

Adapting a Historical Dictionary for the Modern Online User: The Case of the *Dictionary of South African English on Historical Principles*'s Presentation and Navigation Features*

André du Plessis, *Department of Afrikaans and Dutch,
Stellenbosch University, Stellenbosch South Africa (ahdp@sun.ac.za)*
and

Tim van Niekerk, *Dictionary Unit for South African English,
Rhodes University, Grahamstown, South Africa (T.vanNiekerk@ru.ac.za)*

Abstract: A collaboration between the Dictionary Unit for South African English (DSAE) and Stellenbosch University (SU) was initiated in 2016 to address the structure, layout and functionality of the online *Dictionary of South African English on Historical Principles* (DSAEHist). This article focuses on work undertaken during the initial stages of the DSAE-SU collaboration and presents current work in adapting a print historical dictionary for the digital age. As with other print dictionaries that enter the electronic medium, the pilot version of the online DSAEHist still felt and looked like a print dictionary, albeit with a somewhat more user-friendly access structure. Another difficulty that the DSAEHist faces is the need for a dictionary based on historical principles to showcase large diachronic datasets. This requires that the article and microstructure, as well as the access structure, be adapted. Apart from this, the general layout of the dictionary is also being re-evaluated and redesigned to make it more usable on online platforms. To accommodate new users and to bring the DSAEHist into the digital era (with regard to standard web and mobile practices) stylistic, navigational and functional changes are being made. This article presents these changes to show how a valuable cultural and language resource is being brought into the digital era.

Keywords: SOUTH AFRICAN ENGLISH, HISTORICAL DICTIONARY, ONLINE DICTIONARY, PRINT ADAPTATION, MICROSTRUCTURE, MACROSTRUCTURE, ARTICLE STRUCTURE, USER RESEARCH, DATA PRESENTATION, USABILITY, NAVIGATION

* This article was presented as a paper at the Twenty-first Annual International Conference of the African Association for Lexicography (AFRILEX), which was hosted by the Xitsonga and Sesotho sa Leboa National Lexicography Units, Tzaneen, South Africa, 4–6 July 2016.

Opsomming: Die herstrukturering van 'n historiese woordeboek vir moderne aanlyn gebruikers: Die geval van die *Dictionary of South African English on Historical Principles* se data-aanbieding- en navigasie-eienskappe.

In 2016 is 'n samewerkingsooreenkoms tussen die Dictionary Unit for South African English (DSAE) en die Universiteit Stellenbosch (US) van stapel gestuur om aandag te skenk aan die onvoldoende struktuur, uitleg en funksionaliteit van die aanlyn *Dictionary of South African English on Historical Principles* (*DSAEHist*). Hierdie artikel bied 'n blik op die aanvanklike resultate van hierdie samewerkingsooreenkoms en wys ook hoe hierdie historiese woordeboek aangepas word vir die digitale era. Soos met talle ander gedrukte woordeboeke wat omgeskakel word na elektroniese aanlyn woordeboeke, het die loodswaerking van die aanlyn *DSAEHist* nog soos 'n gedrukte woordeboek gelyk en gefunksioneer, al was die toegangstruktuur ietwat meer gebruikersvriendelik. 'n Verdere probleem is dat 'n historiese woordeboek soos die *DSAEHist* 'n groot aantal diachroniese data moet kan vertoon. Daarom moet die artikel- en mikrostruktuur, asook die toegangstruktuur, herontwerp en herbedink word. Daarbenewens moet die algehele uitleg en aanbieding van data ten opsigte van gebruikersvriendelikheid vir die elektroniese sfeer ook verfyn word. Ten einde die *DSAEHist* in staat te stel om nuwe gebruikers te lok en te behou, asook om ooreen te stem met die standaarde en beginsels van moderne internet- en mobiele praktyke, is stilistiese, navigasie- en funksionele veranderinge aangebring. Hierdie artikel het ten doel om te wys hoe 'n kosbare kulturele en taalbron aangepas word vir die digitale era.

Slutelwoorde: SUID-AFRIKAANSE ENGELS, HISTORIESE WOORDEBOEK, AANLYN WOORDEBOEK, AANPASSING VAN GEDRUKTE WOORDEBOEKE, MIKROSTRUKTUUR, MAKROSTRUKTUUR GEBRUIKERSTUDIES, DATA-AANBIEDING, BRUIKBAARHEID, NAVIGASIE

1. Introduction

At the 20th AFRILEX International Conference in 2015, the pilot online version of *A Dictionary of South African English on Historical Principles* (*DSAEHist*) was presented and critically discussed (cf. Van Niekerk 2015). During this session practical and theoretical obstacles and shortcomings of the online *DSAEHist*, especially with regard to general usability, layout and functionality, were highlighted. Subsequently a collaboration between the Dictionary Unit for South African English (DSAE) and the University of Hildesheim (HU) was formalised to "address the need for a thorough adaptation of the dictionary text to support publication on multiple electronic platforms" (Van Niekerk 2015). In 2016 a further collaboration was established between the DSAE and Stellenbosch University (SU). This article elaborates on the work done on the *DSAEHist* as part of the DSAE-SU collaboration. Since the dictionary adaptation is ongoing and the DSAE-SU collaboration builds on and overlaps with prior DSAE-HU work, it may be helpful firstly to distinguish between areas of adaptation and secondly to give a brief background on the pilot version of the *DSAEHist* and clarify the scope of this article.¹ A final introductory note is that

for the development of user-orientated and functional enhancements as part of the SU collaboration, the ideas set forth by Du Plessis (2015) on electronic and mobile dictionaries were relied upon for theoretical guidance.

2. Scope of projects

The scope of the research and adaptations presented here fit into the wider context of other dictionary adaptation projects that are moving from print to digital, for example the *Woordeboek van die Afrikaanse Taal* (Botha and Fourie-Blair 2016) and Macmillan's dictionaries (see Rundell 2015). The main focus of the DSAE-HU collaboration is restructuring the underlying dictionary dataset to support functionality for enhanced search functionality across the dictionary, to provide more flexible presentation of the article microstructure, and to introduce some new visual components and layout strategies customised for the display of complex historical dictionary data. Some new visual dictionary components, which will be presented in a different article (cf. Van Niekerk et al. 2016a), have also been developed. This restructuring of the dataset has made possible certain layout changes and presentation strategies which were proposed as early prototypes but not yet finally modelled under the DSAE-HU collaboration.

The DSAE-SU collaboration builds on this prior work in that it finalises layout changes which have now become possible, as well as modelling some new presentation features not proposed before. Unlike the DSAE-HU project, which focused on changes to the underlying structure of the dictionary dataset and exploiting the results in article presentation, the DSAE-SU project is specifically focused on the user interface, namely enhancements to overall layout and design at the micro- and macrostructural levels, as well as on substantial improvements to navigation features which had not been addressed before. In short, current DSAE-SU work is concerned primarily (but not exclusively) with questions of design, accessibility and navigation as well as, from the user's point of view, the "look and feel" of the dictionary. It should be noted that the screenshots given in this article are wireframe examples which usually emphasise structural and functional design changes, without final font styling, colour schemes and so forth. Many of these prototypes have not yet been implemented since the project is a non-profit undertaking and operates under resource constraints beyond the control of the project team. The intent of this article is therefore to present current work for three key purposes: (1) to present some examples of the dictionary design possibilities that are becoming available in the still-developing area of electronic lexicography; (2) to show how lexicographical theory can meet practical lexicography in a large-scale dictionary project of complex design; and especially (3) to generate discussion and invite feedback from the lexicographical community.

3. The pilot online edition

The *DSAEHist* is a diachronic comprehensive variety dictionary documenting the vocabulary of South African English from its origins in the 17th century to the date of publication. The aim of the dictionary is "to map and illustrate the complex landscape of [this particular] variety of English" (Silva et al. 1996: vii). To achieve this goal, historical data are always incorporated in the form of "bibliographically documented and dated quotations" (Van Niekerk et al. 2016a). The print version of this dictionary has been out of print since 2004, but a pilot online version was launched in 2014 (<http://dsae.co.za>).

As with many other print dictionaries that enter the electronic medium, the pilot version of the online *DSAEHist* still felt and looked like a print dictionary, albeit with a somewhat more user-friendly access structure. Due to the *DSAEHist*'s need to showcase large diachronic datasets, the article and micro-structure, as well as the access structure, are currently not suited to achieve this goal on an electronic platform. Apart from this, the general layout of the dictionary also needed to be re-evaluated and redesigned to make it more usable on online platforms. Many of the aforementioned issues stem from the "internal representation", which was originally designed purely for print typesetting purposes, being only "partly structured with lexicographic categories" (Van Niekerk et al. 2015). Another area of concern was the problematic and somewhat outdated design of the user interface and limited multi-functionality and interactivity. The XML dataset used for the pilot online version lacked the flexibility required to support layout changes, and the general layout is akin to the structure of the printed version which causes navigational, structural and usability obstacles.

In order to move from print to the electronic sphere, different adaptations need to be made. Although many of these adaptations relate to structural and technological aspects, no change can be made without identifying and catering for the dictionary user. Bergenholtz and Tarp (2003: 172) show that dictionaries should satisfy a range of user needs. Designing or adapting a dictionary therefore relies on the balance between dictionary structures and user needs (Gouws 2005: 56, 60). Consequently a review of the *DSAEHist*'s users is crucial for the adaptation process.

4. Dictionary users

The online *DSAEHist* occupies a rather unusual position among electronic dictionaries. It republishes a scholarly dictionary originally compiled over 25 years of comprehensive historical research with a relatively narrow audience of language specialists, historians, writers and interested laypersons. It was not intended as a popular dictionary — this role is fulfilled by the much smaller print dictionary, *A Dictionary of South African English* by Jean Branford (published in four editions between 1978 and 1991). Yet the online *DSAEHist*, unlike

the *OED* on which it was originally modelled, is a free dictionary which is now consulted by a large audience of lay users who may never have seen the historical dictionary model before, let alone a variety dictionary. The result is a forked user base for the electronic edition, with different sets of expectations from two groups, which can be roughly divided as follows:

- Casual Googlers: laypersons who arrive at a *DSAEHist* entry after consulting a search engine about a South African term; and
- Target-users: the original intended audience of writers, editors, translators, linguists, historians and other more sophisticated dictionary users.

This division, coupled with a significant departure from print design and layout conventions in electronic dictionaries, has long-term implications for the adaptation of the original dictionary, prompting many new requirements which can only be addressed in stages. Du Plessis (2015: 27) posits that it is imperative to take note of the range of user requirements as the user always plays a key role in determining the functionality and usability of a dictionary. Therefore, to help identify and prioritise the aforementioned requirements, the DSAE compiled an online user survey which ran from 17 March to 16 May 2016 (Van Niekerk et al. 2016). Another source of information about users is Google Analytics reports, which give general statistical information about users such as their devices, geographical locations and search engine keywords sorted by date or frequency.

The user survey delivered some interesting results but does not necessarily reflect the true scope of the general *DSAEHist* user group since 49 out of the 75 respondents (65.33%) had not used the *DSAEHist* before (Le Du 2016). Nevertheless, this survey sheds some light on possible user groups, needs and expectations. Many of the changes made are also supported by the data from the survey. A few key points from the survey are summarised below.

- The majority of respondents fall into the 60+ age group (38.67%) followed by the 31–40 age group (22.67%), while the 41–59 age group contained 28% of the respondents. That means only around 10% of the respondents are 30 and younger. The skewed data here is mostly due to the distribution of the survey as shown below.
- The majority of respondents are language professionals (58.67%), with a further 9.33% identifying as working specifically in the field of dictionary compilation, publishing, etc. This result however would be influenced by the distribution list of the survey that included many professional and academic institutions or associations. These figures are still consistent with the previously stated target-user group of the *DSAEHist*.
- The majority of respondents identify as speakers of South African English (81.33%) although most still believe that the concept of this variety should be explained to users of the dictionary (86.67%).

- 84% of respondents indicated that they would not prefer a shorter entry with less information when using a computer, although 58.67% indicated that they would like to control the amount of information displayed as clickable options. This relates to adapting the layout in such a way as not to overload the user with data and to make the dictionary environment more functional.

The user survey gives a solid representation of the target-users and reflects on some of their needs or expectations, but it does not necessarily provide enough detail on possible future users (see Section 7). Gouws (2016) mentions that lexicographers need to be aware of "emerging generations" of dictionary users. Nied Curcio (2015) states that the new generation of modern users, the so called "digital natives", are becoming the primary dictionary users. It is likely that a large proportion of the casual Googler user group falls into this category. Therefore, to accommodate both the target-users and emerging dictionary users, the current and any future enhancements to the *DSAEHist* are additionally focused on the evolving needs, expectations and skills of a possible future audience.

From the Google statistics it became clear that the majority of users access the dictionary from a personal computer or laptop. Due to the increase in mobile resources and society's dependence on mobile technology, an increase in mobile users in the near future is however to be expected (Du Plessis 2015: 24-25). This is also conveyed in the user survey where 45% of respondents indicated that they would use a *DSAEHist* app if available. The ideas and enhancements incorporated during the DSAE-SU collaboration are thus implemented in such a way as to make a long-term shift towards improved mobile support easier. Although there is currently a mobile version available, this version will not be discussed here as it needs first to be critically evaluated during a future stage of the adaptation project. The proposed restructuring and simplification of the desktop design will also be reflected in the mobile version and offer significant improvements.

5. Adaptations to the pilot edition

The following two sections illustrate the different adaptations to the pilot edition of the *DSAEHist*. In some cases prior preparatory work is mentioned as introductory material, but it will be clear from the specified dates or other contextual information how the current collaboration refines and extends existing enhancements while also adapting and improving new areas of the online *DSAEHist*. Many of the initial changes and future changes are supported by the user survey.

5.1 General layout

The first step in modernising the *DSAEHist* was to rethink the general page

layout of the pilot edition including the landing page. As Figure 1 shows, the pilot's general dictionary environment is basic and uninviting and does not conform to modern online user expectations in terms of functionality and interactivity as both are limited and not multidimensional (Rundell 2015: 310). There are some useful features like search autocompletion, the alphabet bar and user-tabs, but they need to be refined and integrated in innovative ways to make them more functional. The pilot edition's stylistic choices, such as the colour scheme, font choices, etc., are outdated and drab. Du Plessis (2015: 115) mentions that although it is imperative not to alienate dictionary users by creating an unfamiliar dictionary environment, online dictionaries need to have interesting and modern style choices in order to attract users and to conform to certain current standard web practices and norms. The design and screenshot structure has therefore been adapted in the new working model so that the search routes, user guidance, access structure and user interactivity become visible and prominent features of the dictionary. Design changes will include new fonts and a less obtrusive colour scheme.

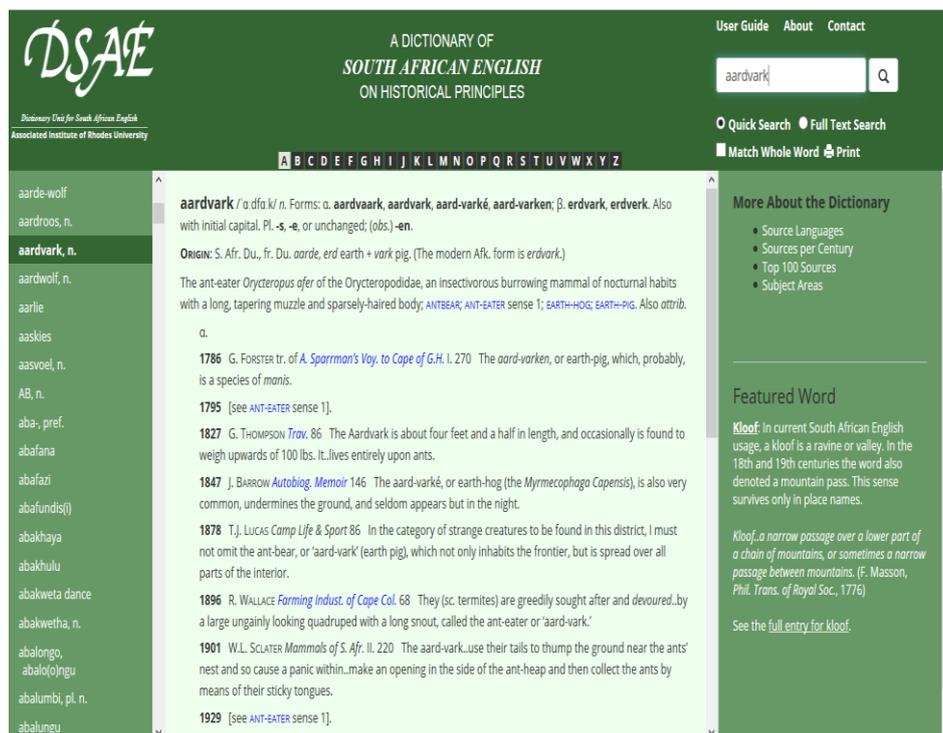


Figure 1: Pilot edition screenshot structure showing entry **aardvark**

The new working model of the landing page presents the user with a simplified logo and banner and possibly shortened title (pending copyright clarification) on the left of the page (see Figure 2). The newly-modelled logo area removes the abbreviation expansions and the mention of the DSAE's status as an affiliated institute of its host university (this was included originally to acknowledge Rhodes University as the source of initial funding for the pilot edition). Institutional information will henceforth be displayed on the About page, reserving the logo and banner for dictionary branding alone.

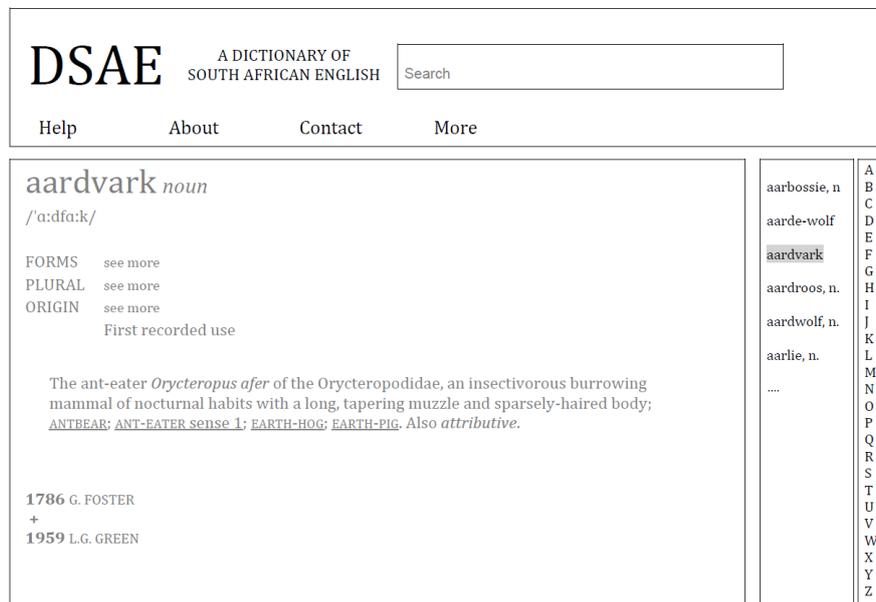


Figure 2: Wireframe prototype of general layout and displayed article

A new lighter colour scheme with readable modern online fonts will be sampled and used. The Print button shown below the search options in Figure 1 will be integrated into the more functional extended microstructure of each displayed dictionary article. Another important change that will improve the functionality and navigability of the landing page, is moving (and enlarging) the search bar to the top-centre of the page. The search bar has become synonymous with web browsing and is the component that online users most frequently access (Terry 2016; Du Plessis 2015: 54). Placing it in a prominent position invites users to start using the dictionary immediately. This change matches current trends in electronic dictionary design and are reflected in the user survey as 50–75% of respondents indicated that clarity, adaptability,

searchability and accessibility are of very high importance, thereby meeting developing user expectations.ⁱⁱ

The new layout also moves menu navigation components (User Guide, About, Contact, and More About the Dictionary content, the latter currently displayed in a side panel) to a set of prominent user-friendly tabs underneath the new search bar and expands on their content (Figure 2). This involves creating easily accessible outer textual tabs to, for example, user guidance, more information about the dictionary and DSAE, and contact information. A label-based Home button will also be introduced. This feature was cut from the pre-publication pilot edition on the basis that the logo functions as a home button, a common convention on the Internet, and that the institutional information under the logo already cluttered the area with text. However, this functionality was insufficient and a Home button is required as an explicit navigation feature to improve usability and navigability, especially for inexperienced online users (Du Plessis 2015: 52, 79).

Online dictionaries should offer multiple search routes to users, especially to target users who are interested in the range of lemmas in the dictionary (Du Plessis 2015: 54-55). With this in mind the alphabet bar and lemma list sidebar were moved to the far-right of the page. The user survey also showed that these components were mostly used by respondents, therefore they cannot be removed. The changes were implemented in order to display the content of greatest interest to the user in the left-centre of the page, in accordance with the F-Layout reading flow pattern followed by most Internet users (see Ambrose and Harris 2010), to provide extra non-intrusive search routes that don't overload the page with data, and to conform to the standard web practice of showing less prominent information, links and options to the right of a webpage (Ambrose and Harris 2010: 168) (see Figure 2).

5.2 User engagement, guidance and feedback features

Given that a historical variety dictionary is unfamiliar to many dictionary users who are better acquainted with general monolingual lookup tools — most South Africans are generally unaware of the extent of South African English — it is also important to offer an engaging initial access point allowing optional exploration of contents which are otherwise opaque on first viewing the dictionary. To facilitate user engagement and improve the general look of the dictionary, low-maintenance but dynamic content (e.g. featured words or interesting facts about South African English) on the landing page is offered as well as under some of the user tabs, with additional dictionary content nested under the More tab. We also propose showing popular searches and user statistics derived from Google Analytics for users to view.

Rather than overloading the user with vast amounts of text, and for the sake of user-friendliness (Tarp 2009: 27), the user guidance will be restructured in the form of frequently asked questions (FAQs), such as "How does the search option work?" or "What do the labels next to the definitions mean?",

with short easy answers that could include step-by-step guidance or example entries. Key FAQs will be incorporated onto the landing page as another means to improve user-friendliness and help new users become accustomed to the page. User interaction can also be gained by including a user feedback form either on the landing page or under the Contact tab since e-mail links are not always effective. Engaging with the user is important to the lexicographic process and adds a mediation function (cf. Fuertes-Olivera 2014) to the dictionary. The possibility of adding images in the form of an image slider were also considered but due to copyright restrictions and added load times to the webpage, as well as the fact that images appear to go against current online dictionary practices, especially for a historical dictionary (Le Du 2016a), this option is under review.

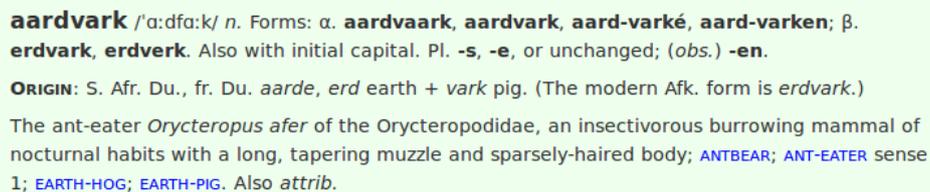
Lastly, as mentioned, dictionary content will be nested under the More tab. Here users will find new dynamic infographics that visually encapsulate and present the user with information on the source languages, sources per century and subject areas, shown as text-heavy tables in the pilot edition. These infographics show interesting extra content that is linked with the dictionary entries. This means that users can find their way to related entries or access the dictionary content in a unique way (see Section 6.2.7 for more information). A multi-layered presentation of data adds to the overall interactivity of the dictionary and is also a different way to navigate through and promote the historical data of the dictionary. These infographics are, however, still a work in progress.

The changes made to the general layout, access structure and screenshot structure inevitably lead to changes in the article and microstructure.

6. Changes to article structure and microstructure

6.1 Initial enhancements

The initial work of restructuring the XML dataset which started during the DSAE-HU collaboration has led to some initial enhancements in the presentation of articles. The following example from *aardvark* in Figures 3 and 4 shows some of these changes (Van Niekerk et al. 2016a):



aardvark /'ɑ:dfɑ:k/ *n.* Forms: α. **aardvaark**, **aardvark**, **aard-varké**, **aard-varken**; β. **erdvark**, **erdverk**. Also with initial capital. Pl. **-s**, **-e**, or unchanged; (*obs.*) **-en**.
ORIGIN: S. Afr. Du., fr. Du. *aarde*, *erd* earth + *vark* pig. (The modern Afk. form is *erdvark*.)
The ant-eater *Orycteropus afer* of the Orycteropodidae, an insectivorous burrowing mammal of nocturnal habits with a long, tapering muzzle and sparsely-haired body; **ANTBEAR**; **ANT-EATER** sense 1; **EARTH-HOG**; **EARTH-PIG**. Also *attrib.*

Figure 3: Pilot edition before enhancements

aardvark /ɑ:dfɑ:k/ *noun*.

- **FORMS:** α. **aardvaark**, **aardvark**, **aard-varké**, **aard-varken**; β. **erdvark**, **erdverk**. Also with initial capital.

- **INFLECTIONS:** Plural **aardvarke**, **aardvarks**, or unchanged; (*obsolete*) **aardvarken**.

- **ORIGIN:** South African Dutch, from Dutch *aarde*, *erd* earth + *vark* pig. (The modern Afrikaans form is *erdvark*.)

The ant-eater *Orycteropus afer* of the Orycteropodidae, an insectivorous burrowing mammal of nocturnal habits with a long, tapering muzzle and sparsely-haired body; [ANTBEAR](#); [ANT-EATER](#) sense 1; [EARTH-HOG](#); [EARTH-PIG](#). Also *attributive*.

Figure 4: Initial improvements to pilot edition

The 2015 working model shown in Figure 4 reflects changes to content presentation, structure and layout which improve the general usability and readability of data in the *DSAEHist*. These changes include:

- the separation and "labelling of main entry components (forms, plurals and origin) into block-level elements" (Van Niekerk et al. 2016a);
- in the *inflections block*: the manual expansion of abbreviations telescoping such that, for example, "Pl. -s, -e, or unchanged; (*obs.*) -en" is rendered fully as "Plural **aardvarks**, **aardvarke**, or unchanged; (*obsolete*) **aardvarken**";
- in the *etymology block*: the expansion of abbreviations here, where they were numerous, as well as in other areas of the dictionary.

The removal of abbreviations is a substantial undertaking requiring largely manual editorial intervention and proofreading, but this has been newly adopted as a general styling policy for the electronic edition (except in quotations which have their own internal editorial consistency) for two reasons: Firstly, a large set of abbreviations were in use, sometimes for language names or grammatical terms which would be unfamiliar to non-specialists, requiring the user to check against an extensive abbreviation glossary. To remove this burden from the user, and for general readability, abbreviations have generally been removed or are deprecated in styling policy. A second, important reason for removing abbreviations is that they obstruct machine-readability for Internet search engines and the online dictionary's internal search functionality. For example, the search engine keywords "plural aardvark" would be less likely to be matched to the relevant page on the dictionary if it contained only the abbreviated form "Pl."

6.2 Further adaptations: current work

The pilot edition presentation, reflecting the inflexibility of the pilot edition

dataset, displayed the initial entry components as a run-on line as per the print edition:

aardvark /'ɑ:dfɑ:k/ *n.* Forms: α. **aardvaark**, **aardvark**, **aard-varké**, **aard-varken**; β. **erdvark**, **erdverk**. Also with initial capital. Pl. **-s**, **-e**, or unchanged; (*obs.*) **-en**.
ORIGIN: S. Afr. Du., fr. Du. *aarde*, *erd* earth + *vark* pig. (The modern Afk. form is *erdvark*.)

Figure 5: Unadapted pre-definition microstructural components from pilot edition

As shown, some preliminary enhancements to the microstructure and article structure have been investigated and tested. The current article structure is a legacy of the print edition and its ubiquitous space-saving editorial policies, augmented with basic hyperlinks for cross-references between articles and to the bibliography of quotation sources. In order to make the dictionary articles more user-friendly and to showcase the historical data, a repositioning and labelling of data categories (as part of the work done with HU) was undertaken as well as including a more multi-layered or dynamic approach to the presentation of content. Again styling changes have also been employed in the actual dictionary articles:

aardvark *noun*

[/ɑ:dfɑ:k/](#)

FORMS: [aardvaark](#) » [more...](#)

PLURALS: [aardvarke](#) » [more...](#)

ORIGIN: South African Dutch » [more...](#)

[First recorded use 1786](#)

The ant-eater *Orycteropus afer* of the Orycteropodidae, an insectivorous burrowing mammal of nocturnal habits with a long, tapering muzzle and sparsely-haired body; [ANTBEAR](#); [ANT-EATER SENSE 1](#); [EARTH-HOG](#); [EARTH-PIG](#). Also *attributive*.

1786 G. FORSTER tr. of *A. Sparrman's Voy. to Cape of G.H.* I. 270 The *aard-varken*, or earth-pig, which, probably, is a species of *manis*.

[+]

1959 L.G. GREEN tr. of *These Wonders* I. 270 That creature of obscure origin, that champion tunneller of the veld, the erdvark or ant-eater. This pig-shaped freak is not rare, but is seldom captured.

Figure 6: Newly-modelled article layout

The redesigned article layout shown in Figure 6 is made possible only through structural changes to the underlying XML dataset which itself represents a comprehensive map of the dictionary design. Some of this restructuring is complete, and some of it represents current work. In all cases adaptations to the dataset require that structural exceptions be handled, prompting an iterative redesign process. The result of these adaptations is ever-increasing flexibility of

visual and stylistic presentation. The results are as follows, taken in order of the article components.

6.2.1 Headword and part of speech

The headword and part of speech have been restyled so that the headword now occupies a much more prominent position on the page, also functioning as a heading (see Figure 7). This adds a contextual signpost for the user while also playing a role in navigation since the headword and part of speech will, according to the new model, always remain in view when scrolling through the rest of the article contents. An anchored heading-style headword is required here since articles may often be hundreds of words longⁱⁱⁱ and in the pilot edition the headword often disappears from view when scrolling. Research has also prompted the adoption of a mobile-friendly sans serif font (for the headword and part of speech) as it reads more easily and looks similar to fonts used across different popular websites (Beaird 2007: 113). Font choice goes beyond purely visual considerations: the working model illustrated in Figure 7 uses the Google Roboto font for the secondary reason that it can be cached locally on the user's device, thus avoiding a negative performance impact caused by font-loading — one of the ways current design work anticipates the needs of mobile device interfaces.

aardvark *noun*

Figure 7: New lemma and part of speech presentation

The part of speech has been italicised and rendered in a smaller point size than the headword (as shown in Figure 7). This is done as apart from styling, it's also a subtle navigation aid, i.e. it allows the user (once he/she has seen a few entries) to distinguish these components intuitively before even reading the words. In the print edition all parts of speech were abbreviated, except in the case of longer parts of speech. These will now be shown expanded. The heading-style foregrounding of the headword and unabbreviated parts of speech also serve the additional purpose of improving search engine optimisation.

Furthermore, lexicographic and contextual labels will be moved to the start of each sense. A final enhancement in this section relates to unassimilated lexical items, namely borrowings which were not regarded as fully integrated into South African English at the time of publication but were included in the dictionary "in order to document the *S. Afr. Eng.* vocabulary as comprehensively as possible" (Silva et al. 1996: viii). In the print and pilot editions these lemmas are preceded by tramlines (||) as a type of label. This is sometimes confusing or obtrusive, and a colour change or less confusing symbol glossed with

a mouse-over popup will be substituted to improve readability and interpretability.

6.2.2 Pronunciation

The user survey indicated that the pronunciation is a vital component of the dictionary entry and that the IPA symbols are usable to users. The pronunciation has however now been moved to below the headword. This is done to separate components, make the article structure more scannable, and create a more sleek and modern look. The pilot edition's pronunciation indicator is inadequate in that the user must refer to a long table of IPA symbols on another web page in order to make sense of the phonetic notation. In the new model the pronunciation is hyperlinked, so that when a user clicks on it a popup with translations of only the relevant IPA symbols will appear to offer context-specific help (see Figure 8).



Figure 8: Model of new pronunciation popup (prototype)

6.2.3 Variant spellings, inflections and etymology

The data categories variant spellings, inflections and etymology have, after restructuring of the XML markup, been changed from inline to block-level elements. They have also been assigned more easily-understandable labels, namely:

- variant spellings are labelled "Forms"
- inflections are labelled "Plurals"
- etymology is labelled "Origin"

These elements have also been refined so that they don't overwhelm the user with data (even lists of variant spellings are sometimes very long) but are optionally expandable (see Figure 9). 58% of respondents in the user survey indicated that expandable content would be welcomed. Under the "Forms" category, different spelling variation group sections are sometimes indicated using Greek symbols, referring to different quotation paragraphs when these are separated according to spelling patterns (e.g. *aardvaark* has spellings starting with either *a* or *e*, and these are distinguished by having separate spelling variant subsections and quotation paragraphs). These variant subsections will also be restructured into a block presentation with each section being hyperlinked to the relevant quotations. The first recorded use of the headword has also been included and hyperlinked to the first documented quotation under the "Origin" category. These changes have been proposed as a subtle way to emphasise the historical data and cognitive function of the dictionary and to aid interested users in exploring the deeper content of the article further.

FORMS: [aardvaark » more...](#)

PLURALS: [aardvarke » more...](#)

ORIGIN: [South African Dutch » more...](#)

[First recorded use 1786](#)

Figure 9: Expandable newly labelled data categories

6.2.4 Definitions

The paraphrases of meaning have been foregrounded by placing them in a text box or panel. Google Search Analytics statistics indicate that the meaning statement is of key user interest (Van Niekerk et al. 2015). Foregrounding the definitions allows those users not interested in the other historical components of the entry to access the data they need more quickly, rather than being discouraged by surrounding text which is beyond their interests (see Figure 10). A long-term goal here is also to separate the definition data-types, for example into core definitions and cross-references, as another way to facilitate improved presentation of the data.

The ant-eater *Orycteropus afer* of the Orycteropodidae, an insectivorous burrowing mammal of nocturnal habits with a long, tapering muzzle and sparsely-haired body; [ANTBEAR](#); [ANT-EATER sense 1](#); [EARTH-HOG](#); [EARTH-PIG](#). Also *attributive*.

Figure 10: Foregrounded paraphrase of meaning for aardvark

6.2.5 Quotations

Directly below and visually juxtaposed with the paraphrases of meaning are the quotations. As mentioned, the historical data are an integral part of the *DSAEHist* and the goal here is for users to access this section regularly, since they contain additional information which cannot be synthesised in definitions. However, some lemmas have an extensive list of quotations.^{iv} This poses somewhat of a quandary as the historical data needs to be highlighted, but overloading the user with data causes users to look elsewhere for the required information. Therefore to satisfy most users' needs (again linking to the respondents in the user survey's need for expandable content) a content filtering strategy which displays, by default, the first and last quotation per sense, with the option to click a button to expand to the full list, has been developed (also see Van Niekerk et al. 2016a), and the final implementation of this filtering will be refined in the DSAE-SU collaboration. Beyond this, we also propose font and background changes here as a means to i) direct the users' attention to this section, and ii) to make the category of data clear and readable (see Figure 11).

1786 G. FORSTER tr. of [A. Sparrman's Voy. to Cape of G.H.](#) l. 270 The *aard-varken*, or earth-pig, which, probably, is a species of *manis*.



1959 L.G. GREEN tr. of [These Wonders](#) l. 270 That creature of obscure origin, that champion tunneller of the veld, the erdvark or ant-eater. This pig-shaped freak is not rare, but is seldom captured.

Figure 11: Expandable quotations with a dynamic structural indicator

6.2.6 Cross-references and combinations

Related current work includes expanding cross-references between quotations (another print-edition space-saving device which is no longer necessary). For example, under the quotations for **berghaan** in the pilot edition (see Figure 12) the user is sent between these entries for different quotations (1889 and 1867). To give the user immediate access to the data the target quotation will be shown instead (see Figure 13). This change will apply to approximately 2,800

quotations and has involved a combination of semi-automatic updates and case-by-case proofreading, incidentally resulting in improved content where ambiguous or broken cross-references were identified.

berghaan /'berx, ha:n, 'be:g-/ *n.* *obsolescent*.
ORIGIN: S. Afr. Du., *berg* mountain + *haan* cock.

1. The **BLACK EAGLE**, *Aquila verreauxii*.

1844 J. BACKHOUSE *Narr. of Visit* 488 Three species of Eagle occur in South Africa..[including] *Aquila vulturina*, the Berghaan.

1867 [see **DASSIEVANGER**]

Figure 12: Cross-references between quotations in the pilot version

berghaan *noun*
– ORIGIN: South African Dutch, *berg* mountain + *haan* cock.
obsolescent

1. The **BLACK EAGLE**, *Aquila verreauxii*.

1844 J. BACKHOUSE *Narr. of Visit* 488 Three species of Eagle occur in South Africa..[including] *Aquila vulturina*, the Berghaan.

1867 E.L. LAYARD *Birds of S. Afr.* 11 *Aquila Verreauxii*...*Dassie Vanger* and *Berghaan* of Colonists...It is called '*Dassie Vanger*' (coney-eater) and '*Berghaan*' (mountain cock) by the Colonists..from feeding principally on the coney, or rock-rabbit.

Figure 13: Quotation cross-references expanded

Work on presenting combined forms and their respective quotations as separate block-level elements (rather than unbroken paragraphs as in the print edition) is also under way after having been initiated during the DSAE-HU collaboration. Again a semi-automated dataset adaptation is required. The same strategy has been applied to compounds and derivatives, all of which were formerly nested in paragraphs which did not separate these elements — again, to save space in print. For example, the combinations for **biltong** – **biltong-curtain** and **biltong farmer** (Figure 14) will be displayed in a block at the foot of the entry similar to the block for forms, plurals and origin to enhance readability (see Figure 15).

2. *comb.* **biltong curtain** [by analogy with Eng. *Iron Curtain*], a jocular name for the borders of South Africa; **biltong farmer**, **-hunter**, **-jackal**, one who hunts game in order to obtain meat for the making of biltong.

Figure 14: Combinations in pilot edition

2. combinations

- biltong curtain** [by analogy with English *Iron Curtain*], a jocular name for the borders of South Africa;
- biltong farmer, biltong hunter, biltong jackal**, one who hunts game in order to obtain meat for the making of biltong.

Figure 15: Combinations in new layout

6.2.7 Extended microstructure and access structure

As stated in Section 5.2, guiding the user to the historical data contained in the dictionary is an important part of the new working model. An interesting and innovative way to present and link dictionary data is providing users with relevant contextual or subject specific data as part of the extended microstructure. The extended microstructure also functions as an expanded access structure as users will have the option of accessing different entries or parts of the dictionary from an extra search route. Interestingly, 44% of respondents in the user survey indicated that access to related content would be important.

aardvark *noun*

[/ˈɑːdʌrk/](#)

FORMS: **aardvaark** [» more...](#)

PLURALS: **aardvarke** [» more...](#)

ORIGIN: South African Dutch [» more...](#)
[First recorded use 1786](#)

The ant-eater *Orycteropus afer* of the Orycteropodidae, an insectivorous burrowing mammal of nocturnal habits with a long, tapering muzzle and sparsely-haired body; [ANTBEAR](#); [ANT-EATER SENSE 1](#); [EARTH-HOG](#); [EARTH-PIG](#). Also *attributive*.

1786 G. FORSTER tr. of [A. Sparrman's Voy. to Cape of G.H.](#) l. 270 The *aard-varken*, or earth-pig, which, probably, is a species of *manis*.

[+]

1959 L.G. GREEN tr. of [These Wonders](#) l. 270 That creature of obscure origin, that champion tunneller of the veld, the erdvark or ant-eater. This pig-shaped freak is not rare, but is seldom captured.

Other animals:

aardwolf
aasvoel
alikeukel
[» more...](#)

Also first recorded in the 1780s:

eland
Jan Compagnie
lock-shoe
[» more...](#)

Find other entries quoting:

G. Forster
L.G. Green
[» more...](#)

aarde-wolf
aardroos, n.
aardvark, n.
aardwolf, n.
aarlie, n.
aaskies
aasvoel, n.
AB, n.
aba-, pref.
abafana
abafazi
abafundisi
abakhaya
abakhulu
abakweta dance
abakwetha, n.
abalongo, n.
abaloongu
abalumbi, pl. n.
abalungu
abaMbo
abangoma
Abantu,
abaNtu
abanumzane

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Figure 16: Related content shown on the right of the article

The related content found in the extended microstructure will be nested under labelled display blocks in a new sidebar adjacent to the article. These blocks will be expandable and could also be linked to the infographics previously discussed. The labels describing each related content block and the article hyperlinks listed under them (see Figure 16 for an example) can be derived

from categories such as shared origin, shared register or subject specific lexicographic labels as well as other historical and bibliographical data encoded in the dictionary database. Except for comprehensive subject-categories — which are currently being added to supplement existing subject-labels as a prerequisite for advanced search functionality — these semantic categories are already encoded in the dictionary's richly-annotated dataset, and their XML structure makes their values queryable with ease and flexibility. However the implementation of appropriate dynamically generated related content requires careful modelling and testing, and will only be undertaken on completion of current project priorities.

7. Future work

Apart from the ongoing work already mentioned in adapting the XML dataset, general layout and article structure, there are other possible long-term goals that should be pursued.

One area that requires more research is user engagement and involvement. Raising awareness about the *DSAEHist* should become a priority from a marketing perspective. The various ways of promoting the dictionary should also be investigated. One of the objectives of the current adaptation process is specifically to make the dictionary design and user interface appealing to an emerging group of 'digital natives'.

A possibility to increase user involvement in the future is to get users to actively partake in the lexicographic process. This can be done by for example letting users collect data for the dictionary, especially in terms of quotations.

Apart from this possibility, other user-friendly enhancements such as audio options alongside the newly modelled pronunciation indicator should be pursued, as well as the proposal (mentioned in Section 6.2.7) to provide access routes to related content or extra data. In due course, possibilities for offering different search routes, functional and interactive elements and user involvement via an extended microstructure or access structure should also be explored and developed.

The content and functions of outer texts or extra matter, currently occupied by content from the preface of the print version, also need further discussion. This is an area where, for example, external links and important concepts such as the history of South African English can be contained.

Other navigation features such as scroll bars and scrolling methods also need to be tested, as does the integration of other visual components and pop-ups.

Finally, as many of the current adaptations, as well as many of the future possibilities, are in accordance with many of the results of the user survey, especially in terms of adaptability, clarity and access to data, possible follow-up surveys should also be considered for the future.

The work presented here will take time to complete and publish online due to resource constraints and the scale of the undertaking of adapting a his-

torical dictionary originally designed for print to a functional digital online format. Nevertheless this paper presents current work towards bringing a valuable cultural and linguistic resource into the digital era, a process in which the DSAE is being greatly aided by the present collaboration with SU. Such collaborations between research institutions offer the opportunity of showing how guidelines from theoretical lexicography can be applied in practice, and the non-profit model of the project, which allows experimentation and innovation without risk to publishers' commercial imperatives, makes it well-suited for this form of exchange. As such, the intent of this paper is to show the evolution of the dictionary from a print to electronic product to meet evolving user needs, and we renew our invitation to the lexicographical community to provide feedback on the project's current and future development.

8. Endnotes

- i. Please note that this article and any future publications stemming from work done within the DSAE-SU collaboration may be read in conjunction with publications from the DSAE-HU collaboration. See for example Van Niekerk 2016a.
- ii. The changes to search functions are not discussed here as this forms part of the DSAE-HU collaboration (cf. Van Niekerk et al. 2016a).
- iii. Some entries are much longer, e.g. *African* (approximately 2,300 words) and *Hottentot* (approximately 6,000 words).
- iv. On average there are 10 quotations per entry but the count is much higher for many headwords, e.g. *Cape* has 230.
- v. Gary Terry is an IT and web design specialist at Rhodes University.

9. References

- A Dictionary of South African English on Historical Principles*. [Online.] Available: <http://dsae.co.za>.
- Ambrose, G. and P.E. Harris.** 2010. *The Visual Dictionary of Pre-press and Production*. Lausanne: AVA Publishing.
- Beaird, J.** 2007. *The Principles of Beautiful Web Design*. Melbourne: SitePoint.
- Bergenholtz, H. and S. Tarp.** 2003. Two Opposing Theories: On H.E. Wiegand's Recent Discovery of Lexicographic Functions. *Hermes, Journal of Linguistics* 31: 17-196.
- Botha, W.F. and H. Fourie-Blair.** 2016. *Demonstration of the Pilot Project for a New Online Woordeboek van die Afrikaanse Taal*. Unpublished paper delivered at the 21st International Conference of the African Association for Lexicography (AFRILEX), 5 July 2016, Tzaneen, South-Africa.
- Nied Curcio, M.** 2015. *The Use of Smartphones in Second-language Teaching*. Colloquium, 3 September 2015, Stellenbosch University, Stellenbosch, South Africa.
- Du Plessis, A.H.** 2015. *'n Analise van die selfoon-WAT: 'n Grondslag vir die verbetering van selfoonwoordeboeke*. Unpublished M.A. thesis. Stellenbosch: Stellenbosch University.

- Fuertes-Olivera, P.A.** 2014. Workshop on Specialised Lexicography. Colloquium, 4 June 2014, Stellenbosch University, Stellenbosch, South Africa.
- Gouws, R.H.** 2005. Oor die verhouding tussen woordeboekstrukture, woordeboekinhoud en leksikografiese funksies. *Lexikos* 15: 52-69.
- Gouws, R.H.** 2016. *Integrating Online Reference Sources into a Comprehensive Dictionary Culture*. Unpublished paper delivered at the 21st International Conference of the African Association for Lexicography (AFRILEX), 5 July 2016, Tzaneen, South Africa.
- Le Du, B.** 2016. Report on User Survey: *Dictionary of South African English on Historical Principles Online*. Dictionary Unit for South African English.
- Le Du, B.** 2016a. Report on Use of Images: *Dictionary of South African English on Historical Principles Online*. Dictionary Unit for South African English.
- Rundell, M.** 2015. From Print to Digital: Implications for Dictionary Policy and Lexicographic Conventions. *Lexikos* 25: 301-322.
- Silva, P., W. Dore, D. Mantzel, C. Muller and M. Wright (Eds.)**. 1996. *A Dictionary of South African English on Historical Principles*. Cape Town/Oxford/New York: Oxford University Press.
- Tarp, S.** 2009. Beyond Lexicography: New visions and Challenges in the Information Age. Bergenholtz, H., S. Nielsen and S. Tarp (Eds.). 2009. *Lexicography at a Crossroads* 90: 17-31. Bern: Peter Lang AG.
- Terry, G.** 2016. Personal interview. 2 March. Grahamstown.†
- Van Niekerk, T.** 2015. *Adaptation of A Dictionary of South African English on Historical Principles for Electronic Platforms*. Unpublished paper delivered at the 20th International Conference of the African Association for Lexicography (AFRILEX), 7 July 2015, University of KwaZulu-Natal, Durban, South Africa.
- Van Niekerk, T., H. Stadler and U. Heid.** 2015. *Data Representation and Data Presentation for Selective Queries to the Dictionary of South African English Online*. Unpublished paper delivered at the Hildesheim Conference on Electronic Dictionaries as Information Tools, 8 October 2015, University of Hildesheim, Hildesheim, Germany.
- Van Niekerk, T., L. Morris, J. Wolvaardt and B. le Du.** 2016. User Survey: *Dictionary of South African English on Historical Principles Online*. Dictionary Unit for South African English.
- Van Niekerk, T., H. Stadler and U. Heid.** 2016a. *Enabling Selective Queries and Adapting Data Display in the Electronic Version of a Historical Dictionary*. Paper delivered at the XVII EURALEX International Congress, 6–10 September 2016, Tbilisi, Georgia. [Online] Available: <http://euralex2016.tsu.ge/publication.html>.