
Firming Up the Foundations: Reflections on Verifying the Quota- tions in a Historical Dictionary, with Reference to *A Dictionary of South African English on Historical Principles**

Sheila Hicks, *Dictionary Unit for South African English, Rhodes University,
Grahamstown, South Africa (s.hicks@ru.ac.za)*

Abstract: *A Dictionary of South African English on Historical Principles* (DSAEHist) is rooted in quotation evidence. It contains just over 8 000 South African English entries, with about 45 000 citations to support those words included as headwords in the dictionary. Using the legacy electronic format in which DSAEHist was typeset, the Dictionary Unit for South African English embarked on a digitising process of DSAEHist, during which it became clear that the quotations would benefit from a full review involving the verification of all quotations against their original sources. This article examines the evolution of the quotation verification project from its beginnings as an entirely manual exercise to its current use of software developed for the purpose. Some of the project's achievements, such as antedatings and primary source identification, are highlighted, and challenges, such as unverifiable quotations and sometimes highly convoluted research paths, are described. In addition to this, the article looks at the necessarily systematic nature of quotation handling and the main types of considerations determining methodology (for example, lexicographic, bibliographic and typographic requirements).

Keywords: DICTIONARY, SOUTH AFRICAN ENGLISH, HISTORICAL PRINCIPLES, VERIFICATION, QUOTATION, CITATION, ACCURACY

Opsomming: *Verstewiging van die fondamente: Gedagtes oor die kontro-
lering van aanhalings in 'n historiese woordeboek, met verwysing na A
Dictionary of South African English on Historical Principles.* *A Dictionary of South
African English on Historical Principles* (DSAEHist) het ontstaan uit aanhalingsbewysmateriaal. Dit
bevat net oor 8 000 inskrywings van Suid-Afrikaanse Engels, met ongeveer 45 000 sitate om daardie
woorde te staaf wat as trefwoorde in die woordeboek ingesluit is. Deur die argaiese elektroniese
formaat te gebruik waarin die DSAEHist geset is, het die Woordeboekeenhed van Suid-Afrikaanse

* This article was presented as a paper at the Fourteenth International Conference of the African Association for Lexicography, organised by the Xhosa Department, University of the Western Cape, Bellville, South Africa, 6–8 July 2009.

Engels 'n digitaliseringsproses van die DSAEHist onderneem waartydens die duidelik geword het dat die aanhalings sal baat by 'n volledige hersiening wat die kontrolering van alle aanhalings in hulle oorspronklike bronne behels. Die artikel ondersoek die vordering van die aanhalingskontroleringsprojek vanaf sy begin as 'n algehele handtaak tot sy huidige gebruik van programmatuur wat vir die doel ontwikkel is. 'n Aantal van die projek se suksesse soos vroeëre daterings en primêrebronidentifikasie word uitgelig, en uitdagings soos onkontroleerbare aanhalings en somtyds hoogs verstrengelde navorsingsroetes word beskryf. Hierbenewens kyk die artikel na die noodsaaklik sistematiese aard van aanhalingshantering en die hoofsoorte oorwegings wat metodologie (byvoorbeeld, leksikografiese, bibliografiese en tipografiese vereistes) bepaal.

Sleutelwoorde: WOORDEBOEK, SUID-AFRIKAANSE ENGELS, HISTORIESE BEGINSELS, KONTROLERING, AANHALING, SITAAT, NOUKEURIGHEID

Introduction

A Dictionary of South African English on Historical Principles (1996) was the product of twenty-five years' research, and is rooted in quotation evidence. It contains just over 8 000 South African English entries (including variant forms and unassimilated terms), each of which presupposes sufficient quotation evidence as a prerequisite for inclusion. These quotations are drawn from an extensive card index archive that was built between the late 1960s and early 1990s, providing a pool from which about 45 000 citations were selected to support those words finally included as headwords in the dictionary itself. Because it is a historical dictionary, citations feature prominently in the entry model, providing a set of concise illustrations of their parent headword. These short excerpts are taken from a variety of sources — diaries, letters, newspapers, magazines, novels, reference works, posters, radio, television, speeches, conversations — and not only contain historical and semantic information which it is not possible to represent neatly in a dictionary definition, but also played a key role in the early stages of lexical acquisition. By recording a quotation and indexing it by catchword — that is, the word as it appears in context rather than its canonical form — a potential headword would be noted or existing evidence for it augmented.

Using the legacy electronic format in which *A Dictionary of South African English on Historical Principles* (DSAEHist) was typeset, in 2006, the Dictionary Unit for South African English (DU) embarked on a process of digitisation of the published set of quotations. This involved both conversion into a future-proof, industry-standard format and the development of a searching and updated editing environment. There were two reasons for this digitisation process. The first, not central to this article, was to prepare the entire dictionary dataset for use in modern software. The second, shorter term motivation was to utilise the set of quotations published in DSAEHist as an electronic citations database in its own right. During this conversion process, it became clear that the resulting citations database would benefit from a full review involving the

verification of all quotations against their original sources. This time-consuming process, some stages of which are unavoidably manual, is now over 60% complete, and is on target for final completion by 2011.

This article examines the evolution of the quotation verification project from its beginnings as an entirely manual exercise, to its current use of software developed for the purpose. To provide the context for this process, the past and future purposes of the citations database for the DU are discussed, and some of the project's achievements, such as antedatings and primary source identification, are highlighted. Challenges encountered while trying to verify quotations are also described, such as unverifiable quotations and sometimes highly convoluted research paths. In addition to these, the article will look at the necessarily systematic nature of quotation handling and the main types of considerations determining methodology (for example, lexicographic, bibliographic and typographic requirements). Finally, the impact of the internet on quotation research is also considered, with some discussion of how this may affect the revision process for a new edition of DSAEHist.

Historical dictionaries

Firstly, some comments about diachronic and synchronic dictionaries.

Like the *Oxford English Dictionary* (OED), DSAEHist is a diachronic dictionary, that is, one that traces the origins and development of words through one or more periods in the history of the language. The main purpose of a diachronic dictionary is to trace the development of form and meaning over an extended period of time: the attention of a historical dictionary is directed to the evolution of meaning (Landau 2001: 130). A dictionary such as the OED "not only provides a historical record of the development of meaning of each word, with illustrative quotations and definitions for each sense, it also shows the changes in spelling, the different forms each word assumed during its history" (Landau 2001: 81). As Zgusta (2006: 126) puts it, "Historical dictionaries are ... synchronic descriptions of older historical epochs."

A synchronic dictionary, on the other hand, such as the *South African Concise Oxford Dictionary* (SACOD), focuses on contemporary language, and attempts to represent the lexicon as it exists (or existed) at a particular point in time — generally, but not necessarily, the present — so etymology is less important than in diachronic dictionaries (Landau 2001: 27).

Hartmann and James (1998: 68) describe a historical dictionary as "a type of reference work in which the vocabulary of a language is traced through time". Whereas general dictionaries provide information on the derivation of words — etymology — under the influence of the historical-comparative perspective in linguistics, historical dictionaries trace the formal and semantic changes in the vocabulary of a language throughout its history or in a particular period, documenting the changes in form and meaning of words (the 'curriculum vitae of vocabulary').

The general objectives of a historical dictionary are described by James Murray in the aims of the *Oxford English Dictionary*:

The aim of this Dictionary is (1) to show, with regard to each individual word, when, how, in what shape, and with what signification, it became English, what development of form and meaning it has since received; which of its uses have, in the course of time, become obsolete, and which still survive; what new uses have since arisen and when: (2) to illustrate these facts by a series of quotations ranging from the first known occurrence of the word to the latest, or down to the present day; the word being thus made to exhibit its own history and meaning: and (3) to treat the etymology of each word.

In dictionaries based on historical principles the ordering of homonyms normally displays the historical ordering, and the lemma sign representing the oldest lexical item is presented as the first homonym (Gouws and Prinsloo 2005: 100): the ordering represents the historical development of the various senses (Kastovsky 2000: 119), and there is therefore a challenge for the editors in identifying fine subdivisions of senses (Coleman and McDermott 2004: v). At the Dictionary of Old English project the evidence of citations is used as the primary guide to sense-division (Coleman and McDermott 2004: 138).

In most synchronic dictionaries focusing on the language as it is currently used, the ordering is usually determined by the usage frequency of the words and the word with the highest usage frequency will be presented as the first homonym (Gouws and Prinsloo 2005: 100).

Etymology is an integral part of the work of a historical dictionary, since it illustrates how form and meaning have changed over an extended period of time, often centuries (Landau 2001: 27). The logical requirement of starting with the etymological (or derivational) meaning is much stronger in diachronic dictionaries than in synchronic dictionaries (Zgusta 2006: 29), and DSAEHist uses labels such as obsolete, obsolescent and historical to elucidate this.

Zgusta (2006: 3) points out that "a historical dictionary can offer ... a description of a ... stage in the development of [a] language" or it could be "concerned with the development of a language in the past only, not focusing on or even not considering at all the present state of that language. ... Thus, the expression historical dictionary is used in reference both to period dictionaries, and to diachronic dictionaries, either of them situated on various points or stretches of the flow of time, or development. That many, perhaps even most, dictionaries simultaneously consist of components that belong to different types may go without saying".

Finally, a comment about the OED (<http://www.oed.com/about/>):

As the OED is a historical dictionary, its entry structure is very different from that of a dictionary of current English, in which only present-day senses are covered, and in which the most common meanings or senses are described first. For each word in the OED, the various groupings of senses are dealt with in

chronological order according to the quotation evidence, i.e. the senses with the earliest quotations appear first, and the senses which have developed more recently appear further down the entry. In a complex entry with many strands, the development over time can be seen in a structure with several 'branches'.

Sources of quotations

The choice of quotation sources tells us as much about the values and reading preferences of the editors/readers/lexicographers as they do about the language itself (Brewer 2007: 186) and reinforces, whether consciously or not, a particular set of cultural values (Brewer 2007: 189). Charlotte Brewer describes the OED as a "cultural treasure-house, simultaneously testifying to the longevity of the language, the glories of the English tongue, and the scholarship of the Dictionary's editors" (Brewer 2007: 4). Rowena Fowler points out that it is "possible to read the *Dictionary* [Richardson's *New Dictionary of the English Language* (1836–7)] as a historical document, quarrying it for evidence of early nineteenth-century knowledge, attitudes and mentalities" (Fowler 2004: 54). She adds: "There is a contradiction built into dictionaries claiming to use quotations from the 'best' or 'most distinguished' writers, since these writers may not be typical, characteristic or otherwise exemplary from the lexicographical point of view. The question cuts two ways: can poetic language provide reliable illustration of meaning, and can or should the dictionary provide adequate treatment of literary usage?" (Fowler 2004: 55). Fowler also stresses the lexical dubiousness of drawing on fictional dialogue as evidence of spoken usage (Fowler 2004: 57). However, while fictional dialogue provided printed evidence of use at the time of publication, historical novels should be treated with caution as they might fallaciously bring archaic usages into modern times.

It is clearly important to avoid hapax legomena — words which occur only once in either the written record of a language, the works of an author, or in a single text — and to get a wide spread of sources. In his discussion of Samuel Johnson's *A Dictionary of the English Language*, Robert Burchfield comments that while "the central body of his [Johnson's] illustrative quotations was drawn from the greatest literary, philosophical, and scientific works of the period from Sidney to the Restoration, that is from the 1580s to 1660 ... the quotations form a kind of concordance to the most memorable utterances of his day" (Burchfield and Simpson 2002: 85-86). Richardson had a slightly different approach, "based on the notion that quotations alone, if sufficient in number, could serve to elucidate 'true etymological meaning'", reinforcing his theory that each word had a single immutable meaning (Landau 2001: 77-78).

"In historical dictionaries they (quotations) can show the type of source in which the term is recorded (formal literature, textbooks, magazines, science fiction, etc.). They can show typical (and sometimes eccentric) contexts in which the term has been used, and in a chronological sweep of evidence it is often possible to obtain a telescopic panorama of semantic and structural shift

through which a term has passed over the centuries" (Simpson 2003: 269). While quotations in synchronic and learners' dictionaries can be "invented" examples of usage, invented examples should not be used in historical dictionaries.

With the OED, "Contributors were asked to analyse carefully 'the works of any of the principal writers, extracting all remarkable words, and all passages which contain definitions or explanations' ... The lists also make it clear that writers who use language in markedly idiosyncratic ways were not to be excluded" (Brewer 2007: 124). However, a "reliance on literary quotations is problematic, because it skews the representative character of the sampling" (Brewer 2007: 125).

Mooijaart (2004: 202) lists three important factors that need to be considered when choosing citations:

- (1) the quality of the citations themselves: length, reliability and correctness of the text;
- (2) the accessibility and usability of the electronic citations and the non-linguistic information; and
- (3) the quality of the citations as a corpus: issues of representativeness and diversity.

Mooijaart (2004: 202-203) also points out some of the possible functions of citations:

- As lexicographic evidence
- As "examples of good taste"
- To exemplify meaning
- To exemplify meaning but also to illustrate the variety of grammatical context, connotation and register

A basic principle, however, is that quotations should generally describe and explain linguistic conventions, rather than trying to account for every individual language event (Atkins and Rundell 2008: 48).

In some cases special evidence is taken from reference works such as dictionaries. However, with some exceptions, Mooijaart (2004: 203) feels that, with some exceptions, these are not appropriate sources for illustrative citations as most of them offer invented phrases, or abbreviated phrases with no proper bibliographical documentation of sources.

How DSAE originally collected citations

At the commencement of the DSAEHist project, no electronic corpora were available, and so the process similar to that employed by Murray for the OED was followed: citations were glued on to cards, or handwritten, or typed. Lan-

dau points out that "vast citation files still exist only in paper form because some of the citations cannot be scanned (the type is too faint, or too large, or the citation is handwritten, or there are important handwritten notes attached, or the paper is too fragile, etc.) and because keyboarding would be prohibitively expensive" (Landau 2001: 190). This is certainly the case with DSAEHist: at a rough guesstimate, the DU currently has over 300 000 cards from which quotations for the dictionary were extracted.

Digitisation of DSAEHist

DSAEHist was produced before the advent of dictionary publishing systems, in Wordperfect 5. Before publication these Wordperfect files were sent via FTP to the publisher, the Oxford University Press in Oxford, and then to the US for semi-automated conversion to SGML (a textual format which is the pre-decisor to XML) to be used for typesetting. This SGML file was also the basis of the new XML version to which DSAEHist was converted in 2005. The quotations were converted, but the rest of the dictionary was only partly converted because of other work pressures, and this therefore became low priority.

Another key benefit of the conversion work, conducted with a good text editor, pattern matching, and custom text-processing utilities developed in-house, is that inconsistencies in the content (including quotations) could be identified and either highlighted for manual attention or fixed globally. The markup used is standards-compliant, transparent markup which is not bound to any particular software: Serna has been most useful so far and so that is what has been used.

The digitisation of the dictionary made it easier to analyse the dictionary and produce statistics such as which authors or newspapers were most frequently quoted and whether there was a good geographic spread across the country (unsurprisingly, there was a heavy bias towards the Eastern Cape, which is where the Dictionary Unit is located), what were the earliest quotations, distribution of oral informants (again, there was a heavy bias towards the Eastern Cape), analysis of chronological range, and so on.

DSAEHist style guide

An extensive style guide was developed by the editors of DSAEHist. This was based largely on the style of OED, but with some modifications based on local requirements. This lengthy document covers matters such as the layout, order and numbering of entries; how titles, authors, page numbers and so on are styled in the dictionary; use and treatment of proprietary names, prefixes and suffixes; parts of speech; labels for subject, register and frequency; etymology; combinations, combinations and compounds; bibliography; abbreviations; spelling (use of *-ize* rather than *-ise* forms); punctuation — and much more. Some of these aspects can be handled by a style sheet within the software.

Database of quotations

It may be useful to distinguish briefly between the digitised version of the dictionary and the database of quotations (the terms 'citation' and 'quotation' are used interchangeably in this article).

As described above, the database existed as SGML quotations, which needed conversion and integration with Serna and the development of a search interface. The conversion was done using in-house text-conversion skills and utilities, as well as commercial text-processing software, in a sort of semi-automated fashion, part manual, part automated. It was a very slow, painstaking process with many safety mechanisms.

The quotations are of course a component of the entire dictionary, and the quotations were simply extracted and treated as a separate database, to be edited on its own. Whether the reviewed quotations are treated separately or as part of DSAEHist is a matter of choice and convenience. The way the markup and styling have been approached means that the typography and styling rules of DSAEHist have for the most part been preserved, unless there has been a conscious decision to change them. Some changes present challenges, such as cases where a quotation date was modified and this was integral to the entry (e.g. the definition referred to it), but these are minor. In a sense the review of quotations was the beginning of the revision of the DSAEHist.

A separate database was created for bibliographical information (see the section **Bibliography** below).

Extracts in the early days of the verification process were less sophisticated, as the database was still in the process of being developed and refined, tags redefined, and so on. So the early extracts did not, for instance, include the headword, or the quotation number, or distinguish between monographs and serials. Now, with the online search system using a web browser, the extracts produced are much clearer and easier to work with: they do not include the tags, they include both old and new quotation IDs, and because of the layout, it is much easier to notice inconsistencies of styling.

Purposes of the citations database

Some of the purposes for which the citations database may be used in future include:

- (1) Integration with a larger database for use in future quotation research
The idea was for the DSAEHist quotations database to be integrated with a new 'intake' (or, as called by the OED, 'incoming') database to provide one searchable, growing set of quotations. New intake will have different emphasis: recent citations for the same words as well as new words, use of internet resources etc. After the verification stage, post-dating quotations for DSAEHist entries could be approached systematically, with use of the internet, etc.

- (2) Integration with DSAEHist editing environment
Quotation handling is an important component of DSAEHist editing, technically as well as lexicographically. The new quotation markup allows software to do automatic sorting and styling of quotations in dictionary entries, generally removing the burden of manual editorial tasks.
- (3) Integration with bibliography
Quotations are OED-style linked with entries in the bibliography entries.
- (4) Semi-automated proofing and statistical checking
The markup makes it possible to identify imbalances in date range, source type coverage and so on with customized queries.
- (5) There are also other possibilities such as linking quotations to library records and internet-published resources, and ranking quotations (e.g. isolating quotations from oral or written informants as opposed to printed quotations, or 'informational' quotations).

Bibliography

The bibliography was converted to a text editor, which is maintained on an ongoing basis as verification of individual titles is completed. The bibliography printed in DSAEHist was of necessity limited — publishers are always concerned about space constraints. But in an electronic file not only is it easier to have more data, it is also possible to record information which may not necessarily be part of the final bibliography available to users of the dictionary, but is extremely useful to the editors as they work. In other words the markup allows custom views (e.g. published and unpublished views) determined by purpose and context.

For example:

- Information regarding the location of titles. It can sometimes take quite a bit of detective work to find out where a title is held, and once it has been tracked down it is useful to record this for future reference. This information includes the Dewey Decimal Classification number as well as the name of the repository.
- Name of publisher. In some cases, there is more than one edition of a title, published in the same year or in different years, with different pagination, and clearly this information is necessary for a reader who wants to look at the original title from which a quotation was taken. This information was not included in the printed bibliography in DSAEHist possibly because of space constraints.
- First names of authors. This is essential for authors such as Margaret Roberts and Michael Roberts, but useful for all authors. If a reader wants

to look up an author called, say, M. Roberts, there may only be one M. Roberts in the DSAEHist bibliography, but many on a library catalogue.

- Additional information about the title, such as the full title as well as the abbreviated 'short' title used in citations. There may be more than one version of the short title: for example, where a diary has been edited and quotations are in some cases taken from the original author, and in other cases from the words of the editor (for example, notes, footnotes or editorial comments):

1835 A. SMITH *Diary* (1940) II. 62 After waiting about ten minutes we were apprised by the exclamations 'Byat! Byat!' that he (i.e. Mazilikatze) was approaching from an inner enclosure.

1940 P.R. KIRBY *Diary of Dr Andrew Smith* II. 72 Bayede, Hail, King.

1940 P.R. KIRBY *Diary of Dr Andrew Smith* II. 72 Bayede, Hail, King. The Matabele royal salute.

The first Kirby quotation is as it appears in DSAEHist, the second version is as it has been finalised after verification, with additional relevant information.

- Notes, such as who has performed the verification, and any other useful information.

Textual accuracy

John Simpson, Chief Editor of the OED, points out in the Preface to the Third Edition (<http://dictionary.oed.com/about/oed3-preface/accuracy.html>): "It is important that quotations represent the texts as originally published, and many thousands of texts are being re-examined and the results fed ... on to the online database."

Simpson also stresses the importance of accurate transcriptions from original texts (Burchfield and Simpson 2002: 180):

It is only by making accurate material available that reliable linguistic analysis can be made. The first edition of the *Oxford English Dictionary*, for example, contains in its illustrative quotations seventeen early instances of the elided form *it's* (= 'it is') from sixteenth-century sources (1575–98). In many cases these quotations were cited from later editions of texts in which the editors had partially modernized the spelling to simplify matters for their readers. But anyone investigating the emergence of forms such as *it's* would be ill-served by relying on this evidence, since most of the examples — especially from the early part of the sixteenth century — in fact do not show any elision.

The Examining the OED project at Hertford College points out that searching the OED is sometimes difficult (<http://oed.hertford.ox.ac.uk/main/content/category/11/43/161/>):

Authors' names, works, and dates of publication are not always consistently rendered in OED. Thus in the first edition of OED, the work by Carlyle on the French Revolution was cited in at least five different forms, while in the second edition, volumes of poems by Auden were usually referred to by their UK titles, but also on occasion by their American titles, with dates different from the UK equivalents. The OED is currently standardising the way in which quotation sources are identified in OED Online, in a huge sub-project of the revision process.

Similar problems have occurred in DSAEHist, and an example of this is given in the section **Examples of improvements** below.

The policy in DSAEHist was to retain the original spelling (but not orthography — the dictionary uses the modern terminal or short 's' rather than what is known as the long, medial or descending 'f'), lineation, and punctuation, except that double quotation marks have been replaced by single quotation marks.

The traditional editorial practice of large historical dictionaries was for selected citations to form complete phrases or full sentences with respect to syntax. Although words or phrases within the citation which were not relevant for the word to be illustrated were often skipped and replaced by ellipses, the basic structure of the sentence was maintained. This implies that all parts of speech are represented, but in the citations they will however have a less complex structure than they have in texts. This was frequently the practice with DSAEHist.

Dating of quotations in DSAEHist

The policy with regard to dating of quotations is based on and similar to that of the OED (see <http://www.oed.com/archive/oed2-preface/gen-main-5.html>). Quotations are normally taken from the earliest available printing of a work; where a later text has been used the publication date is given in parentheses after the title. In DSAEHist, quotations from Pettman (1913), Jeffreys (1964, 1967, 1970) and Swart (1934), three secondary sources who wrote extensively on South African English, were not verified against primary sources but had an attribution in parentheses after the citation. During the present verification process, it has been possible to verify a considerable number of these against primary sources. Where quotations have been taken from more than one edition of a work, the edition used is given in parentheses after the title (this practice is followed where a particular quotation is only present in a later edition).

Verification process

The verification process consists basically of three steps:

- Verification

- Data capture
- Finalisation

(1) Verification

With DSAEHist, as with OED, verification takes place principally by printing out all of the quotations from a particular source — title or author — and comparing these against the appropriate edition of the source text, making amendments where necessary.

(2) Data capture

The next step is for a keyboarder to locate each quotation via an interface programme running under a web browser (Firefox), and to edit it in a specially designed text editing interface (running in Serna). Any changes must be recorded, as well as the initials of the verifier and the data capturer/keyboarder. There is a facility for annotations and queries, which has proved very useful.

(3) Finalisation

The final step is for the original verifier to compare the annotations on the printouts against the updated quotation on the database to ensure that the keyboarder has correctly interpreted handwriting and notes, and correctly captured all changes.

During the verification process, it has been possible to verify a number of quotations against a primary source or an earlier edition — about 2 100 quotations so far.

It has also been possible to antedate a small number of quotations. However, these statistics are not definitive: Because the researchers are looking at quotations from a particular source, rather than quotations for a particular lemma, it would be too time-consuming, and not productive simply for the sake of statistics, to consider each lemma to see whether verifying against an earlier edition would result in an antedating.

Quotation verification system

The quotation verification system (QVS), accessed through a web browser, has options to search in a number of different key fields, such as full text, quotation text, author, title, headword and quotation index. The search interface will compare against any of whole word, whole field, start of word or 'anywhere', and also allows for multiple combinations of search options.

An alternative search option is to filter by verification status. Here one may select from a drop-down box the initials of any of the staff members working on the system, or 'somebody', 'nobody' or 'unfiltered', in each of three fields: Verifier, Capturer and Finaliser. This is useful if, for example, a verifier wants to see which of the verified quotations have been captured and are

waiting to be finalised. It is also useful for extracting statistics, to show how many quotations have been captured and finalised. It does not, obviously, show just how many quotations have been verified, because they are only recorded as being verified at the same time that they are captured, so these two fields will show the same count. The statistics for quotations verified are thus a manual count, maintained on an Excel spreadsheet.

All the verification has been done by one of the editors, Sheila Hicks, with the assistance of a researcher and three student assistants.

Sources for verification

The DSAE has a small library, containing just over 1 600 monographs and about 40 periodical titles. However, we are very fortunate in being able to use the Rhodes University libraries. Some of the quotations were verified against titles in the DSAE library, but a large number came from the Rhodes Library (and the Inter-Library Loans system), the Cory Library for Historical Research, and the National English Literary Museum, as well as the Schönland Herbarium, the Albany Museum, and Rhodes departmental libraries such as Classics, Education, Ichthyology, the International Library of African Music (ILAM), Law, Music and Zoology.

The internet is another useful resource which was not available to the original editors of the dictionary. Sometimes it has been possible to access the 'full view' of a work on Google Books, at other times a 'limited preview' or 'snippet view' has enabled the researcher to request the relevant page numbers through Inter-Library Loans (ILL). In some cases, searching for a particular phrase on Google Books has revealed that the wrong title was cited in DSAE-Hist. Other electronic resources include the Internet Archive, Project Gutenberg, and databases such as JSTOR (available through the Rhodes library system).

When using the ILL system, it has been much easier to locate monographs than periodicals. A monograph title can be requested through ILL, and in cases where the holding library is not prepared to lend the book itself, it is often prepared to make photocopies or scans of the relevant pages, so a list of page numbers is sent (this usually works well, although sometimes additional pages have to be requested because quotations run over from one page to another). However, with periodicals, specific page numbers have to be requested, preferably with the title of the article (which is almost never known) and the author (which is known in only a few instances, because it was often not included in the quotations). For periodicals such as the *Cape Times*, from which the dictionary has 597 quotations, or *S.A. Panorama*, from which the dictionary has 556 quotations, it is clearly not practical to fill out a form for each quotation. For *S.A. Panorama*, the Nelson Mandela Metropolitan University (NMMU) in Port Elizabeth had good holdings of this title, and it was possible to verify the issues which were not held by Cory at NMMU. For other titles, it is necessary to make

several trips to libraries which have extensive holdings such as the National Library of South Africa's 'Places of Legal Deposit' in Cape Town or the Msunduzi Municipal Library in Pietermaritzburg.

What does the QVS system give us?

Because the system was home-grown, it could be tailored to meet our specific requirements, often with modifications being made along the way as the need for a particular facility became apparent. First of all, because the different categories require different information, the quotations were divided into certain basic categories — monographs, serials, ephemera, manuscript, broadcast or informant. For example, a serial could have a volume and part number, or a day and/or month of publication, and the formatting of the quotation in the dictionary is different from that of a monograph, because a monograph will have an author. For example:

1959 E. MPHAHLELE *Down Second Ave* 188 It was later in 1955 that I joined the African National Congress (ANC). I had for some years been torn between it and the All-African Convention (AAC).

whereas a serial may or may not have a 'nested' author, who is the author of the article but not of the whole publication:

1941 C.A. KUSE in *Bantu World* 1 Mar. 4 We will have no constituted cooperation with political organisations like the A.A.C.

1932 *Grocott's Daily Mail* 9 Jan. 3 The flowers which form the subject of these delightful fantasies are Flames, Morning Flowers, Aandbloem.

Facilities available in QVS include:

- Drop-down boxes listing possible values for certain fields, such as 'Content type', where the options are advertisement, caption, cartoon, heading, legend, letter or title. Without such a system, there was plenty of scope for inconsistent values to be used.
- Facility to add a 'minor title', such as the title of a short story. Although this is not printed in DSAEHist, it is particularly useful in, for example, an anthology or collection of short stories, where the minor title can enable the researcher to track down the first date of publication and verify the quotation against the first printing.
- Ability to insert special characters, such as non-breaking spaces (for example, what looks like a space in '65 372', to prevent the numbers being broken over a line-end), or special symbols used in original texts, for example to indicate pronunciation (i.e. not true phonetics) such as

1919 H.H. JOHNSTON *Compar. Study of Bantu & Semi-Bantu Lang.* I. 797 The divergent dialects of ?ōsa, such as Isi-pondw, Isi-baʔa, Feñgu, &c.

-
- Facility to add line-break indicators in poetry. Poetry often uses a style of language diverging from that of ordinary speech, and it may be useful to show this. Some readers indicated line-breaks by putting commas at the end of lines and/or beginning new lines with a capital letter, but these may or may not have been there in the original text.
 - Facility to add page-break indicators. Where a quotation runs over two pages, this was inconsistently treated in DSAEHist: sometimes the page number on which the quotation begins was given, and sometimes the page number on which the headword occurs. This facility makes it easier to spot such inconsistencies.
 - Facility to record column numbers in periodicals. Trying to find a quotation in an old number of a broadsheet newspaper, typically using a very small font, can be quite difficult, but it is much easier if the column number is given.
 - Accurate recording of fonts. In DSAEHist, fonts were typically 'normalized', with emphasis such as bold font, small capitals, italics or underlining downgraded or removed. With QVS, it is possible to record fonts accurately, and leave editorial decisions about styling to be made at the time of publication.

Extracting by author makes it easier to spot duplications. For example, here is an excerpt from the entry for **sitkamer**:

1908 F.C. SLATER *Sunburnt South* 12 The room into which he showed me was the *zit-kamer* or sitting-room; it was scrupulously neat and tidy. 1909 F. MASEY in *State* Vol.2 No.7, 67 In the 'zitkamer', .. there stands that commonly met feature the curio cabinet. 1912 F. BANCROFT *Veldt Dwellers* 31 It was smoke-room, bar room, and general zit-kamer combined. 1927 F.C. SLATER in *Outspan* 1 Apr. 5 The room into which he showed me was the zit-kamer or sitting-room; it was scrupulously neat and tidy.

It will be noticed that the quotations dated 1908 and 1927 are the same, but the 1927 one was taken from a secondary source. Here's another example for the same author, from the entry for **place**:

1908 F.C. SLATER *Sunburnt South* 139 He had recently met me at a neighbouring farm and, finding in me an attentive and admiring auditor, had cordially invited me to visit his 'place'. 1929 W.M. MACMILLAN *Bantu, Boer & Briton* 200 (Swart), A great number of the places they deserted are now used as cattle places by proprietors who do not reside upon them but leave them and their cattle in charge of freedmen, Bechuanas and Bushmen. 1931 F.C. SLATER *Secret Veld* 219 He had recently met me at a neighbouring farm, and, finding in me an attentive and admiring auditor, had cordially invited me to visit his 'place'.

Here the incorrect title (and date) had been given for the 1931 quotation.

Examples of the verification of the primary source

At **berg**:

1944 J. MOCKFORD *Here Are S. Africans* 71 The total number of wagons that crossed the Berg during the Great Trek is estimated at more than one thousand.

After verification:

1937 M. NATHAN *Voortrekkers* 177 The total number of wagons that crossed the Berg during the Great Trek is estimated at more than one thousand.

At **sammy**:

1949 J. MOCKFORD *Golden Land* 179 Sammy's baskets balanced on the end of a bamboo pole, like the nests of a weaver bird, go bobbing along the streets and lanes of every township in Natal.

After verification:

1921 C. DAWBARN in *My S. Afr. Yr* 214 Sammy's baskets balanced on the end of a bamboo pole, like the nests of a weaver bird, go bobbing along the streets and lanes of every township in Natal.

In both cases, the primary source was listed in Mockford's books as a footnote, with no page number, so it took some detective work and speed-reading to find the quotation in the primary source.

Here is another example:

blue-coat *n.* *Prison slang.* [tr. Afk. *bloubaadjie*, see BLOUBAADJIE.]

1. A long-term male prisoner serving an indeterminate sentence, usu. after having been declared an habitual criminal; BLOUBAADJIE sense 2 b. See also BLOUROKKIE sense 1.

1949 H.C. BOSMAN *Cold Stone Jug* (1969) 12 In prison the blue-coat occupies a position of some degree of importance. A blue-coat is even higher than a murderer. 1976 V. ROSENBERG *Sunflower* 57 In prison, the murderer, unlike the blue-coat, does not wear a distinctive garb. He is not dressed by the authorities in a way to single him out from the other convicts.

2. An indeterminate prison sentence; BLUE-JACKET; COAT; cf. BAADJIE sense 2.

1949 H.C. BOSMAN *Cold Stone Jug* (1969) 11 He goes back to prison to serve the indeterminate sentence all over again. As they say in prison, he goes back to do his second blue-coat. 1975 *Sunday Times* 15 June 3 If I give myself up, I want to know that I'm not going to get a bluecoat (indeterminate sentence).

Careful examination of Rosenberg's *Sunflower to the Sun*, and some speed-reading of Bosman's *Cold Stone Jug*, reveals that the 1976 quotation should be:

1949 H.C. BOSMAN *Cold Stone Jug* 8 In prison, the murderer, unlike the blue-coat, does not wear a distinctive garb. He is not dressed by the authorities in a way to single him out from the other convicts.

However, this means that both the two quotations given to support the first sense of **blue-coat** are taken from Bosman (as well as one of the two quotations

for the second sense). This is not an adequate spread of sources, and so the editors will have to look carefully at what other quotations are available for this lemma.

A number of quotations in DSAEHist were taken from OED: For example, this one at **sassaby**:

1820 S. DANIELL *Sk. S. Africa* 18 The Sasayby is an Antelope, heretofore not described, found in the Booshwana country.

which in DSAEHist became

1820 S. DANIELL *Sketches S. Afr.* 18 The Sasaybe is an Antelope .. found in the Booshwana country.

It will be noticed that although the title has been changed from the way it was styled in OED ('Sk.' has become 'Sketches'), it still does not conform to DSAEHist styling, which is to include prepositions. Also, as the full title of the work is *Sketches Representing the Native Tribes, Animals, and Scenery of Southern Africa from Drawings Made by the Late Mr. Samuel Daniell, Engraved by William Daniell*, the abbreviation should be 'Sn' not 'S.'. It will also be noticed that the spelling 'Sasayby' in OED has been changed to 'Sasaybe' in DSAEHist. It was possible to obtain the relevant pages of the primary source, and the revised styling of this quotation will now look like this:

1820 S. & W. DANIELL *Sketches .. of Sn Afr.* 18 The Sasayby is an Antelope .. found in the Booshwana country.

Examples of antedating

African Hunting from Natal to the Zambesi, Including Lake Ngami, the Kalahari Desert, etc., from 1852 to 1860 by William Charles Baldwin was first published in 1863. The Cory Library for Historical Research holds the second edition, published in 1894, as well as a facsimile reprint of the first edition. Of the 24 quotations from this work in DSAEHist, nineteen were dated 1863 and five were dated 1894. By checking the first edition for these five, it was possible to antedate the first recording of **inyoka** by three years from 1866 to 1863:

inyoka /in'jɔ(:)ka/ *n.* Also **nhoca**, **nyoka**. [Xhosa and Zulu.] A snake. Also *attrib.*

1866 W.C. HOLDEN *Past & Future* 300 This was the highest kind of sacred inyoka, or 'serpent'. 1891 R. MONTEIRO *Delagoa Bay* 114, I described the noise to Jack in the morning, and he at once said it was made by a 'nhoca' snake as thick as his arm. 1894 W.C. BALDWIN *Afr. Hunting* 112 One of the Amatongas .. gave a most unearthly howl, .. saying that an inyoka snake had bitten him. 1937 [see IZINYANYA]. 1958 R. COLLINS *Impassioned Wind* 60 Inyoka! Snake! Where's the snake?

and of **kaalkop** by 31 years from 1894 to 1863:

2. Special Comb. [...]; **kaalkop** /-kɔp/ [Afk., kop] head, a bald person; also *fig.* and *transf.*; [...]

[...] **1894** W.C. BALDWIN *Afr. Hunting* 301 The bush was very good, a moderate breeze of wind, which I kept always below, but I had great difficulty in getting the bull out from the company of the 'carl kop' (naked head). **1896** R. WALLACE *Farming Indust. of Cape Col.* Kaalkop wheat, a beardless variety well liked by millers. **c1936** S. & E. *Afr. Yr Bk & Guide* 1022 On the silver coins of the Union, the King's head is crowned. The lack of this in British Silver has earned them the name of 'Kaalkop,' bare head. [...]

Here is a quotation for **half-cord** taken from Sir Francis Fleming's *Southern Africa: A Geography and Natural History of the Country, Colonies and Inhabitants from the Cape of Good Hope to Angola*, published in 1856:

1856 F.P. FLEMING *Sn Afr.* 480 *Scomber Capensis, Cuv. and Val. (Halfcord).* .. A large fish measuring from two to three feet.

Fleming's book includes as an appendix Ludwig Pappe's *Synopsis of the Edible Fishes at the Cape of Good Hope*, published in 1853, so it was easy to go to the primary source, which was held at the library of the South African Institute for Aquatic Biodiversity, and verify the quotation there, which resulted in an ante-dating for **half-cord** from 1856 to 1853:

1853 L. PAPPE *Synopsis of Edible Fishes* 23 *Scomber Capensis, Cuv. and Val. (Halfcord).* .. A large fish, measuring from two to three feet.

Examples of mistakes

- OED often gave chapter numbers (in roman numerals) instead of page numbers, and the earlier DSAEHist editors sometimes interpreted these as roman page numbers. For example:

1896 BADEN-POWELL *Matabele Campaign* iv,

in OED was interpreted in DSAEHist as

1896 R.S.S. BADEN-POWELL *Matabele Campaign* p.iv.

This has now been corrected to

1897 R.S.S. BADEN-POWELL *Matabele Campaign* 97.

It will be noticed that the date of the quotation has also been changed: Although the full title of the work is *The Matabele Campaign 1896: Being a Narrative of the Campaign in Suppressing the Native Rising in Matabeleland and Mashonaland*, it was only published in 1897. (DSAEHist has seven other quotations from this title, which are all dated 1897.)

- Page number missing or incorrect in DSAEHist. There could be several reasons for this. Sometimes the keyboarder has transposed the numbers: '298' has been typed as '289'. Sometimes it was because two quotations had been presented on one card, and the keyboarder had captured the same page number for both quotations. Sometimes it has simply been

left out by the keyboarder. Sometimes it has been mistranscribed due to unclear handwriting: '40' has been read as '46'. And sometimes the page number was incorrect due to a page break (as mentioned in the section **What does the QVS system give us?** above).

Sometimes the page number can be found by looking at the cards. However, in some cases there is no page number on the card, or the card is missing. Then some detective work is required. Google Books has proved very useful in tracking down some of these: Even if there is only a snippet view, it may give the page number. In other cases, it is necessary to speed-read or scan the book.

— Volume number missing or wrong.

— Wrong title. This may be a simple typing error, as at **rhenosterbos**:

1941 C.W. DE KIEWIET *Hist. of. S. Afr.* 189 Tumbleweed, prickly pear, rhenosterbos, and jointed cactus invaded the territory of the edible grass and nutritious plants.

where the full stop after 'of' should be deleted.

Or the title may have changed. For example, at **dolf**, the title of the 1st edition was *South West Africa* but the title of the 2nd edition was *Story of Namibia*. Quotations are dated 1976 (1st edition) but the title was taken from the 2nd edition (1978). So

1976 O. LEVINSON *Story of Namibia* 64 Kavango was the most fortunate of the homelands as far as water was concerned ... Timber is a promising industry with the beautiful 'dolf' and 'usivi' (chivi) trees.

should be

1976 O. LEVINSON *S.W. Afr.* 64 Kavango was the most fortunate of the homelands as far as water was concerned ... Timber is a promising industry with the beautiful 'dolf' and 'usivi' (chivi) trees.

Another cause of wrong titles being recorded is that the reader has simply written the author and date on the card, without a title, and the incorrect title has been used. For instance, Keppel-Jones, 1948, p. 60 became:

1948 A. KEPPEL-JONES *When Smuts Goes* 60 The 'Boer with his roer' — his old muzzle-loader — meant security from the old attacks.

This quotation actually comes from Keppel-Jones's *South Africa: A Short History*, published in 1949.

— Wrong font, such as italics instead of roman, or roman instead of italics. One reason for this could be that the original reader underlined the headword on the card in order to highlight it, and of course the conventional proofreading symbol for italics is to underline. Another reason is that it was not possible to capture fonts such as italics, bold, or small

capitals. And a third reason is that bold was used in DSAEHist as a lemma marker, to highlight the first entry in a section of 'special combinations', for example at **protea**:

1962 S. ELIOVSON *Discovering Wild Flowers in Sn Afr.* 50 Often the size of a dinner plate, the flower-head of the **King Protea** is the most spectacular of a large genus.

With the online version, it is possible to indicate this entry differently, so is it now represented thus:

1962 S. ELIOVSON *Discovering Wild Flowers in Sn Afr.* 50 Often the size of a dinner plate, the flower-head of the King Protea is the most spectacular of a large genus.

and the original font preserved. It may be decided in the future to revise the way in which this is indicated: OED online separates quotations for different senses by a horizontal rule, which enables the typestyle of the original source to be preserved.

— Correct punctuation and capitalisation.

— Correct word or spelling, which may or may not change meaning, and may affect the list of variant spellings given for the lemma. Examples include: at **biliary fever** (change 'in' to 'is'), at **looper** (change 'the' to 'they'); at **impala** ('Frangipangi' should be 'Frangipani'); at **Volksraad** ('1938' should be '1838'); at **outspan** ('wanted' should be 'wished'); at **khoisan** ('originated' should be 'diverged'); at **Bantu** ('passionately' should be 'dispassionately', correctly cited at **native**); at **Afrikanerdom** ('unwanted' should be 'unwonted'); at **mealie** ('mealie' should be 'mie-lie'); at **veldskoen** ('feldtschoon' should be 'feldt schoon'); at **liberal** ('negrophile' should be 'negrophilist'); at **saltriem** ('brought' should be 'bought'); at **tessies** ('tessies' should be 'tassies'); at **Uhlanga** ('distinct' should be 'indistinct'); at **werf** ('moved' should be 'removed'); at **pondokkie** ('packing' should be 'sacking'); at **native** ('recommended' should be 'recommenced'); at **stoker** ('decision' should be 'celebration'); and at **rand** ('1918–24 conflict' should be '1914–18 conflict').

— Delete duplicate word:

1898 C. RAE *Malaboch* p.xvii (Jeffreys), The Commissioner sent some of the missionary Kafirs up to to the Hoofstadt with a request for Malaboch to come down as he wanted to take the census of the tribe.

As this was verified in the primary source, it was also possible to remove the details of the secondary source (in this case, Jeffreys), and the spelling 'Hoofstadt' corrected, so that the quotation will now look like this:

1898 C. RAE *Malaboch* p.xvii The Commissioner sent some of the missionary Kafirs up to the Hoofstadt with a request for Malaboch to come down as he wanted to take the census of the tribe.

- Correct author, and add nested author for serials. For example: *A Beginner's Guide to Our Birds* — Jo Oliver was printed as J.O. Oliver.

Examples of improvements

- Nested source information (author and title) added.
- Verify against primary source or earlier edition, especially Pettman, Swart and Jeffreys quotations.
- Take quotations from the same, earlier, edition instead of from several editions:

kaffir: 1908 B. BLACKBURN *Leaven* (1991) 75 The nature of the charge against him was known to every kafir in the jail before he had been there an hour, for news travels fast in kafirdom.

hamba: 1908 D. BLACKBURN *Leaven* 299, I should be only a kitchen boy, as I was in Maritzburg, with the police always waiting to catch me for being out after the 'hamba kyah' bell had rung.

Notice also that the author's initial is incorrect in the first quotation.

- Headword does not appear in quotation:

kill-me-quick: 1948 E. HELLMANN *Rooiyard* 48 At the present time *babaton* comes first in popularity, with *shimeya* or *shimeyani* as close second.

However, the addition of the next sentence resolves this error:

kill-me-quick: 1948 E. HELLMANN *Rooiyard* 48 At the present time *babaton* comes first in popularity, with *shimeya* or *shimeyani* as close second. *Isiqataviki* (kill-me-quick) is made to a much lesser degree.

- Consistency of titles. For example:

1955 T.B. DAVIE *Education & Race Rel.* (Hoernlé Mem. Lecture 1955) 10 Since 1933 .. the Afrikaans-medium universities have disaffiliated themselves [from NUSAS] and have joined a new and separate organization, Die Afrikaanse Studentebond (A.S.B.).

1955 T.B. DAVIE *Educ. & Race Rel.* (Hoernlé Mem. Lecture) 10 As far back as 1924 the students .. created .. a National Union of Students (N.U.S.A.S.) which for many years operated successfully in the interests of students generally.

1955 T.B. DAVIE *Hoernlé Mem. Lecture 1955* 30 The early stages of a restored South Africanism having been prepared in the infant and primary schools and carried forward into the high schools, what can be done to further this aim at the university level?

are now all represented thus:

1955 T.B. DAVIE *Educ. & Race Rel.*

As the full information (that it is The Hoernlé Memorial Lecture 1955) will be given in the enlarged bibliography under the author's name, it is not necessary to include this in the quotation.

- Add additional relevant and useful information. For example, at **Spear of the Nation:**

1964 H.H.W. DE VILLIERS *Rivonia 80 Umkhonto we Sizwe* (The Spear of the Nation), was formed under the auspices of the National Liberation Movement.

now reads

1964 H.H.W. DE VILLIERS *Rivonia 80 Umkhonto we Sizwe* (The Spear of the Nation), was formed under the auspices of the National Liberation Movement which included the African National Congress and the South African Communist Party.

(A missing full stop after the author's third initial has also been inserted.)

Unverifiable quotations

It has been impossible to verify all the quotations in DSAEHist. Some of the titles which are not available through any of the lending libraries in South Africa to which the DU has access, or through any of the electronic book repositories, are:

L.P. Greene *The L. Patrick Green Adventure Omnibus* (1925?)

Gwen Westwood *Bright Wilderness* (1970)

Essop et al. *Challenge to 25 Natal Moolvies* (1969)

The Field (13 Oct 1877 and 9 March 1902)

Suzanna Lynne *Glittering Gold* (1972)

Maximilien Kollisch *The Musselman Population at the Cape of Good Hope* (1867)

E.V.C. *The Promised Land, or, Nine Years (Gold Mining, Hunting, and Volunteering) in the Transvaal* (1884)

Mary Ann Parker *Voyage round the World* (1795)

A number of quotations in DSAEHist were taken from OED, and it has been possible to verify some, but not all, of these in the primary sources. For example, at **sand-mole:**

1850 A. WHITE *Pop. Hist. Mammalia* 232 Another member of this family..is also a native of South Africa: this is the Coast Rat or *Sand-Mole (*Bathyergus maritimus*).

Adam White's *A Popular History of Mammalia: Comprising a Familiar Account of Their Classification and Habits* is not available in South Africa, so OED will have to be relied on as secondary source. The title, however, will be changed to conform to the DSAEHist style, which is to include prepositions, so it will become *Pop. Hist. of Mammalia*.

Another quotation taken from OED for which the primary source is not available in South Africa was (at **imfe**):

1893 WATT *Dict. Econ. Prod. Ind.* VI. iii. 277 This .. is said to be extensively grown in Africa and America, the plant of the former country being the Imphee, and of the latter the Sorgho, which is mainly cultivated on account of sugar.

The abbreviated title does not conform to DSAEHist styling, and the abbreviation 'Prod.' does not appear in the list of abbreviations used in DSAEHist. Although 'Ind.' is given in the list of abbreviations, since this is the only time it is used, the entry will be changed to

1893 G. WATT *Dict. of Econ. Products of India* VI. iii. 277 This .. is said to be extensively grown in Africa and America, the plant of the former country being the Imphee, and of the latter the Sorgho, which is mainly cultivated on account of sugar.

which is rather more reader-friendly. (And the initial of the author has been inserted.)

Conclusion

As time went by, it became more and more evident how necessary it was to do this work, because of the number of quotations that had been, for one reason or another, incorrectly printed in DSAEHist. There will inevitably be some quotations that cannot be tracked down, but the DSAE's quotation verification team feels that it is well on the way to making this a much more accurate representation of the story of South African English.

References

- Atkins, B.T. Sue and Michael Rundell. 2008. *The Oxford Guide to Practical Lexicography*. Oxford/New York: Oxford University Press.
- Brewer, Charlotte. 2007. *Treasure-house of the Language: The Living OED*. New Haven Conn./London: Yale University Press.
- Brewer, Charlotte. 2009. *Examining the OED* [online]. <http://oed.hertford.ox.ac.uk/main/> [accessed 27 June 2009].
- Burchfield, Robert, with an Afterword by John Simpson. 2002. *The English Language*. Oxford Language Classics. Oxford/New York: Oxford University Press.
- Coleman, Julie and Anne McDermott (Eds.). 2004. *Historical Dictionaries and Historical Dictionary Research: Papers from the International Conference on Historical Lexicography and Lexicology at the University of Leicester, 2002*. *Lexicographica*. Series Maior 123. Tübingen: Max Niemeyer.
- Fowler, Rowena. 2004. Text and Meaning in Richardson's Dictionary. Coleman, Julie and Anne McDermott (Eds.). 2004: 53-62.
- Gouws, Rufus H. and Danie J. Prinsloo. 2005. *Principles and Practice of South African Lexicography*. Stellenbosch: SUN PRess.

- Hartmann, R.R.K. and Gregory James.** 1998. *Dictionary of Lexicography*. London/New York: Routledge.
- Jeffreys, M.D.W.** 1964. Africanderisms. *Africana Notes and News/Africana Aantekeninge en Nuus* 16(2): 43-95.
- Jeffreys, M.D.W.** 1967. Africanderisms II. *Africana Notes and News/Africana Aantekeninge en Nuus* 17(5): 216-220.
- Jeffreys, M.D.W.** 1970. Africanderisms III. *Africana Notes and News/Africana Aantekeninge en Nuus* 19(1): 29-41.
- Kastovsky, Dieter.** 2000. Words and Word-Formation: Morphology in *OED*. Mugglestone, Lynda (Ed.). 2000. *Lexicography and the OED. Pioneers in the Untrodden Forest*: 110-125. Oxford/New York: Oxford University Press.
- Landau, Sidney I.** 2001. *Dictionaries: The Art and Craft of Lexicography*. Second Edition. Cambridge: Cambridge University Press.
- Mooijaart, Marijke.** 2004. Citations in the *Woordenboek der Nederlandsche Taal*. Coleman, Julie and Anne McDermott (Eds). 2004: 201-211.
- Mugglestone, Lynda.** 2005. *Lost for Words: The Hidden History of the Oxford English Dictionary*. New Haven Conn.: Yale University Press.
- Murray, J.A.H. (Ed.).** 1884. *A New English Dictionary on Historical Principles. Volume 1*. Oxford: Clarendon Press.
- Pettman, Charles.** 1913. *Africanderisms: A Glossary of South African Colloquial Words and Phrases and of Place and Other Names*. London: Longmans, Green.
- Rhodes University Dictionary Unit for South African English (Eds.).** 2002. *South African Concise Oxford Dictionary*. Cape Town: Oxford University Press Southern Africa.
- Silva, Penny and Rhodes University Dictionary Unit for South African English (Eds.).** 1996. *A Dictionary of South African English on Historical Principles*. Oxford: Oxford University Press.
- Simpson, John.** 2003. The Production and Use of Occurrence Examples. Van Sterkenburg, P.G.J. (Ed.). 2003. *A Practical Guide to Lexicography*: 260-272. Terminology and Lexicography Research and Practice 6. Amsterdam/Philadelphia: John Benjamins.
- Simpson, John.** 2009. *Preface to the Third Edition* [online]. <http://dictionary.oed.com/about/oed3-preface/> [accessed 26 May 2009].
- Svensén, Bo.** 1993. *Practical Lexicography. Principles and Methods of Dictionary-Making*. Translated from the Swedish by John Sykes and Kerstin Schofield. Oxford/New York: Oxford University Press.
- Swart, C.P.** 1934. *Africanderisms: A Supplement to the Rev. Charles Pettman's Glossary of South African Colloquial Words and Phrases and Place and Other Names*. M.A. Thesis. Pretoria: UNISA.
- Zgusta, Ladislav.** 2006. *Lexicography Then and Now: Selected Essays*. Edited by Fredric S.F. Dolezal and Thomas B.I. Creamer. Lexicographica. Series Maior 129. Tübingen: Max Niemeyer.