developed was limited, we therefore restricted our subsequent

<sup>2</sup> Number of Types divided by Number of Tokens times 100.

investigation to three PMs, namely *okay*, *anyway* and *shame*, based on their prevalence in the corpus<sup>3</sup>, and, in the case of *shame*, on its uniquely South African nature.

Once the concordance list of total occurrences for each word had been obtained, they were examined, line by line, and all the instances of PMs were selected. Figure 3 graphically represents, for each corpus, the total number of instances of each word found versus the number of instances of that word as a PM. Interestingly, while the prevalence of the PMs *okay* and *shame* make up 90% and 95% of the total number of occurrences of these elements, respectively, *anyway* occurs as a PM only 55% of the time.



**Figure 3.** The total and PM occurrences of *okay*, *anyway* and *shame* in each corpus

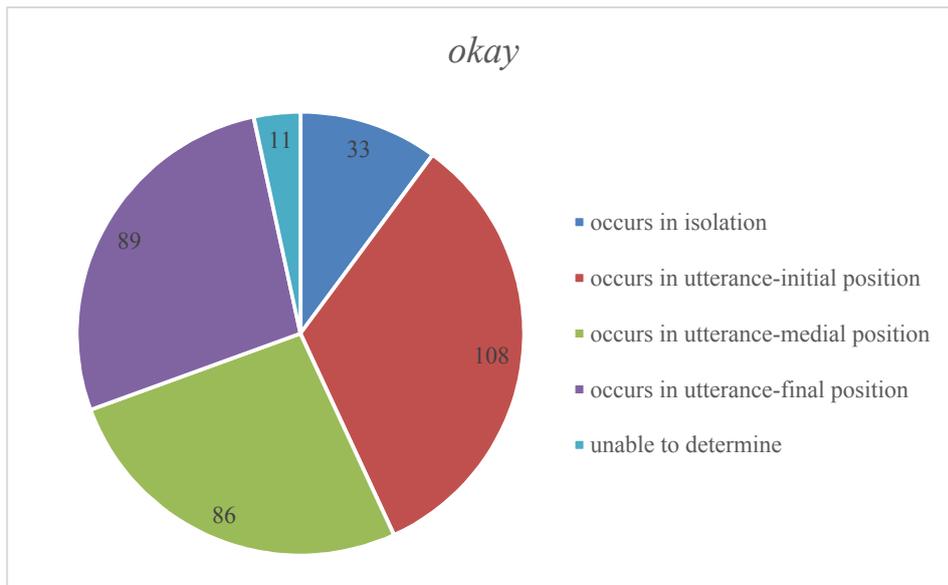
## 6. Data analysis and discussion

In this section, we discuss the various occurrences of the three PMs, characterising their distribution and identifying and illustrating the primary functions that these PMs perform in the ICE-SA corpus, as representative of educated SAE.

### 6.1 *Okay*

The PM *okay* (and its alternate *OK/ok*) is the most frequent of the three PMs. The PM *okay* occurs in various utterance positions in the ICE-SA corpus. Approximately 40% of the instances of *okay* occur in utterance-initial position or as the only element in an utterance. The second-most prevalent position for *okay* is in utterance-final position, followed by its occurrence in utterance-medial position. However, given the nature of transcribed speech, with its lack of prosodic indications, it is possible that a more accurate analysis of some instances of utterance-medial *okay* would be as utterance-initial or -final. For a small number of occurrences of *okay*, it is not possible to determine what positions they occupy, although in all such cases, *okay* occupies an utterance-peripheral position. Figure 4 presents a graphical representation of the number of times *okay* as a PM occurs in each utterance position in the ICE-SA corpus.

<sup>3</sup> Results of less than 500 concordance lines.



**Figure 4.** Utterance position of PM *okay* in ICE-SA

Examples (1)-(3)<sup>4</sup> illustrate the occurrence of *okay* in the three utterance-related positions, respectively.

- (1) <\$A> **Okay** As I mentioned in the beginning uhm as a scientist (SAE, s2a-027)
- (2) ... Art is essentially mysterious **okay** and truth has to be comprehensible otherwise it's not truth ... (SAE, s1b-003)
- (3) <\$H> ... You mustn't take it any more **OK**  
<\$K> The doctor said I must take it (SAE, s1a-083)

Gaines (2011:3292) notes that various studies of the PM *okay* have shown that it performs “an almost bewildering array of functions”. Some of these functions observed in the corpus will be discussed and illustrated, after which an analysis of the distribution of this element in the corpus will be provided in order to highlight some interesting aspects of this PM.

The PM *okay* is able to serve several functions in the utterance-initial position. One function of *okay* in this position is to draw attention to the speaker, as illustrated in (4). In terms of this function, *okay* plays an important role in the indication of turn-taking. This function of *okay* is a way for the speaker to acknowledge their turn and to prepare to speak.

- (4) <\$C>I can't say that  
<\$A><#>**okay** that's uh uh now I want to ask you why did didn't you stop the vehicle  
you were just nine metres behind the vehicle (SAE, s1b-066)

Another function of the PM *okay* is that of introducing a new topic. As with the turn-taking use of *okay* illustrated in (4), in cases such as that illustrated in (5), *okay* functions as a parallel marker (Fraser 1996:168). Specifically, *okay* is used as a conversation-management marker, a

<sup>4</sup> The examples have been presented in this article as they appear in ICE-SA.

subtype of parallel markers (Fraser 1996:168), as the speaker uses it to steer the conversation towards a forgotten or unrelated topic.

- (5) ... and the high density plastic Both of them are recyclable **okay** the question is what happens to the stuff when once we collect (SAE, s2a-027)

When *okay* appears as a PM in the utterance-final position, it serves one of two functions. The first function, as with *okay* in utterance-initial position, has to do with turn-taking. *Okay* acts as an indicator that the speaker has finished speaking, and that it is now the other individual's turn to start talking. As was mentioned before, the next speaker will often start their turn with the PM *okay* to reinforce the turn transition. Beach (1993:341) refers to this function as a "projection device for turn and, at times, speaker transition".

The simplest reason for *okay* appearing in the utterance-final position is the fact that the speaker is giving others the option of asking for clarification of what they he/she has just said, a so-called "tag-positioned comprehension check" (Broderick and Broderick 2003, cited in Gaines 2011:3292). As indicated in Table 1, a portion of the corpus is made up of classroom interactions and unscripted speeches such as those found in the lecture hall. In such educational situations, the educator is often seen to end an utterance with *okay*. The utterance functions as an informal way of asking whether the students have understood what has been said, and whether they are ready to move on to the following aspect of the topic. This is illustrated by the utterance in (6).

- (6) the history of or the narrative of spirit on the way to truth <,>**okay** That's not a problem for him (SAE, s1b-003)

The final position in which *okay* appears is the utterance-medial position. *Okay* occurs in this position for several reasons. A primary reason is that the speaker needs to pause in order to collect their thoughts, but does not want the pause to be silent, as demonstrated in (7). In these instances it performs a gap-filling function.

- (7) <\$A> Then the only thing that I want **OK** is just an explanation from you (SAE, s1b-004)

In some cases, *okay* appears in the utterance-medial position, acting, however, as if it were in the utterance-initial position. Such action occurs when the speaker is reporting speech. Often a speaker starts reported speech in exactly the same way in which the speech was given, starting with the PM *okay*, as we see in (8).

- (8) What about if we collect the stuff and we say to people **OK** I'll give you the bread but in exchange I want one bag of plastic (SAE, s2a-027)

These examples above illustrate that *okay* as a PM appears most prominently as a conversation-management marker (Fraser 1996:185), as it is used to control the flow of the conversation. In such control, *okay* is often used to take the floor or to introduce reported speech, thought processes or an offer.

In terms of the distribution of *okay* in the corpus, there are some interesting observations to be made. Firstly, the highest number of occurrences of *okay* for one single text was found in an

interaction recorded in the pharmacy department at the University of Port Elizabeth (now Nelson Mandela Metropolitan University) where medicine was being dispensed. In this 2000-word text, *okay* occurred 69 times, constituting 20% of the total occurrences of *okay* in the corpus. Furthermore, the majority of these occurrences were either utterance-final comprehension checks, or the occurrence of *okay* in isolation as a “signal of understanding” (Condon 2001:495). A further 30% of the total occurrences of *okay* occurred in the section of the corpus devoted to lectures and seminars – although, in this case, the overall word count for the texts which contained these instances of *okay* was considerable higher (more than 40 000 words). Such a weighting is to be expected, however, given that, as noted above, tag-positioned comprehension checks are a feature of educational contexts.

A second interesting observation is that 43 of the occurrences of *okay*, 13% of the total, were found in the telephone exchange section of the corpus, which itself consists of only 6000 words. However this should not actually be all that surprising, given the attention devoted to *okay* in telephonic exchanges (cf. Beach 1993) as a receipt marker in conversation. Finally, the largest section of the corpus – at close to 200 000 words – was the face-to-face private conversations, which only exhibited 35 occurrences of *okay* (10% of the total). This last observation shows that in typical spoken conversation, the PM *okay* is not particularly frequent. A point to follow up in future investigations would be to see what other elements in spoken SAE perform the functions of the element that Levin and Gray (1983:195) refer to as “the most versatile utterance in English”.

## 6.2 *Anyway*

As seen in Figure 3, *anyway* appears as a PM less than half as frequently as the PM *okay*, but three and a half times as frequently as *shame*. Furthermore, unlike *okay* and *shame*, under 70% of the occurrences of *anyway*, around two thirds, can be classified as PMs. The other 30% of the time *anyway* functions as an adverb, with either a dismissive, contrastive or modificative use (Coll 2009: 161). This later function is limited to utterance-final instances of *anyway*, while the utterance-initial occurrences of *anyway* function as PMs. Of course, it is possible to argue that the utterance-final adverb *anyway* functions as a cohesive device, making it plausible that it should also be classified as a PM, however we leave that question open for further research.

The PM *anyway* performs several functions, related to discourse discontinuity and digression management. It functions primarily as a DM (Fraser 1996:186), specifically, as a resumptive marker (Ferrara 1997: 350). It signals a conversational reentry after a diversion of some sort, therefore connecting two segments of speech. This is illustrated in (9).

- (9) reservoir and it’s lovely and they all swim in the reservoir / and **anyway** there / there’re these sheer rock faces (SAE, s1a-004)

In example (9), the speaker is telling a story about an event that happened at a rock cliff, but went off at a tangent when describing their position and appearance. They then used the DM *anyway* to indicate that the following piece of speech was connected to the previous one.

*Anyway* can also function as a conversational management marker, which, as noted above, is a sub-type of Fraser’s (2006) parallel marker category. *Anyway* can be used to resolve an interactional impasse (Park 2010), and to control or to regulate the flow of the conversation. With this function *anyway* is used to acknowledge the interactional impasse, and to indicate to

the other participant(s) that the topic can be changed. It is most often the speaker who created the impasse who uses *anyway* in this manner (Park 2010: 3297). In example (10), speaker one has made a somewhat absurd observation, and so has created an impasse. Speaker two uses *anyway* to resolve the impasse, and to ensure that the conversation continues.

- (10) S: no // your camera hasn't got a brain / so it can't tell  
 J: **anyway** // and then we went to ... \* I told you // and this was the ... (SAE, s1a-016)

One further possible function of *anyway* may be provisionally identified from the data, specifically, *anyway* may function as a mitigation marker (Fraser 1996:183). Mitigation markers are used to reduce the loss of face that is associated with a certain message. In example (12), the speaker has had to contact a nun for some form of assistance. The speaker is embarrassed by what they have to say, and so use the mitigation marker *anyway* in an attempt to save face.

- (11) Yes it's a problem at the moment uhm But **anyway** then I got hold of Sister Ethel at the Missionvale (SAE, s2a-027)

As noted above, utterance position is determined by the nature of *anyway* as a PM marker, however, what we can discuss finally in relation to this element, is its distribution in the corpus. Of the 130 instances identified of PM *anyway*, 118 occur in the face-to-face conversation section of the corpus (none occur in the medicine dispensing exchange), five occur in the private telephone conversations, while the remaining seven are found in the rest of the corpus. Interestingly, this distribution is mirrored by the distribution of all occurrences of *anyway* – 172 instances occur in the face-to-face conversations, five in the telephone conversations and 17 in the rest of the corpus.

### 6.3 *Shame*

In the ICE-SA corpus, *shame* appears 38 times of which 36 can be identified as PMs. This is a very small number of occurrences, although, interestingly, if we compare this item's occurrence in SAE to another variety of English, for example, East African English (ICE-EA, one of the other ICE corpora available online), we see that *shame* appears only three times in the ICE-EA corpus. These occurrences are given in (12) and (13).

- (12) We women we have that **shame**. We are not like guys have no **shame** and guys you see... (EAE, conversation1k)
- (13) He sees the death of the minister as a bringer of **shame** and wanton destruction (EAE, sch-broadcastk)

In the first example from the ICE-EA corpus, example (13), *shame* appears twice, with neither of the occurrences being a PM, as is also the case in (14). Instead, the elements function as nouns, with the canonical meaning of “a painful feeling of humiliation or distress caused by the consciousness of wrong or foolish behaviour” (Oxford Dictionary of English 2010)<sup>5</sup>. This is also the meaning found in one of the two non-PM occurrences of *shame* in ICE-SA, as

<sup>5</sup> The Oxford Dictionary of English (2010) also provides a definition of “shame” as a South African exclamation which is used to express sympathy or sentimental pleasure, especially when referring to “something small and endearing”.

illustrated in (14), the other being an occurrence of the standard English expression “what a shame”.

- (14) At the funeral of Stompie Sipei I said that his death was an unspeakable crime and that these past few weeks have probed beneath the surface of South Africa's **shame** (SAE, s1b-064)

This shift in function and the bleaching of semantic meaning in the SAE data from canonical meaning to pragmatic marker can be categorised, in Aijmer’s (1997, cited in Aijmer and Simon-Vandenberg 2009) terms, as “pragmaticalisation”. The lexical item now performs a pragmatic function, similar to the way a lexical item may become grammaticalised in order to perform a grammatical function.

In the ICE-SA corpus, the PM *shame* occurs primarily as an isolated utterance or as an utterance-initial element. In both cases it is often preceded by another element, *oh*, or its (originally) Afrikaans equivalent *ag*. In fact, out of the 36 PM instances of *shame*, 7 occur together with *ag*, while 8 occur together with *oh*, functioning as complex PMs, as the examples in (15) and (16) illustrate.

- (15) <\$C>Ja She thought it was cancer all the way  
<\$A> **Ag shame**  
<\$C> The first set of plates showed there was a kidney stone (SAE, s1a-063)
- (16) <\$B> Well she's up and about but I don't think her back is is uh uh right yet  
<\$A> **Oh shame**  
<\$B> Ja (SAE, s1a-059)

Of the 36 instances of (*oh/ag*) *shame*, 12 occur in isolation, while nine occur as the first element in a speaker’s utterance. The examples in (17) and (18) illustrate the occurrence of *shame*, on its own, as either the first word or the only word spoken in a turn.

- (17) <\$A> Yes I was there, I watched I watched him die.  
<\$C> **Shame.**  
<\$A> And he (SAE, s1a-051)
- (18) <\$B> Yah Did Did old Lister do us a favour when he planted these Port Jacksons or  
<\$A> **Shame** you know he thought so uhm  
<\$B> Why didn't he put in Rooikrantz (SAE, s1a-077)

Of the remaining occurrences of *shame* in the ICE-SA corpus, 10 occur utterance-medially, while only two occur utterance-finally. In the case of the utterance-medial occurrences of *shame*, all but one are preceded by another PM, as illustrated by the example in (19), making them essentially part of a string of utterance-initial PMs, rather than strictly utterance-medial.

- (19) B: oh // no but **shame** // ja / ja // she she / sort of wanted to have ... (SAE, s1a-027)

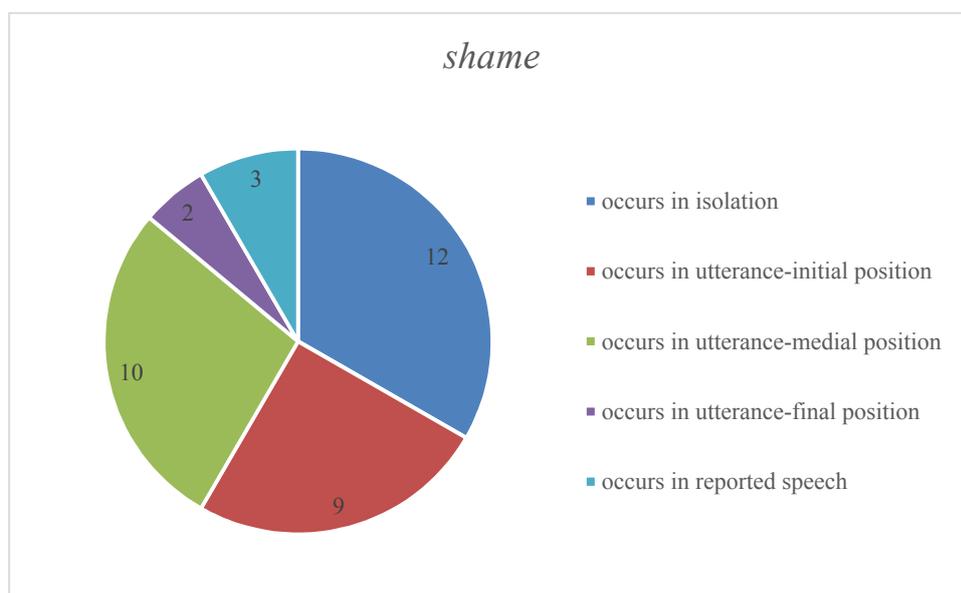
The only truly utterance-medial occurrence of *shame*, given in (20), is also the only occurrence of *shame* in a text other than a privately recorded informal conversation. This instance of *shame* was uttered by a teacher giving instructions in a school classroom interaction.

- (20) ... the borderline of my frame and I'm going to take the flat side of my pen and I'm going to draw a line from there to there and I'm going to go crooked on purpose - up there I am and I go along here, and here, **Shame**, on purpose you see, they say "but you know you're so artistic" and I smile ... (SAE, s1b-011)

Finally, *shame* occurs three times in reported speech, as illustrated by the example in (21).

- (21) a whole fifty rand for this class medal party / so I said **shame** ... (SAE, s1a-027)

Figure 6 presents a graphical representation of the number of times *shame* as a PM occurs in each utterance position in the ICE-SA corpus.



**Figure 6.** Utterance position of PM *shame* in ICE-SA

From the context of the examples of *shame* presented above, as well as the other instances of (*oh/ag*) *shame* observed in the corpus, it is clear that these elements function primarily, in Fraser's (1996:176) terms, as basic markers, specifically as interjections or pragmatic idioms. However, in SAE, the PM *shame* can also be characterised as a solidarity marker (Fraser 1996:185). The function of *shame* in such contexts is used to express sympathy with the hearer or with the person or character in the story that is being told.

Another function of *shame* is somewhat harder to define. It is used as another way of saying, "Oh, that thing is adorable". Such an expression is most notably used with regard to animals and children, as illustrated by the example in (22).

- (22) S: uh // uh // **shame** the poor penguins (SAE, s1a-015)

In this example, two people are discussing a postcard with a picture of penguins on it. It is important to note that there is nothing wrong with the penguins. The speaker in (24) is simply expressing the fact that the penguins are "cute". This use of the PM *shame* ties in with the "sentimental pleasure" aspect of the Oxford English Dictionary (2010) definition given in footnote 4.

Notable in relation to interjections, is the fact that the meaning assigned is often conveyed by the intonation that is imposed on them, rather than the actual form. Interjections therefore express much of their meaning in their intonation. Due to this characteristic, interjections can be used sarcastically. In SAE, the same holds true for the use of *shame*. Take, for example, the constructed utterance in (23):

(23) He's got a cold, and it's the end of the world to him. **Shame**, he's taking it very hard.

In this example, *shame*, in its “sentimental” function, is being used to assign childlike reactions or qualities to an adult. The use seems to be a combination of the two uses that have been mentioned: to show empathy and to express pleasure, due to a desire to smooth over a situation and to be endearing. Unfortunately, the one drawback of a transcribed corpus is the lack of intonational cues available to the researcher interested in analysing the functions of elements that are so dependent on context for interpretation.

As noted in section 4.2, speakers' sociological and linguistic backgrounds are not consistently indicated across all the texts included in the corpus, making it difficult to speculate on age- or gender-based differences in patterns of language use. Anecdotally, the solidarity marker function of *shame* seems to apply equally for male and female first-language speakers of SAE, while the sentimental function seems to be used predominantly by women. This is supported by the metadata that is currently available: of the 36 occurrences of *shame*, metadata indicating the gender of speakers is available for 20 of the 36 occurrences of *shame*; of these 20 occurrences, 19 are by women, while only one use of the PM *shame* can be allocated to a male speaker (see example (24) for this particular utterance). Based on contextual information and inferences about the conversations recorded for the ICE-SA corpus, one can say with reasonable certainty that of the remaining 16 occurrences of *shame*, 10 are more than likely produced by one or more female speakers. However, given the incomplete nature of the metadata available – especially as regards the total number of male versus female speakers – observations based on these inferences cannot be generalised for SAE.

Finally, in terms of the distribution of the PM *shame* in the ICE-SA corpus, as noted above, we find all but one of the occurrences of *shame* in private conversations, recorded as part of the Dialogues section of the corpus. These occurrences are distributed over 24 texts, of which one is a telephone conversation and the rest are face-to-face conversations. Four of the face-to-face conversation texts contain two instances of *shame*, three texts contain three instances of this PM, while one text contains four instances thereof. In the latter case, three out of the four instances of *shame* are produced by the same speaker, who is in fact the data collector, and so is functioning in the role of facilitator. This ties in strongly with the observation that *shame* functions as a solidarity marker; in this case, the data collector is creating a sense of solidarity in order to elicit conversation. The remaining 14 texts contain one occurrence each of the PM *shame*. While these statistics indicate that the use of *shame* is not particularly widespread in the corpus, it is still worthy of attention and analysis, as Norrick (2009:863) points out, “[c]orpora are now finally large enough to assemble sufficient data on these less frequent pragmatic markers for significant analysis”.

The use of *shame* as a PM is distinctly South African, and there is plenty of scope for speculating as to how this use came about. One possibility is that it was as a result of the influence of the Afrikaans language, specifically of the expression *sies tog*, which functions in the same way as (*ag/oh*) *shame* in the above examples. While we leave such speculation for

further study, support for this supposition comes from the joke narrated in example (24), where the Britishness of the soldiers, with their overzealous field cornet, Cox, is contrasted with the (probable) Afrikaansness of the farmer through the use of the expression *ag shame*.

- (24) M: .../ and um // the corporal said um // “we’re a platoon of the Queen’s Own Mounted Yorkshire Fox Terriers um / cut up from our main task force under Lord Roberts / chased across the veld by an overzealous field cornet / excaped under a cover of nightfall / marched through the whole of the day without fodd and shelter can you help us?” / and the farmer / also wanting a degree of qualification said / “how many them are you?” // and the corporal said “twenty-nine without Cox” / “ag shame” said the farmer [laughter] // that’s where it comes from  
 G: ag shame  
 S: you can’t say “ag shame” ever again qhite / quite as you used to  
 M: I’m I’m sorry about that [laugh] (SAE, s1a-028)

## 7. Conclusion

This article presented a description of the distribution and functions of the PMs *okay*, *anyway* and *shame* as they occur in the spoken component of the ICE-SA. The analysis in this article was limited by the restricted scope of the Master’s study from which it took its starting point, and recommendations for further study would be to compare the findings using the ICE-SA corpus with other ICE corpora, especially other southern hemisphere varieties of English, and to other corpora of SAE, as and when they become available. Of course, in order to allow for comparative studies, the corpus data need be altered to include token level annotation, as well as text level annotation, such as comprehensive metadata information on the text type, the speakers’ sociological backgrounds (including education), and so forth.

Given the limited scope of the present study it would also be valuable to conduct further research into other PMs in SAE. One of the limitations of this and similar studies, however, is the fact that much meaning assigned to PMs is based on a speaker’s intonation, and so cannot be completely analysed with a corpus consisting only of written transcriptions of spoken texts.

The strength of the current study lies in the fact that it is one of the first to use the ICE-SA corpus. Furthermore, despite the fact that this corpus has not necessarily been completed, nor all the transcriptions fully marked-up and checked, its use in this examination of PMs in SAE adds to the growing number of corpus-based studies of varieties of English, as well as studies of PMs in natural spoken discourse. Furthermore, the study also offers preliminary insights into *shame* as a PM unique to SAE.

## References

- Aijmer, K. and A.M. Simon-Vandenberg. 2009. Pragmatic markers. In J.-O. Östman and J. Verschueren (eds.) *Handbook of pragmatics: 2009 installment*. Amsterdam: John Benjamins. pp. 1-29. Available online: <http://benjamins.com/online/hop/> (Accessed 6 November 2012).
- Andersen, G. 2001. *Pragmatic markers of sociolinguistic variation: A Relevance-Theoretic approach to the language of adolescents*. Amsterdam: John Benjamins.

- Baker, P. 2007. *Using corpora in discourse analysis*. London: Continuum.
- Beach, W.A. 1993. Transitional regularities for 'casual' "Okay" usages. *Journal of Pragmatics* 19: 325-352.
- Bekker, I. 2009. The vowels of South African English. Unpublished doctoral dissertation. Potchefstroom: North-West University.
- Coll, M. 2009. 'Anyway': A formal approach to the syntax and semantics of Discourse markers. Unpublished doctoral thesis. Colchester: University of Essex
- Condon, S.L. 2001. Discourse *ok* revisited: Default organization in verbal interaction. *Journal of Pragmatics* 33: 491-513.
- Crystal, D. 2008. *By hook or by crook: A journey in search of English*. New York: Harper Perennial.
- Da Silva, A. 2007. South African English: A sociolinguistic investigation of an emerging variety. Unpublished doctoral dissertation. Johannesburg: University of Witwatersrand.
- De Klerk, V. 2005. Procedural meanings of *well* in a corpus of Xhosa English. *Journal of Pragmatics* 37: 1183-1205.
- De Klerk, V. and D. Gough. 2002. Black South African English. In R. Mesthrie (ed.) *Language in South Africa*. Cambridge: Cambridge University Press. pp. 356-380.
- De Klerk, V., R. Adendorff, M. de Vos, S. Hunt, S. Simango, L. Todd and T. Niesler. 2006. Educated mother-tongue South African English: A corpus approach. *Language Matters* 37(2): 206-226.
- Fairhurst, M. 2013. The pragmatic markers *anyway*, *okay* and *shame*: A comparative study of two African varieties of English. Unpublished MPhil thesis. Stellenbosch: Stellenbosch University.
- Ferrara, K. 1997. Form and function of the discourse marker *anyway*: implications for discourse analysis. *Linguistics* 35: 343-378
- Fraser, B. 1996. Pragmatic markers. *Pragmatics* 6(2): 167-190.
- Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics* 31: 931-952.
- Fraser, B. 2006. Towards a theory of discourse markers. In K. Fischer (ed.) *Approaches to discourse particles*. Amsterdam: Elsevier. pp. 189-204.
- Gaines, P. 2011. The multifunctionality of discourse operator *okay*: Evidence from a police interview. *Journal of Pragmatics* 43: 3291-3315.
- Greenbaum, S. 1996. Introducing ICE. In S. Greenbaum (ed.) *Comparing English worldwide: The International Corpus of English*. Oxford: Clarendon Press. pp. 3-12.

Gough, D. 1996. Black English in South Africa. In V. De Klerk (ed.) *Focus on South Africa*. Amsterdam: John Benjamins. pp. 53-78.

Hunston, S. 2002 *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Jeffery, C. 2003. On compiling a corpus of South African English. *Southern African Linguistics and Applied Language Studies* 21(4): 341-344.

Jeffery, C. and B. Van Rooy. 2004. Emphasiser *now* on colloquial South African English. *World Englishes*, 23(2): 269-280.

Kamwangamalu, N.K. 2006. South African Englishes. In B.B. Kachru, Y. Kachru and C.L. Nelson (eds.) *The handbook of World Englishes*. Oxford: Blackwell. pp. 158-171.

Kennedy, G. 1998. *An introduction to corpus linguistics*. London: Longman.

Lass, R. 2002. South African English. In R. Mesthrie (ed.) *Language in South Africa*. Cambridge: Cambridge University Press. pp. 104-126.

Levin, A. and D. Gray. 1983. The lecturer's OK. *America Speech* 58(3): 195-200.

Mesthrie, R. 1992. *English in language shift: The history, structure and sociolinguistics of South African Indian English*. Cambridge: Cambridge University Press.

Mesthrie, R. 2002. From second language to first language: Indian South African English. In R. Mesthrie (ed.) *Language in social history*. Cape Town: David Phillip. pp. 339-355.

Nelson, G. 2006. World Englishes and corpora studies. In B.B. Kachru, Y. Kachru and C.L. Nelson (eds.) *The handbook of World Englishes*. Oxford: Blackwell. pp. 733-750

Norricks, N.R. 2009. Pragmatic markers: Introduction. *Journal of Pragmatics* 41: 863-865.

*Oxford Dictionary of English*. 2nd Edition (Kindle Edition). 2010. Oxford: Oxford University Press.

Park, I. 2010. Marking an impasse: The use of *anyway* as a sequence-closing device. *Journal of Pragmatics* 42: 3283-3299.

Schiffrin, D. 1987. *Discourse markers*. Cambridge: Cambridge University Press.

Schourup, L. 1999. Tutorial overview: Discourse markers. *Lingua* 107: 227-265.

Scott, M. 2012. *WordSmith Tools*. Liverpool: Lexical Analysis Software.

Sinclair, J. 2004. *Trust the text: Language, corpus and discourse*. London: Routledge.

Stegmeier, J. 2012. Toward a computer-aided methodology for discourse analysis. *Stellenbosch Papers in Linguistics* 41: 91-114.