Thinking out of the box – perspectives on the use of lexicographic text boxes
Rufus H. Gouws and Danie J. Prinsloo
University of Stellenbosch & University of Pretoria

‘And they all get put in boxes, little boxes all the same’

Although text boxes have become a common phenomenon in dictionaries relatively little attention has been paid to their presentation and to the motivation for their use and the type of data to be included in a dictionary in this specific way. Text boxes are salient dictionary entries and as such they are used to place more than the default focus on a specific data item. Dictionaries offer a variety of data types in text boxes such as guidance in terms of sense, contrasting related words, restrictions on the range of application, register, pronunciation, et cetera. The default presentation seems to be as article-internal microstructural entries within a typical relation of lemmatic addressing. Whereas some text boxes present data relevant to only the specific article, other text boxes, i.e. those with a synoptic assignment, also have relevance for other articles, namely a hybrid addressing relation, presenting both immediate and distant addressing. As devices employed in an extended compulsory microstructure care should be taken that text boxes do not become part of the compulsory microstructure and in so doing lose their significance and decrease the emphasis on the data included in the text boxes. The added value of text boxes may never be undermined by an over exposure of this device. Using both micro- and macrostructural text boxes offers exciting possibilities. Where dictionaries have a text production function data could be included in a text box to emphasise the use or non-use of certain combinations and collocations as well as prescriptive guidance. Of real importance is that lexicographers should realise that text boxes are lexicographic devices that can really enhance the data transfer in dictionaries. Lexicographers should think out of the box and they should get out of the box of tradition and employ text boxes in bold, innovative and functional ways.

1. Introduction

One of the important paradigm shifts in modern lexicography is the acknowledgement that all articles in a given dictionary do not necessarily have to display the same structure. A sequence of articles in a single article stretch does not merely have to be a presentation of more of the same. Instead of a compulsory and consistently applied homogeneous article structure the lexicographer has the liberty to opt for a less rigid heterogeneous article structure in which the compulsory microstructure, the default version of the specific dictionary, can be supplemented by items representing an extended microstructure. Different types of micro- and article structures give lexicographers the opportunity to present the microstructural items in a way that would best fit the needs and reference skills of their intended target users. Yet again, the user-perspective plays a dominant role in the planning and eventual compilation of dictionaries. This also applies to the data distribution and data presentation in any given dictionary.

One lexicographic device frequently employed in the presentation of data presentation, is the lexicographic text box. Although text boxes have become a common phenomenon in dictionaries relatively little attention has been paid to their presentation and to the motivation for their use and the type of data to be included in a dictionary in this specific way. This paper firstly has a contemplative approach by taking a look at some aspects of the current use of text boxes but then follows a transformative approach by going further to present some ideas deemed necessary for an improved use of this lexicographic tool.

2. The presentation of text boxes

Text boxes are currently presented in a diverse way. The default presentation seems to be as article-internal microstructural entries within a typical relation of lemmatic addressing. This
procedure can be employed to position the text box in various article slots. However, the positioning of a text box must be done in accordance with the type of microstructure prevailing in the dictionary and with cognizance to the scope of the text box in its addressing relation to the rest of the article.

One of the most typical positions allocated to text boxes is in a slot at the end of the article, cf. (1) from HAT:

(1)

\textit{eg·ter (ietwat deftig)}

voegwoordelike bw. om ’n teëstelling uit te druk; ewenwel, nietemin, tog, nogtans: \textit{Ons het hom geroep; hy het egter nie geantwoord nie.}

\begin{itemize}
\item[(a)] in Afrikaans begin ’n sin nie met \textit{egter} nie (in Engels kan \textit{however} wel ’n sin inlei). Ook die volgende is foutief: \textit{Piet, egter, was nie teenwoordig nie.}
\item[(b)] aangesien \textit{maar en egter} dieselfde teenstellende funksie verry, word hulle nie in dieselfde sin gebruik nie, bv. *\textit{maar die ouers het egter nie saamgestem nie.}
\end{itemize}

This text box includes data addressed at the lemma sign. Its presentation at the end of the article implies that it has the full article in its scope.

Where a dictionary article displays an integrated microstructure a text box may support this structure by being positioned within a subcomment on semantics along with the paraphrase of meaning/translation equivalent and the complementary context or cotext entries. This close vicinity will ensure a relation of direct and unimpeded addressing between the text box and its lemmatic address, cf. the treatment in HAT of the lemmata \textit{verwys} and \textit{plasties} in (2):

(2)

\begin{itemize}
\item[a.] \textit{ver·wys} n.nw. [\textit{verwys}]
\begin{itemize}
\item[1] na ’n ander plek of persoon wys of stuur; refereer: \textit{Hy is na my verwys vir ’n getuigskrif.} • \textit{Ter strafssitting verwys} (==na ’n voorlopige ondersoek ’n verdagte persoon na die strafhof stuur.) • ’n Student na ’n hoofstuk in ’n handboek verwys. • \textit{Ek verwys na u brief van die 11de deser.}
\end{itemize}

\item[b.] \textit{plas·ties} b.nw., bw.
\begin{itemize}
\item[1] wat die eienskap besit om vorm te gee: \textit{Plastiese kunste soos keramiek, beeldhoukuns.} • \textit{Plastiese chirurgie} (==, waardeur misvormde of beskadigde liggaamsdele verbeter of vervang word, of waardeur iemand se uiterlike kosmeties verbeter word.)
\item[2] wat maklik vervorm kan word; kniebaar: \textit{Plastiese materiaal.}
\item[3] van, berei uit plastiek; plastiek==: \textit{Plastiese houers, pype, borde, koppies, sakke.}
\end{itemize}

\item[c.] \textit{Plasties} as s.nw. (i.p.v. plastiek), bv. ’hierdie pyp is van plasties’, is foutiewe Afrikaans.
\end{itemize}

In this article the text box is part of the first subcomment on semantics and only the first polysemous sense falls within its scope.

Here the integrated microstructure accommodates the text box in the third subcomment on semantics with the accompanying scope restriction. This ensures a direct and immediate addressing relation between the text box and its address and assists the user in an unambiguous retrieval of information.
Section 2. The Dictionary-Making Process

In (3) from the Macmillan English Dictionary for the lemma *above* an integrated approach results in the text box being presented as part of the comment on form, preceding the comment on semantics. This positioning can be justified because the data contained in the text box typically belongs to the comment on form:

(3)

In (3) from the Macmillan English Dictionary for the lemma *above* an integrated approach results in the text box being presented as part of the comment on form, preceding the comment on semantics. This positioning can be justified because the data contained in the text box typically belongs to the comment on form:

This position once again emphasises the importance of having a text box in a direct addressing relation.

Quite often text boxes are not positioned at the end of an article but at the end of a well-identifiable component of the article. As an example, where a word can be used in more than one part of speech form a text box, directed at the occurrence of the word in one specific part of speech, can immediately follow the treatment allocated to that part of speech occurrence of the word. In example (4), an abbreviated article from HAT, the word *reël* is treated in its occurrence as noun and as verb with a text box directed at the first component of the article in which the occurrence as noun is treated.

(4)

This position once again emphasises the importance of having a text box in a direct addressing relation.

Quite often text boxes are not positioned at the end of an article but at the end of a well-identifiable component of the article. As an example, where a word can be used in more than one part of speech form a text box, directed at the occurrence of the word in one specific part of speech, can immediately follow the treatment allocated to that part of speech occurrence of the word. In example (4), an abbreviated article from HAT, the word *reël* is treated in its occurrence as noun and as verb with a text box directed at the first component of the article in which the occurrence as noun is treated.

(4)

*reël*

- **s.nw. [reëls]**
  1 streep waarop geskryf word: Skryf op die reëls. • Trek reëls met 'n potlood om op te skryf. • 'n Notaboek met reëls.
  
  **....**

  5 wat gewoonlik gebeur; gewoonte, beginself; ook, aanvaarde manier van handel: *In die reël is sy uiers betroubaar (==, gewoonlik, normaalweg).* • Om dit te doen, is teen die reëls van wellewendheid. • As ons nou meer toegee, word hierdie soort optrede (die) reël.

  **Uitdr.:**

  *'n guilde reël*

  *'n verstandig rigsnor: die beste handelwyse.*
  
  **....**

  dis teen die reëls
dis in stryd met wat gewoonlik gedoen word.

  **Tussen die lynne lees word afgekeur.**

- **ww. [gereël]**
  1 skik, orden; inrig, organiseer; in orde bring, rangskik: *Alles is vir die party gereël.* • As sy die byeenkoms gereël het, sal alles in orde wees. •.......

  [Latyn *regula*]
Where procedures of lemma nesting and niching are employed to present horizontally ordered lemmata the data in the text box is part of the treatment directed at the main lemma and the nested/niched lemmata presented in a partial article stretch. The text box falls within the microstructural scope of all these articles. This leads to a text box that is positioned as an article-external micro structural item with the nested or niched lemmata as well as the preceding vertically ordered main lemma in its scope. Compare (5) from HAT:

(5)  
\textbf{Neger} s.nw. [Negers] (rasisties)  
1 boorling/Invowner van Midde-Afrika; lid van die swart rasse wat ongeveer noord van die Demokratiese Republiek van die Kongo voorkom.  
2 afstammeling van so ’n persoon: \textit{Die Negers van Amerika} \textit{(==tans meestal) • Afro-Amerikaners}  
\textit{(==genoem.)}  
\textit{[Latyn niger swart]}  
\textbf{Neger:} Negerdans, Negerin, Negermusiek, Negervriend  

\begin{tabular}{|l|}
\hline
Die woord Neger word as kwetsend ervaar en moet liefs vermy word.
\hline
\end{tabular}

The text box is directed at the relevant lexical item and applies to both the lemmata included in the niched partial article stretch and the lemma of the preceding fully treated article. The box is a typical microstructural entry but it is not positioned within an article. As an article-external microstructural entry it has more than one article in its scope.

Text boxes can also be allocated a position as immediate constituents of article stretches. A boxed partial article stretch can be found in the Tweetalige aanleerderswoordeboek/Bilingual Learner’s Dictionary in the article stretch ‘M’ for the lemmata Mr and Mrs. The text box contains the two rudimentary articles and is situated in its alphabetical place between the lemmata mow and much as in (6).

(6)  

\begin{tabular}{|l|}
\hline
\textbf{mow} verb (past tense mowed, past participle mown, present participle mowing).  
\textbullet{} Mr is the written abbreviation for Mister: \textit{Mr} (George) Smith.  
\textbullet{} Mrs is an abbreviated title that comes before the surname of a married woman: \textit{Mrs} (Aimee) Smith.  
\textbullet{} Much! baie “How ver is it met jou work”? “Ek het nog baie om te doen.” “How far are you with your work?”  
\textbullet{} Still have much to do.”  
\textbullet{} How much hoeveel “Hoewel kos die appels?”  
\hline
\end{tabular}

These boxed articles display a rudimentary microstructure because they contain fewer items than that allocated to the default compulsory microstructure.

Text boxes can also be employed as articles in the secondary macrostructure of a dictionary. The Reader’s Digest Dictionary displays two central columns on each page, flanked by columns in which, among other, text boxes, directed at articles in the central columns, are included, cf. (7):
Section 2. The Dictionary-Making Process

The secondary macrostructure, presented in the two outer columns, contains articles with lemmata from the primary macrostructure as guiding elements and the treatment focuses on a backtranslation of the translation equivalents presented in the articles of the primary macrostructure. In addition the secondary macrostructure contains alphabetically ordered text boxes that are immediate constituents of the relevant article stretches.

3. Single x synopsis text boxes

In LSP lexicography, cf. Bergenholtz/Tarp/Wiegand (1999) the valid distinction between single and synopsis articles was introduced. Whereas the treatment in a single article merely focuses on the lemma sign functioning as guiding element of the specific article, the addressing in a synopsis article exceeds the article and specific lemma sign and is also directed at other related lemmata. In a chemistry dictionary the definition given for a lemma like *acid* also applies to e.g. the lemmata *nitric acid* and *sulphuric acid* but the definition is not repeated in these articles. The lemma *acid* is the guiding element of a synopsis article and its definition is relevant for the interpretation of the lemmata *nitric acid* and *sulphuric acid*. The distinction single x synopsis is also relevant for text boxes. Whereas some text boxes present data relevant to only the specific article other text boxes, i.e. those with a synoptic assignment, also have relevance for other articles. In the *Tweetalige aanleerderswoordeboek/Bilingual Learner’s Dictionary* the article of the lemma *meat* includes the text box given in (8).

The meat from some animals has a different name from the animal itself: the meat from a cow is called **beef** and that from pig **pork**, but the meat from a lamb is called **lamb**. For fish and for birds such as chicken or **duck** the same word is used for both meat and animal.

Although the contents of the text box is directed at the lemma sign *meat* its application goes much further to impact on articles with lemma signs like **beef**, **cow**, **pig**, **pork**, **lamb**, **chicken**, **duck** as guiding elements. This is an example where a hybrid addressing relation, presenting both immediate and distant addressing, holds between the text box and the different addresses.

The notion of a synopsis text box can also be realised in a different way. In the following partial article stretch, given in an abbreviated form, the guiding elements of the default articles, representing three homonyms, have identical spelling but the syllable stress of the first form differs from that of the second and third forms. The first two articles are separated.
by a partial article stretch consisting of sublemmata. In the final article slot of the last homonym a text box is presented to make the user aware of the stress differences holding between the three main lemmata. This synopsis text box has the lemma of its own article as well as the lemmata of the immediately preceding articles, including the niched lemmata, in its scope, cf. (9):

(9)

\[\text{roman}^1 \text{s.nw;} \text{[roman] (Frans)}\]
1 verdigte prosaverhaal van betreklik groot lengte, waarin ’n aantal verbandhoudende episodes behandel word; ......

\[\text{roman;} \text{romankuns, romanliteratuur, romanreeks, romanskrywer, romanwerk}\]

\[\text{roman}^2 \text{s.nw;} \text{[romanne, romans] (Latyn)}\]
rooi seevis met ’n donkerblou streep tussen die oë, ’n lekker eetvis;: \text{Chrysoblephus laticeps}.

\[\text{roman}^3 \text{s.nw;} \text{[romanne, romans]}\]
enigeen van ’n paar harige spinnekopagtige diere; jagspinnekop, haarskeerder, vetvreter.
Let op die klemverskil tussen \text{roman}^1 \text{ en roman}^2 \text{ en roman}^3.

It has already been shown how the article of the lemma sign above in the Macmillan English Dictionary has a text box integrated into the comment on form. The same article contains a second text box which is partially directed at the comment on semantics of this article but also on the comment on semantics of the article of the lemma sign over, resulting in a synoptic text box as illustrated in (10):

(10)

\[\text{above} /\text{əˈbāv}/ \text{adj, adv, preposition} \]

5 louder or higher than another sound if you can hear one sound above other sounds, it is louder or higher than the other sounds: Carl shouted, but it was impossible to make himself heard above all the noise.

6 too good to do sth if you are above a particular type

Both above and over can be used to mean ‘at a higher level than something’: the light above/over the door. Use above when something is not directly over something else: on the hillside above the river.
Use over when something moves or stretches across the space above something: flying over London \text{ the bridge over the river.}

4. The contents of text boxes

Dictionaries offer a variety of data types in text boxes. In the Afrikaans to English section of the Reader’s Digest Dictionary the majority of text boxes guide the user by contrasting related words focusing on crucial differences and avoiding incorrect use as in (11).

(11)

\[\text{WORDS IN ACTION}\]

\[\text{Maanlig, maanskyn/maneswyn}\]
Maanlig (moonlight) is the ordinary word; maanskyn (moonlight, moonshine) has acquired romantic overtones.

\[\text{WORDS IN ACTION}\]

\[\text{mag; kan}\]
Under the influence of English the verb mag (may; is allowed to) is wrongly used instead of kan (may; can, is able to) …
Section 2. The Dictionary-Making Process

An analysis of the Afrikaans-English section of the Reader’s Digest Dictionary offers the results given in Table 1 and Figure 1 for the randomly selected alphabetic stretch ‘M’.

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast related words</td>
<td>37</td>
</tr>
<tr>
<td>Communication/application range or restrictions</td>
<td>28</td>
</tr>
<tr>
<td>Grammar data</td>
<td>17</td>
</tr>
<tr>
<td>Register and written/spoken</td>
<td>10</td>
</tr>
<tr>
<td>Spelling</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1. Alphabetic stretch M in the Afrikaans-English section of the Reader’s Digest

In the reverse side of the same dictionary a significant number of text boxes are also dedicated to the contrasting of related words but more attention is given to guidance in terms of communication/application range or restrictions. Compare, for example, the textboxes of *mean* and *maritime* in (12).

(12)

**WORDS IN ACTION**

*mean*

*He plays a mean violin* beteken nie dat die viool ordinêr of armoedig is nie, maar wel dat die person ’n besonder knap speler is. ….

**WORDS IN ACTION**

*maritime*

*Al beteken dit ook ‘aan die see geleë’ moet tog nie jou vakansiehuis aan die kus as ’n maritime cottage beskryf nie. …*

An analysis of the use of text boxes in the English-Afrikaans component of this dictionary for the alphabetic stretch ‘M’ leads to the results given in Table 2 and Figure 2:

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication/application range or restrictions</td>
<td>38</td>
</tr>
<tr>
<td>Contrast related words</td>
<td>23</td>
</tr>
<tr>
<td>Contrast BE, AE</td>
<td>12</td>
</tr>
<tr>
<td>Register</td>
<td>11</td>
</tr>
<tr>
<td>Syntactic restrictions</td>
<td>8</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>7</td>
</tr>
<tr>
<td>Grammar data</td>
<td>7</td>
</tr>
<tr>
<td>Spelling</td>
<td>3</td>
</tr>
<tr>
<td>Obsolete, antiquating</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2. Alphabetic stretch M in the English-Afrikaans section of the Reader’s Digest
In the Macmillan English dictionary guidance in respect of sense distinction and collocations constitute the content of most text boxes as for (13).

(13)

\[\text{mantle}^1\]

<table>
<thead>
<tr>
<th>Sense</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 authority/responsibility</td>
<td>4 piece of clothing</td>
</tr>
<tr>
<td>2 sth that covers/hides</td>
<td>5 part of the Earth</td>
</tr>
<tr>
<td>3 cover for flame</td>
<td></td>
</tr>
</tbody>
</table>

An analysis of the text boxes in the article stretch ‘M’ of the Macmillan English Dictionary is given in Table 3 and Figure 3:

<table>
<thead>
<tr>
<th>Senses</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metaphor</td>
<td>4</td>
</tr>
<tr>
<td>Collocations</td>
<td>19</td>
</tr>
<tr>
<td>History</td>
<td>12</td>
</tr>
<tr>
<td>Part of Speech</td>
<td>7</td>
</tr>
<tr>
<td>Get it right</td>
<td>7</td>
</tr>
<tr>
<td>[Not sure / mixed]</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other ways of saying</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>British versus American</td>
<td>1</td>
</tr>
<tr>
<td>Talking or writing about</td>
<td>1</td>
</tr>
<tr>
<td>Other ways of saying</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. Text boxes in M of the Macmillan English Dictionary
Section 2. The Dictionary-Making Process

A comparable analysis of other dictionaries will confirm some of these data types but will also introduce a wide-ranging selection of additional data types, e.g. the use of thesaurus boxes in the Longman SA School Dictionary. This evidence indicates that lexicographers frequently employ text boxes as lexicographic devices.

5. Why text boxes?

Text boxes are salient dictionary entries and as such they are used to place more than the default focus on a specific data item. As devices employed in an extended compulsory microstructure care should be taken that text boxes do not become part of the compulsory microstructure and in so doing lose their significance and decrease the emphasis on the data included in the text boxes. Lexicographers should make a clear distinction between using text boxes and using other lexicographic devices, e.g. labels, to focus the attention of the user on a specific item. The added value of text boxes may never be undermined by an over exposure of this device. Looking at the current situation where it is not always clear what motivated the use of a text box. The Oxford School Dictionary Afrikaans-Engels/English-Afrikaans employs a variety of shaded boxes, presented as usage notes, to guide users ‘on potential areas of difficulty, helping them avoid common mistakes’. However, some of the data presented in these shaded text boxes might have had a better effect had they been presented in the default article structure. The article of the lemma sign land\(^1\) contains three subcomments on semantics for the different senses of the word. The comment on form contains the plural form lands. Following the third subcomment on semantics a text box is presented with the entry ‘In sense 1, land cannot be used in the plural form’ This is the type of entry that does not necessitate a text box but could have been accommodated in the relevant subcomment, elevating the first subcomment to a subcomment on comment and form, e.g.: ‘land\(^1\) noun (lands) 1 (no plural) ....’ In various articles (cf. Commonwealth, north) the entry ‘no plural’ has been used as a default micro structural entry and in the case of land\(^1\) it could have sufficiently guided the user.

Where the importance or significance of the data in a text box does not really live up to the salience of the position of a text box as dictionary entry the lexicographers should consider to bring such data out of the box and into the default article slots. Text boxes should rather be reserved for data in need of a position of salience. But it is also important that lexicographers should employ text boxes in a more innovative way.

6. Getting out of the box

Lexicographers should get out of the box of tradition and employ text boxes in bold, innovative and functional ways. The primary question to all lexicographers planning, compiling or revising their dictionaries should be: what do I want my target user to be able to do with this dictionary? As a response to this question the lexicographer should ask: how can I ensure that the user retrieves the information from my dictionary needed to solve the problem that initiated the dictionary consultation procedure? Then the question can be put how text boxes can be employed in the data distribution structure of a dictionary to accommodate the kind of data that cannot be included in the default article structure or that needs additional emphasis?

The idea of using both micro- and macrostructural text boxes offers exciting possibilities. In modern-day printed dictionaries lexicographers are hesitant to label a word as being a neologism because the track record of actual usage it needs to qualify for inclusion in a
dictionary makes it hardly a neologism. However, dictionary users are often interested to know which ‘new’ words have been included in a new revision of a dictionary. In this regard macrostructural text boxes could be used to inform users of the ‘flavour of the edition’ words. Popular words making their lexicographic debut could be introduced in this way.

Text boxes could also be used to achieve specific lexicographic functions of a given dictionary. The strong semantic focus still prevailing in many monolingual dictionaries does not allow the lexicographer to include some encyclopaedic data, regarded as important to the user, within the default microstructure. Text boxes could be employed to accommodate data typically regarded as too encyclopaedic in nature for a given dictionary. This approach can also be applied in bilingual dictionaries to include cultural data that might be relevant for the user of that dictionary. Where dictionaries have a text production function data could be included in a text box to emphasise the use or non-use of certain combinations and collocations as well as proscriptive guidance.

Of real importance is that lexicographers should realise that text boxes are lexicographic devices that can really enhance the data transfer in dictionaries.

7. In conclusion

In the planning and compilation of dictionaries lexicographers should find ways and means to present data that can ensure an optimal retrieval of information by the users. This will lead to successful dictionary use. To achieve this, the use of a device like the text box needs to be reconsidered and put to use in innovative ways to focus the attention of users on data that needs special attention.
Section 2. The Dictionary-Making Process

References

Dictionaries


Other references