

# On Recent Proposals to Abolish Polysemy and Homonymy in Lexicography

Herman L. Beyer, *Department of Language and Literature Studies, University of Namibia, Windhoek, Namibia, and Department of Afrikaans and Dutch, Stellenbosch University, Stellenbosch, South Africa (hbeyer@unam.na)*

---

**Abstract:** Two articles appeared recently in *Lexikos* that propose the abolishment of homonymy and polysemy in lexicography, particularly in dictionaries with a text reception function only. This contribution identifies two main theoretical premises of the proposal in these articles and challenges them. They are: (i) a theory of the lemma as linguistic sign; and (ii) the results of dictionary criticism. Under examination, it is found that both premises fail to support the proposal with regard to polysemy. With regard to homonymy, the first premise is proven invalid, and the second is found to be valid. This implies that the theoretical basis for the proposal should either be reviewed (for which the lexicographical communication theory is offered), or the proposal should rely on the sole practical and unproven argument of data accessibility. The contribution simultaneously develops a potential broad framework for the lexicographical communication theory. The framework constitutes a lexicographical text grammar, which is presented as a parallel communication code to elements of the lexicographic text theory and linguistic grammars. It is argued that dictionary articles constitute texts in which these two grammars overlap to varying degrees, representing a hybrid form of textual communication.

**Keywords:** LEXICOGRAPHICAL COMMUNICATION THEORY, GRAMMAR, HOMONYMY, LEXICOGRAPHICAL COMMUNICATION, LEXICOGRAPHICAL GRAMMAR, LINGUISTIC SIGN, LINGUISTICS, POLYSEMY, SEMIOTICS, LEXICOGRAPHICAL TEXT THEORY

**Opsomming:** Oor onlangse voorstelle vir die wegdoen van polisemie en homonimie in leksikografie. Twee artikels het onlangs in *Lexikos* verskyn wat voorstel dat weggedoen word met homonimie en polisemie in die leksikografie, spesifiek in woordeboeke met slegs 'n teksresepsiefunksie. Hierdie bydrae identifiseer twee teoretiese hoofpremisses vir die voorstel en bevestig hulle. Die premisse is: (i) 'n teorie van die lemma as taalteken; en (ii) die resultate van woordeboekkritiek. By nadere ondersoek word bevind dat beide die premisse faal met betrekking tot polisemie. Met betrekking tot homonimie word die eerste premis as ongeldig bewys, en die tweede een word geldig bevind. Die bevindinge hou in dat die teoretiese basis vir die voorstel óf hersien moet word (waarvoor die teorie van leksikografiese kommunikasie aangebied word), óf op die enkele praktiese en onbewese argument van datatoeganklikheid moet steun. Terselfdertyd ontwikkel die bydrae 'n potensiele breë raamwerk vir die teorie van leksikografiese kommunikasie. Die raamwerk verteenwoordig 'n leksikografiese teksgrammatika, wat as 'n kommunikasiekode parallel tot elemente van die teorie van leksikografiese tekste en taalkundige grammatikas aangebied word. Daar word aangevoer dat woordeboekartikels uit tekste bestaan waarin hierdie twee

grammatikas in wisselende mates oorvleuel en as sodanig 'n hibriediese vorm van tekstuele kommunikasie verteenwoordig.

**Sleutelwoorde:** GRAMMATIKA, HOMONIMIE, LEKSIKOGRAFIESE GRAMMATIKA, LEKSIKOGRAFIESE KOMMUNIKASIE, POLISEMIE, SEMIOTIEK, TAALKUNDE, TAALTEKEN, TEORIE VAN LEKSIKOGRAFIESE KOMMUNIKASIE, TEORIE VAN LEKSIKOGRAFIESE TEKSTE

## 1. Introduction

Two articles appeared recently in *Lexikos* that propose the abolishment of homonymy and polysemy in lexicography. The first article claims that "polysemy and homonymy do not exist" and that "in lexicography we can do well without these terms" (Bergenholtz and Agerbo 2014: 31). The apparent overall rejection of these concepts is also clear from the title of the article: "There is No Need for the Terms Polysemy and Homonymy in Lexicography". The second article builds on the work presented in the first, but it displays a more moderate attitude towards the relevant concepts, stating that "the existence of homonymy and polysemy as concepts in the field of linguistics is acknowledged," that arguments can be advanced for the abolishment of the "traditional distinction between homonymy and polysemy", and that the proposal to abolish polysemy and homonymy is limited to "the communicative situation where a mother-tongue speaker or a foreign language speaker encounters text reception problems" (Bergenholtz and Gouws 2017: 110, 112, 125).

The first article (Bergenholtz and Agerbo 2014) describes three models according to which homonymy and polysemy can be dealt with in dictionaries:

- Model I: the "traditional" model, where homonyms are linguistically distinguished as formally identical but separate lexemes on the grounds of semantic non-relatedness and/or different etymologies, each represented by a separate lemma sign and dictionary article, and polysemy on the grounds of the relatedness of semantic values that can be assigned to one lexeme, i.e. polysemic values presented in one article.
- Model II: a model that rejects the notions of homonymy and polysemy, and assigns only one semantic value to a given lemma: In model I, a set of two homonyms, each with three polysemic values, would be presented as two formally identical lemma signs representing each of the homonyms, each lemma sign with its own article containing three polysemic values. Given model II, the same set of lexical items would be presented as six formally identical lemma signs, each with its own article representing one semantic value only; no polysemic or homonymic relations would be signalled.
- Model III: "words that are orthographically similar but have different inflectional paradigms (also within the same part of speech) are defined as

homonyms, whereas orthographically similar words belonging to the same part of speech and with the same inflectional paradigm are defined as polysems [sic]" (Bergenholtz and Agerbo 2014: 29).

In the first article, model III is favoured because it is "closer to the solution that dictionary users are familiar with" (Bergenholtz and Agerbo 2014: 34).

The second article (Bergenholtz and Gouws 2017) attempts to build a case for the model II solution on the basis of two main theoretical premises:

- a lexicographic theory of the lemma as linguistic sign by Bergenholtz and Agerbo (2014);
- criticism of a selection of Danish and English dictionary articles.

The first aim of this contribution is to challenge these premises and therefore the validity of model II on the following points, which will be elaborated in the indicated sections to construct the argument:

- Bergenholtz and Agerbo's (2014) lexicographic theory of the lemma as linguistic sign is flawed as well as irrelevant: section 2.
- The model II solution does not address Bergenholtz and Gouws's (2017) criticism of existing dictionary articles, but merely transfers a number of perceived metalexigraphic problems from one lexicographic text structure type to another, potentially adding unnecessary complications for lexicographical communication in the process: section 3.

In the course of arguing the above points, a potential broad framework for the theory of lexicographical communication (or: lexicographical communication theory), as introduced by Beyer (2014) and Beyer and Augart (2017), is developed in subsection 2.3 on the basis of linguistic grammar. This is the second aim of this contribution. The basic tenets of the lexicographical communication theory are that (i) at its core, lexicography is an exercise in communication, and (ii) this communication is indirect communication mediated by text (Beyer and Augart 2017: 8). The description of dictionary article text structures in the theory of lexicographic texts (or: lexicographic text theory), developed primarily by H.E. Wiegand within a general theory of lexicography, is "completely taken over from formal syntax" (Wiegand 1996: 136), which can be observed in that theory's presentation of (abstract) microstructures in the form of hierarchical tree structures similar to the presentation of sentence constituents in context-free (i.e. phrase structure) grammars (cf. Gouws, Heid, Schweickard and Wiegand 2013: articles 3–10). This method has inspired the grammar framework that will be presented for the lexicographical communication theory. Consequently, similarities between the framework presented and the relevant elements of the lexicographic text theory will be evident, and will be accounted for where necessary for the purposes of the discussion.

## 2. Bergenholtz and Agerbo's lexicographic theory of the lemma as linguistic sign

Bergenholtz and Agerbo (2014) employ De Saussure's (2013) model of the *linguistic sign* to evaluate the status of a set of word types. This evaluation forms the main premise of their proposal to abolish the concepts *polysemy* and *homonymy* in lexicography. It will be shown in this section that this premise is conceptually flawed and that therefore the conclusion based on it is logically false. First, however, it is necessary to clarify the relevant terms within the Saussurean model.

### 2.1 (Linguistic) sign, code and sign system

The term *sign* is defined as follows by Bock (2014: 57):

def<sub>1</sub> A **sign** is something that represents or stands for something else, where the 'something else' may refer to an idea, object, value or phenomenon. The sign is not 'the something' itself, but rather a representation of that thing.

While signs in themselves have *values*, they can only assume *meaning* in relation to other signs (De Saussure 2013: 134ff). This requires signs to possess paradigmatic and syntagmatic properties which allow them to function in various relations with other signs (cf. De Saussure 2013: 144-148). The sum of the paradigmatic and syntagmatic properties of all signs that belong to the same sign system can be referred to as that sign system's *code*. A *sign system*, then, consists of two primary components: (i) a set of signs, and (ii) a set of rules, known as a code, which describes the paradigmatic and syntagmatic properties of the signs that allow them to be combined to signal meanings (cf. Bock 2014: 57-58). In linguistic terms, *sign system* is equated to a particular language (e.g. English), *set of signs* is equated to that language's lexicon, and *code* is equated to the language's grammar (Bock 2014: 57-58).

A *linguistic sign* is a sign (<def<sub>1</sub>) that functions within a linguistic code: English words are linguistic signs inasmuch as they function within the linguistic code of the English grammar. De Saussure (2013: 77) defines a linguistic sign as a combination of two "intimately linked" elements, namely a "concept and a sound pattern"<sup>1</sup>. Chandler (2007: 14ff) uses the equivalent terms *signified* and *signifier*, and Bergenholtz and Agerbo (2014) use the equivalent *content* and *expression*. Although this article is a response to Bergenholtz and Agerbo (2014) and Bergenholtz and Gouws (2017), Chandler's terms will be used in the following discussion, because they bear the closest resemblance to the original terms proposed by De Saussure (i.e. French *signifiant* and *signifié*). A (linguistic) sign, then, is "the whole that results from the association of the signifier [expression] with the signified [content]" (Chandler 2007: 15), which can, in the

style of De Saussure (2013: 77), be presented in the following diagram:



**Figure 1:** The constitution of the *sign*, according to De Saussure (2013), in the terms of Chandler (2007)

An alternative presentation of the same concept in table format, which will be used in this article, looks as follows:

**Table 1:** An alternative representation of the concept *sign* according to De Saussure (2013), in the terms of Chandler (2007)

Sign	
Signifier	Signified

## 2.2 Bergenholtz and Agerbo's application of the term *linguistic sign*

Bergenholtz and Agerbo (2014: 31) claim that "we cannot speak about polysemy and homonymy if we relate these terms to the linguistic sign. However, in lexicography we can do well without these terms." This claim is based on the following argument (Bergenholtz and Agerbo 2014: 31):

quote<sub>1</sub> In the lexicographical tradition [...] a lemma is not a linguistic sign because a lemma can represent different lexical words (sometimes it represents only one lexeme, in other cases it represents several lexemes). Hence, there is no solidarity between one expression [signifier] and one content [signified].

The argument is followed by the model II proposal as a "radical solution [...] where we discard polysemy and homonymy and instead connect each lexical word to its own lemma," because only then "the lemma could be defined as a linguistic sign" (Bergenholtz and Agerbo 2014: 31).

In the following subsections different aspects of Bergenholtz and Agerbo's application of the term *linguistic sign* will be scrutinised.

### 2.2.1 All (types of) words are linguistic signs

The model II solution depends on Bergenholtz and Agerbo's evaluation of the lemma as a linguistic sign in certain uses and not a linguistic sign in other uses.

This evaluation is conducted within the context of a broader evaluation of the status of a set of word types vis-à-vis the concept *linguistic sign*, namely so-called orthographic words, text words, grammatical words, lexical words (lexemes) and dictionary words (lemmata) (Bergenholtz and Agerbo 2014: 30-31). The broader evaluation can be summarised in the following table:

**Table 2:** Summary of Bergenholtz and Agerbo's (2014) evaluation of a set of word types

Word type	Description	Linguistic sign?
orthographic word	A sequence of letters between blanks and sentence signs (like commas), also search strings in e-dictionaries.	No
text word	A concrete word in a text with a specific spelling, meaning, grammar, etc.	Yes
grammatical word	An expression with at least one nucleus morpheme and for adverbs, verbs and nouns also at least one grammatical morpheme. A grammatical word belongs to a certain inflection paradigm.	No
lexical word (lexeme)	An abstraction for an amount of grammatical words belonging to the same stem and the same inflection paradigm.	Yes
lemma	An abstraction for an amount of grammatical words, but it is not the same as a lexical word, because, contrary to lexical words, different stem meanings do not result in different lemmata.	No

In every case in table 2, a word type is judged to be a linguistic sign or not on the basis of the perceived presence or absence of a combination of signifier and signified to form a sign. In fact, each judgement is based on the prerequisite for the existence of a sign per se (cf. def<sub>1</sub>; Chandler 2007: 15), and not necessarily of a linguistic sign, because the requirement of functioning specifically in a linguistic code is not tested (except perhaps with the type *text word*).

Table 2 clearly shows that every word type *represents or stands for* some concept as summarised under the heading "Description" (<def<sub>1</sub>; Chandler 2007: 15), which presupposes signification, i.e. a combination of signifier and signified, *in every case*. This is an obvious refutation of every "No"-judgement, i.e. of every judgement that a particular word type is not a linguistic sign. Moreover, Bergenholtz and Agerbo's (2014: 31) argument in quote<sub>1</sub> above that "a lemma is not a linguistic sign because a lemma can represent different lexical words" is self-contradictory: If a lemma (or any other word type) *represents or stands for* x, y and/or z, it follows that it is a sign. This can be illustrated by listing an exem-

plar of each word type and indicating how that exemplar is a sign by aligning its signifier and a representation of its signified, as in table 3:

**Table 3:** Examples of word types and their sign values

Ref.	Word type	Sign value	
		Signifier	Representation of the signified
1	orthographic word	<i>flush</i>	'the grapheme sequence ⟨f, l, u, s, h⟩'
2	text word	<i>flushes</i>	' <i>flushes</i> in "Tom has played two flushes so far"'
3	grammatical word	<i>flushes</i>	'the grammatical word paradigm { <i>flushes</i> (n., pl.: 'reddening'), <i>flushes</i> (n., pl.: 'hand of cards'), <i>flushes</i> (n., pl.: 'piece of wet ground')}'
4	lexical word (lexeme)	<i>flush</i>	'the inflection paradigm { <i>flush, flushes</i> '}
5	lemma	<i>flush</i>	'the lexeme <i>flush</i> '

Table 3 shows the various signs' values. Additionally, each of the signs can be proven to be a *linguistic sign*, because each can function in terms of its word type and assume *meaning* in paradigmatic and syntagmatic relations to other signs in the code of the English grammar. More directly, the mere fact that each category could be designated a type of *word* indicates the linguistic sign status of every category member. Compare their respective occurrence in the following grammatical English sentences (numbered in correspondence to "Ref." in table 3) (cf. also Murphy 2010: 11f and Cruse 2011: 47):

- (1) [The orthographic word] *flush* consists of five graphemes.
- (2) [The text word] *flushes* in "Tom has played two flushes so far" means 'more than one hand of cards all of the same suit'.
- (3) [The grammatical word] *flushes* represents a grammatical word paradigm.
- (4) [The lexeme] *flush* represents an inflection paradigm.
- (5) [The lemma] *flush* represents a lexeme.

Sentences (1) to (5) demonstrate that each word functions not only as a sign, but also as a linguistic sign.

The conclusion is therefore that, in the first place, and contrary to Bergenholtz and Agerbo's (2014) evaluation, all word types in table 2 are signs because signification is proven in all cases. In the second place, they are specifically linguistic signs because they function within a linguistic code, in this case that of English.

There are, however, more obvious and general problems with Bergenholtz and Agerbo's (2014) lexicographic theory of the lemma as linguistic sign. These are dealt with in the following subsections.

### 2.2.2 Representation of the signified is not the signified

Compare the following dictionary article from the *Oxford South African Concise Dictionary* (Van Niekerk and Wolvaardt 2010: 449):

da<sub>1</sub> **flush**<sup>3</sup> ■ n. (in poker or brag) a hand of cards all of the same suit.

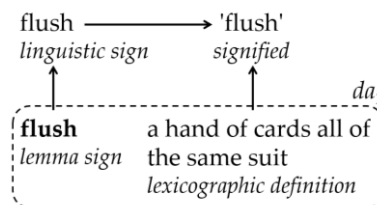
Leaving the homonymy indicator |<sup>3</sup>| and the register item |(in poker or brag)| aside for the moment, Bergenholtz and Agerbo (2014) would argue that the lemma in da<sub>1</sub> is a linguistic sign because there is solidarity between one expression (signifier: the lemma sign form) and one concept (signified: |a hand of cards all of the same suit|). Semiotically speaking, however, there is a fundamental problem with this argument.

The signifier is the "sensory part" of the sign which "implies reference to the whole [i.e. the sign itself — HLB]" (De Saussure 2013: 77). It is "the *material (or physical) form* of the sign — it is something which can be seen, heard, touched, smelled or tasted" (Chandler 2007: 15). The signified is "generally of a more abstract kind" (De Saussure 2013: 76). Chandler (2007: 16) explains that De Saussure's "*signified* is not to be identified directly with [...] a referent but is a *concept* in the mind — not a thing but a notion of a thing." (Cf. also Peirce 1985, Sebeok 2001: 5-6, Danesi 2004: 4-6, Hébert 2018.)

The point being made is that whereas the signifier has a physical form, the signified is abstract: It is physically imperceptible. A lexicographic paraphrase of meaning — ostensibly referred to as a *meaning* by Bergenholtz and Agerbo (2014) and Bergenholtz and Gouws (2017)<sup>4</sup> — is a physically perceptible signal; therefore, it is impossible to equate it to a signified (or, in Bergenholtz and Agerbo's (2014) terms, a *content*). Rather, the lexicographic definition |a hand of cards all of the same suit| in da<sub>1</sub> constitutes a complex sign (in the form of a syntagma) associated with the signified 'flush' in the very same way that the lemmatically represented word form *flush* constitutes a simple sign associated with the same signified.<sup>5</sup> The logical conclusion is that the lemmatically represented form and the lexicographic definition are two *equivalent* signs. This fact becomes clearer when the lexicographic definition is replaced by a word synonym in a monolingual dictionary and by a translation equivalent in a bilingual dictionary. (Bergenholtz and Agerbo (2014: 34) assert that their theory applies to "both monolingual and bilingual dictionaries; there are no significant differences".) As wholes, then, the lemma sign and lexicographic definition in da<sub>1</sub> are indirectly equivalent signs: the lemma in the form of a sign representing a simple linguistic sign with the value 'flush'<sub>i</sub> and the lexicographic definition in the



form of a syntagma as signifier of a complex sign with the meaning 'flush'. The relevant relations can be represented in figure 2:



**Figure 2:** A simplified representation of the semiotic relations involving the lemma sign and lexicographic definition in  $da_1$ , and the signified ("x → y" = x refers to y)

It follows that a dictionary article, or any text for that matter, cannot contain a signified/content. A monolingual dictionary article simply coordinates signs in one and the same sign system that share the same signified, in exactly the same way that a bilingual dictionary article coordinates signs in a source sign system with signs in a target sign system that share signifieds, explained in linguistic terms by Zgusta (1971: 294) as the semantic coordination of a set of lexical items in one language with that of another. With regard to the purposes of a specific dictionary, the lexicographic definitions, word synonyms and/or translation equivalents function as representations of (or *comments* on) the signifieds associated with the lemmatically represented signs; they are not — and cannot possibly be — the signifieds in themselves. In the case of a dictionary article of a polysemic lemma, the lemma sign represents a set of linguistic signs with identical signifiers (which, in model I, normally constitute a lexeme), while the semantic and pragmatic comments on the various identified senses represent the set of signifieds co-constituting the respective signs. From the number of senses so distinguished, together with data on inflection, the number of signs that are (partially) represented in the dictionary article can be inferred, if necessary, although this would hardly fulfil one of the purposes of a dictionary with only a text reception function. This, in short, is the semiotic nature of the typical dictionary article as text.

The above exposition clearly shows that the semiotic requirement that a dictionary article should represent "solidarity between one expression [signifier] and one content [signified]" (Bergenholtz and Agerbo 2014: 31) is untenable, regardless of the dictionary's purposes. In semiotic terms, a monosemic dictionary article in effect coordinates at least two signifiers that can signify the same signified. This represents one of the core problems in lexicography: how to represent the signified of a particular signifier in terms of another signifier or signifiers.

A further problem with the semiotic requirement pertains to the question of inflected word forms as linguistic signs, which is the focus of the next subsection.

### 2.2.3 Inflected words are (also) linguistic signs

Gallmann (1991) assigns all formal (i.e. physical) features of the linguistic sign to the signifier, while all grammatical and semantic features are assigned to the signified, in line with the concept of the sign (cf. again Peirce 1985, Sebeok 2001: 5-6, Danesi 2004: 4-6, Chandler 2007: 15-16, De Saussure 2013: 77, Hébert 2018). Therefore, inflected and non-inflected word forms constitute separate linguistic signs, since an inflected word form as sign differs both in terms of signifier (i.e. formal features) and signified (i.e. grammatical features) from its non-inflected form. Bergenholtz and Agerbo (2014: 30) also evaluate so-called text words, which include inflected forms, as linguistic signs (cf. table 3 and sentence (2) in 2.2.1). This can be illustrated with a simple example in table 4:

**Table 4:** Inflected and non-inflected word forms as separate linguistic signs

Sign	
Signifier	Representation of the signified
<i>ampersand</i>	'&'
<i>ampersands</i>	'& & ...'

Bergenholtz and Gouws (2017: 125) regard inflected forms as "different variant forms of the expression [signifier] with the same contents [signified]." From the above it is clear that this is an untenable position. It also contradicts Bergenholtz and Agerbo's (2014: 30) evaluation of text words as linguistic signs. Even orthographic variants, like *realise* and *realize*, are separate signs: Although they share the same signified, they have distinctive signifiers. After all, a (linguistic) sign exists only as "solidarity between *one* expression [signifier] and *one* content [signified]" (Bergenholtz and Agerbo 2014: 31; my emphasis — HLB). Bergenholtz and Gouws's mistaken semiotic definition of inflected forms seems to originate from Bergenholtz and Agerbo's (2014: 30) evaluation of a lexeme as a linguistic sign (cf. table 2), which is of course correct in itself; however, a lexeme's signified constitutes an entire inflection paradigm and not only the stem of such a paradigm (cf. table 3). It would seem that properties of the concept *lexeme* (a linguistic notion) have been confused with that of the concept *sign* (a semiotic notion).

If Bergenholtz and Agerbo's (2014) semiotic requirement that a lemma should be a linguistic sign with one signifier and one signified is to be met, then it follows that every inflected word form should also be lemmatised instead

of merely indicating inflection possibilities in the article of a stem. This is obviously not Bergenholtz and Agerbo's (2014) and Bergenholtz and Gouws's (2017) positions, from which it would appear that they contradict their own requirements. Therefore, Bergenholtz and Agerbo's (2014: 34) claim that model II is not "connected to any theoretical contradictions" does not hold water.

Besides the foregoing, it will be argued in the following subsection that typical lexicographical communication, especially via the medium of the typical dictionary article, is conducted within a sign system that is different from the natural language that is the object of the lexicographical communication in a particular instance. This implies that in lexicographical communication the lemma is in fact *not* a *linguistic sign*, but a sign in a different code, namely a lexicographical code, and is therefore a *lexicographic sign*.

### 2.3 The lemma as non-linguistic sign (in a linguistically-based theory of lexicography)

The lexicographical communication theory takes a global view of the potential of linguistic theory for meta-lexicography, i.e. linguistic theory not merely to explain the representation of lexical data in dictionaries, but also to form a basis for explaining how lexicographical communication functions (cf. Beyer 2014: 40). An attempt to construct such a basis will be outlined in this subsection as part of the discussion of the lemma as sign. Although the linguistic perspective is inspired by the lexicographic text theory, there are important areas of divergence between the lexicographic text theory and the lexicographical communication theory, as will be indicated where relevant.

#### 2.3.1 A lexicographic sign system

The fact that dictionary articles typically comment on the lexical features of a particular natural language obscures the fact that such comments are typically not encoded in that language, but in a hybrid sign system that merely partially resembles and overlaps with the relevant language, yet is significantly distinct from it. Compare the following two texts (*text<sub>2</sub>* being a slightly adapted version of a dictionary article from the *South African Oxford Secondary School Dictionary* (Reynolds 2006: 57)):

*text<sub>1</sub>* This is a paragraph about the word *bigwig*. The word *bigwig* is a word in English, and it is spelt as b, i, g, w, i, g. It is a noun. It is also an informal word, so be careful not to use it in a formal context; if you hear it or read it in a text, you will know that the speaker or author is using informal language in that instance. The word *bigwig* has only one semantic value, namely 'an important person'.

*text<sub>2</sub>* **bigwig** *n.* (*informal*) an important person

Text<sub>1</sub> is a text in natural language which adheres to the grammar of English. Text<sub>2</sub> obviously does not adhere to the grammar of English, yet it successfully communicates the same contents than text<sub>1</sub> does — but only for someone who knows how to interpret it. A literate mother-tongue speaker of English would easily interpret text<sub>1</sub> fully and correctly, but this does not imply that they would be able to fully and correctly interpret text<sub>2</sub>. Conversely, it is possible for someone who does not know English at all to at least partially interpret text<sub>2</sub> correctly and even to answer a limited set of user questions (e.g. that the form *bigwig* is a lexeme in English and that it has only one sense), provided that they are "text<sub>2</sub>-literate", in spite of the fact that they would not be able to interpret text<sub>1</sub> at all. Since humans make meanings through the creation and interpretation of signs (Sebeok 2001, Chandler 2007: 14), human communication requires sign systems. Because text<sub>2</sub>, which seems to be an English text, successfully communicates only between parties with some type of competence in addition to their competence in English, it follows that text<sub>2</sub> adheres to a sign system that is at least partially different from English.

The lexicographic text theory would argue that text<sub>1</sub> has been subjected to textual condensation in a process of lexicographic textualization in order to produce text<sub>2</sub>, which means that text<sub>2</sub> is some condensed version of text<sub>1</sub> (cf. Wiegand 1996a). Textual condensation would involve operations identified as shortening, abbreviating, omitting, shifting, substituting, summarising and embedding (Wiegand 1996a: 139). Some of these operations correspond to a greater or lesser degree to some of the operations identified and described in text linguistics, particularly abbreviation, substitution and ellipsis. However, the critical distinction is that text linguistics explains the relevant operations within the framework of the grammar of the relevant language, for example De Beaugrande and Dressler (1981) with regard to English, and Carstens (1997) with regard to Afrikaans. In contrast, the operations of textual condensation that would render text<sub>2</sub> as a condensed version of text<sub>1</sub> cannot be explained within the framework of the grammar of English. It follows then that text<sub>1</sub> and text<sub>2</sub> are created within the frameworks of different codes: text<sub>1</sub> within the framework of the grammar of English, and text<sub>2</sub> within the framework of some other code. This fact has required the lexicographic text theory to develop elaborate sub-theories of textual condensation (cf. Wiegand 1996a) and addressing structure (cf. Wiegand and Gouws 2013) to construct an inter-code bridge between text<sub>1</sub> and text<sub>2</sub>. These sub-theories in fact amount to the description of an alternative code to the grammar of English in order to make the rendering of text<sub>2</sub> possible. For this reason, the lexicographical communication theory does not recognise text<sub>2</sub> as any *version* of text<sub>1</sub>, but rather views text<sub>1</sub> and text<sub>2</sub> as distinctly separate texts that happen to encode the same set of lexicographic messages by means of distinctly separate sign systems: text<sub>1</sub> by means of the English language, and text<sub>2</sub> by means of a lexicographic sign system (which, in this case, overlaps with English in some ways), effectively making text<sub>1</sub> and text<sub>2</sub> *textual translation equivalents* of each other.

Although text<sub>2</sub> does not adhere to the grammar of English but ostensibly contains English words and even an English syntagma, it might be argued that it constitutes a version of text<sub>1</sub> because the reader can successfully interpret text<sub>2</sub> through processes of inference such as described by for example the theory of conversational implicature (cf. Grice 1991) and relevance theory (cf. Sperber and Wilson 1995, Clark 2013), to arrive at the propositions in text<sub>1</sub>. In this regard Sperber and Wilson (1995: 12-13) note the following:

Inferential and decoding processes are quite different. An *inferential process* starts from a set of premises and results in a set of conclusions which follow logically from, or are at least warranted by, the premises. A *decoding process* starts from a signal and results in the recovery of a message which is associated to the signal by an underlying code, and signals do not warrant the messages they convey.

It is clear that the highly sophisticated and intricate lexicographic text theory has developed a general code for lexicographic texts, because every functional text segment identified and described by the theory is assigned a specific unit of lexicographic data that it transmits. This means that there is a fixed association between signal and message, and that the receiver of such a text *decodes* the signal to recover the lexicographic message. Therefore, during optimal lexicographical communication, encoding and decoding takes place rather than implicature and inferencing. This implies "an underlying code", which, as has been seen, is not the grammar of English, but a distinct lexicographical code.

When text<sub>1</sub> and text<sub>2</sub> are evaluated against the foregoing argument, the conclusion is that text<sub>1</sub> is an English text, but that text<sub>2</sub> is not an English text, although it is a text about English. It is clear that there is an overlap of codes (and sign systems) in text<sub>2</sub>, but this in itself is not an unusual phenomenon. Although it is not equally evident, there is also an overlap of codes in text<sub>1</sub>. Chandler (2007: 149) points out that "various kinds of codes overlap, and the semiotic analysis of any text or practice involves considering several codes and the relationships between them." Based on a range of code typologies found in the literature of semiotics, Chandler (2007: 149-150) distinguishes between three main classes, of which two are relevant for the current discussion, namely:

- **social codes**, including natural/verbal language (with phonological, syntactic, lexical, prosodic and paralinguistic subcodes), bodily codes, commodity codes and behavioural codes;
- **textual codes**, including scientific codes, aesthetic codes, genre codes, rhetorical codes, stylistic codes and mass media codes.

A language like English obviously belongs to the class of social codes, but text<sub>1</sub> is created through an overlap between the social code and a particular textual code in order to produce a paragraph. Arguably, the social code is the primary code and the textual code is the secondary code (cf. also De Saussure 2013 on

the spoken vs. written modes of natural language). Given that lexicographical communication almost exclusively takes place through the medium of specialised types of text (and not in sound form as in the case of natural language), it can be argued that a particular textual code (which is significantly different from that of text<sub>1</sub>, even to the extent that it in fact constitutes a different sign system) is the primary code of text<sub>2</sub>, which is overlapped to a certain degree by a social code, in this case English. Therefore, lexicographical communication like in text<sub>2</sub> takes place by means of a distinct lexicographic sign system. The sign systems that have been studied the most extensively and scientifically are natural languages because they are the "primary and most pervasive" codes in any society (Chandler 2007: 149). This has given rise to the extensive discipline of modern linguistics. It therefore makes sense to consider the potential value of linguistic theory in attempting to describe a lexicographic sign system. Such a specific text-based sign system could be referred to as a *lexicographic language*, or *l-language* (as opposed to a natural language, or "n-language"). It should be noted that, because of its text-based nature, an *l-language* is not a type of natural language and is not represented by an element of Chandler's class of social codes or described by linguistics; rather, it is represented by a type of textual code. The sign |■| in da<sub>1</sub> (cf. 2.2.2), for example, is not a linguistic sign, but it belongs to the lexicon of the relevant *l-language*. The partial term *language* is merely used for lack of a better alternative.

With regard to an *l-language* as sign system, *set of signs* is equated to *lexicographic lexicon* (or: *l-lexicon*), and *code* is equated to *lexicographical grammar* (or: *l-grammar*). The sign |■| in da<sub>1</sub>, for example, would be an element of the *l-lexicon* of the *l-language* used in the dictionary involved. In the following section natural language grammars will be highlighted briefly to provide a background for the introduction of an *l-grammar* in section 2.3.3.

### 2.3.2 Natural language grammars

Traditionally, a natural language grammar consists of the following components:

- *phonetics* and *phonology*, describing the sound system of the language;
- *morphology*, describing word formation;
- *syntax*, describing sentence formation;
- *semantics*, describing the meaning of words and sentences;
- *pragmatics*, describing the use of the language in context.

In a traditional grammar, the largest unit of study is any of the various types of sentence. Consider the following simple English sentence:

s<sub>1</sub>     A lemma represents a lexeme.

An English phonetics and phonology would study the speech sounds and phonological processes involved in pronouncing the sentence, for example that *a* is pronounced [ə], and that [ə] does not assimilate with the following sound [l] because it is a lateral.

Morphology would for example note that the verb *represents* is an inflected form of *represent*, and that *represent* is a diachronic derivative of the order [present]<sub>V</sub>.

Syntax would identify and describe the order of the various sentence constituents, for example in the following linear representation of the constituent syntax of *s*<sub>1</sub>:

[<sub>S</sub>[<sub>NP</sub>[<sub>DET</sub>[<sub>ART</sub> A]] [<sub>N</sub> lemma]]<sub>NP</sub> [<sub>VP</sub>[<sub>V</sub> represents] [<sub>NP</sub>[<sub>DET</sub>[<sub>ART</sub> a]] [<sub>N</sub> lexeme]]]]

From the above description the following set of syntactic rules could be derived: S → NP VP; NP → DET N; DET → ART; VP → V NP

Semantics would describe the semantic values of respective words and the propositions that are encoded in the sentence, and the relations between them, for example:

Lexical semantics: *lemma* → [- animate], [+ abstract], [+ countable], etc.

Sentence semantics: REPRESENT(a lemma, a lexeme)

Pragmatics would describe the meaning of the sentence as an utterance in context, for example that it constitutes an assertion, that its interpretation can be described in terms of a cooperative principle of communication, how the subject relates to interlocutors' common ground through reference by means of the indefinite article *a*, etc.

In addition to traditional sentence-based grammars, the discipline of text linguistics expands the basic object of linguistic enquiry to the text or discourse as a whole (cf. De Beaugrande and Dressler 1981, Carstens 1997). According to Carstens (1997: 53-59), Van Dijk (1972) had a tremendous influence on the development of text research, particularly with his notion of a *text grammar*, which proposes that, like sentences, texts can be described in terms of a type of formal grammar, facilitated by a distinction between textual surface and deep structures. The following tasks are assigned to a text grammar by Van Dijk (1972: 11):

- to formally enumerate all and only grammatical texts of a language;
- to assign structural descriptions to each of these generated texts;
- to formulate rules in terms of which the textual deep structure can be derived from the textual surface structure; and
- to investigate textual surface structures.

The potential of a text grammar for lexicographic theory development is par-

ticularly attractive to the lexicographical communication theory, especially because of the generally highly conventionalised nature of lexicographic texts as it relates to the second basic tenet of the theory. Within the broader discipline of text linguistics, the seven elements of textuality, i.e. cohesion, coherence, intentionality, acceptability, informativity, situationality and intertextuality (cf. De Beaugrande and Dressler 1981, and Carstens 1997), are also of central relevance.

### 2.3.3 A text grammar as a lexicographical code

In line with the object of study in text linguistics, the largest unit of study in an *l*-grammar is any of the various types of *lexicographic text*, which entails that an *l*-grammar is essentially a type of text grammar. The lexicographic text theory, having empirically identified and meticulously described a range of lexicographic text types, provides a solid foundation in this regard.

Adopting and adapting concepts from linguistic theory, it is proposed that an *l*-grammar consists at least of the following components:

- an *l-syntax*, describing the order of the various text elements in a lexicographic text and the textual surface structure relations among them;
- an *l-morphology*, describing the formation of lexicographic items contained in a lexicographic text;
- an *l-semantics*, describing the lexicographic propositions encoded in lexicographic items and the textual deep structure relations among them;
- an *l-pragmatics*, describing the communicative functions of the various text elements and the textual deep structure relations among them.

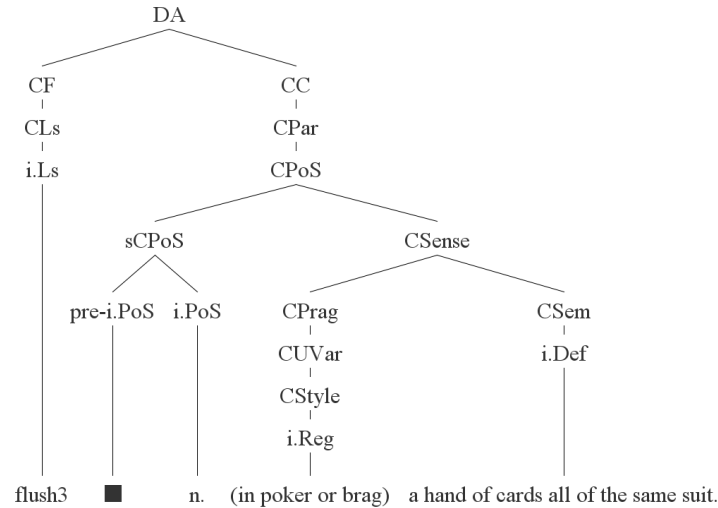
An *l-phonology* could be added in cases where lexicographical communication takes place via the audio channel, for example the representation of pronunciation data relating to the target language by means of audio(-visual) signals in an e-dictionary.

The above *l*-grammar components can be illustrated by applying them to  $da_1$  (repeated below):

$da_1$     **flush**<sup>3</sup> ■ n. (in poker or brag) a hand of cards all of the same suit.

An *l-syntax* would identify and describe the order of the various text constituents in  $da_1$ , for example in the hierarchical structure in figure 3:





**Figure 3:** A constituent *l*-syntax of  $da_1$

(Key: DA = dictionary article; CF = comment: form; CC = comment: concept; CLs = comment: lemma sign; i.LS = item: lemma sign; CPar = comment: paradigmatic properties; CPoS = comment: part of speech; sCPoS = sub-comment: part of speech; pre-i.PoS = pre-item: part of speech; i.PoS = item: part of speech; CSense = comment: sense; CPrag = comment: pragmatic value; CUVar = comment: usage variation; CStyle = comment: style; i.Reg = item: register; CSem = comment: semantic value; i.Def = item: *l*-definition)<sup>2</sup>

The following set of *l*-syntactic rules could be derived: DA → CF CC; CF → CLs; CLs → i.LS; CC → CPar; CPar → CPoS; CPoS → sCPoS CSense; sCPoS → pre-i.PoS i.PoS; CSense → CPrag CSem; CPrag → CUVar; CUVar → CStyle; CStyle → i.Reg; CSem → i.Def

An *l*-morphology would describe the formation of the *l*-items involved, e.g. the lemma sign | **flush**<sup>3</sup> | consists of the lemma sign form | **flush** |, printed in roman and bold, and a suffix |<sup>3</sup> | in superscript; the pre-item to the part-of-speech item is a dark square | ■ |; the part-of-speech item | n. | is an abbreviation and printed in roman; the register item | (in poker or brag) | is a PP, circumfix by parentheses and printed in roman; the lexicographic definition | a hand of cards all of the same suit | is a NP and printed in roman. With regard to the part-of-speech item | n. |, there is an overlap between the morphology of the *l*-grammar and the morphology of the target language's grammar, and with regard to the lexicographic definition | a hand of cards all of the same suit |, there is an overlap between the morphology of the *l*-grammar and the syntax of the target language's grammar. These overlaps accentuate the hybrid nature of the *l*-language.

The lexicographic text theory regards typographical features like parentheses as non-typographical structural markers, and bold print and italic print as typographical structural markers, all of which are elements of a set of non-functional text elements (cf. Wiegand 1990). The lexicographical communication theory, however, regards these features as *l*-morphemes and therefore as inherent component structures of *l*-items.

An *l*-semantics would describe the semantic value of each *l*-item as a union of form and *l*-proposition(s), for example in the table below:

**Table 5:** *L*-items and *l*-propositions in da<sub>1</sub>

<i>L</i> -items	<i>L</i> -propositions
<b>flush</b>	<i>lp</i> <sub>1</sub> : This is the dictionary article about the word <i>flush</i> . <i>lp</i> <sub>2</sub> : The word <i>flush</i> is a word in SA English. <i>lp</i> <sub>3</sub> : The word <i>flush</i> has the orthographic form (f, l, u, s, h).
<sup>3</sup>	<i>lp</i> <sub>4</sub> : The word <i>flush</i> is a member of a homonym paradigm.
n.	<i>lp</i> <sub>5</sub> : The word <i>flush</i> is a noun.
(in poker or brag)	<i>lp</i> <sub>6</sub> : (As a noun) the word <i>flush</i> is a word in the register of poker or brag.
a hand of cards all of the same suit	<i>lp</i> <sub>7</sub> : (As a noun) the word <i>flush</i> has the semantic value 'a hand of cards all of the same suit'.

An *l*-pragmatics would describe, among other things, the illocutionary force that accompanies every *l*-proposition to form the *l*-message encoded in the *l*-utterance. In terms of da<sub>1</sub>, the illocutionary force STATEMENT would for example accompany *l*-propositions *lp*<sub>1</sub> to *lp*<sub>5</sub> and *lp*<sub>7</sub> in table 5, and the illocutionary force ADVICE could accompany *l*-proposition *lp*<sub>6</sub>, depending on the dictionary's purposes and target user sociology.

The *l*-semantic information in table 5, coupled with the relevant *l*-pragmatic variables (specifically speech acts), explain how text<sub>2</sub> above communicates the same messages than text<sub>1</sub>, but by means of a sign system that is distinct from English, namely an *l*-language.

#### 2.3.4 The lemma (sign) as sign

From table 5 in the previous section it is clear that the lemma sign form | **flush** |, as it functions in da<sub>1</sub>, is not a linguistic sign like in sentence s<sub>1</sub> (cf. 2.3.2), because in da<sub>1</sub> it does not display the paradigmatic and syntagmatic properties required to function in the English grammar. Whereas the lemma *flush* functions as a linguistic sign in sentence (5) in section 2.2.1, it functions as an *l*-sign

in the *l*-language of  $da_1$ , representing a complete, multi-propositional *l*-utterance, as *l*-propositions  $lp_1$  to  $lp_3$  in table 5 demonstrate.

Furthermore, the *l*-status (as opposed to the linguistic status) of the lemma sign form |flush| can be illustrated by contrasting its salient paradigmatic and syntagmatic properties to those of the lemma as linguistic sign, as in table 6 below:

**Table 6:** Paradigmatic and syntagmatic properties of the lemma *flush* as linguistic sign and as *l*-sign

	Lemma <i>flush</i> as linguistic sign in (5)	Lemma <i>flush</i> as <i>l</i> -sign in $da_1$
<b>Paradigmatic properties</b>	— Can be replaced by any countable noun	— Can be replaced by any lemma sign form
<b>Syntagmatic properties</b>	<ul style="list-style-type: none"> <li>— Forms the compulsory head of a NP</li> <li>— Functions as stem of inflected forms</li> <li>— Can be inflected by the plural-forming suffix <i>-es</i></li> <li>— Can take AP, NP, NUM, etc. as pre-modifiers</li> <li>— Can take ADV, PP, S, etc. as post-modifiers</li> </ul>	<ul style="list-style-type: none"> <li>— Forms the compulsory head of a CF</li> <li>— Functions as stem of i.LS</li> <li>— Takes the superfix<sup>3</sup> [<b>&lt;b&gt; ... &lt;/b&gt;</b>]</li> <li>— Can take the suffix [<sup>&lt;sup&gt;</sup><math>x_i</math><sub>&lt;/sup&gt;] to indicate that it is an element (number <math>x_i</math>) of a homonym paradigm</sub></li> </ul>

Consider the variation of  $da_1$  in  $da_2$  below:

$da_2$  \*<sup>3</sup> n. (in poker or brag) *flush* a hand of cards all of the same suit.

Dictionary article  $da_2$  is preceded by an asterisk in the linguistic tradition of marking an ungrammatical construction, in this case an *l-ungrammatical* variation of  $da_1$  because the lemma sign form does not conform to its *l*-syntactic and *l*-morphological properties within *l*-grammar $_{da_1}$ , which can be expressed in the following rules:

*l*-syntax $_{da_1}$ : DA → CF CC; CF → CLs; CLs → i.Ls

*l*-morphology $_{da_1}$ : [ $x$ ]<sub>i.Ls</sub> = [**<b> $x$ </b>**]<sub>i.Ls</sub>; [ $x$ ]<sub>i.Ls[+HOM, 3]</sub> → [<sup><sup></sup> $x$ <sub></sup>]<sub>i.Ls</sub></sub>

(Key: <b> $x$ </b> = superfix: print  $x$  in bold; <sup> $x$ </sup> = superfix: print  $x$  in superscript. Compare Booij (2012: 119) for an interpretation of the morphological rule.)

The foregoing illustrates that, at least in principle, a lemma can function as both linguistic sign and *l*-sign. It functions as linguistic sign in a natural language sentence, and as *l*-sign in a dictionary article. Obviously, its primary function is

that of an *l*-sign. Therefore, again, any requirement that a lemma should be a linguistic sign in order to function in an *l*-grammar cannot be valid. This distinction would of course not affect the basic general norm that in order for a lemma to be considered for inclusion in the lemma list of a dictionary, such lemma (as an *l*-sign) should represent a linguistic sign in the treated lexicon.

## 2.4 Perspective

The discussion in the foregoing subsections (especially 2.2) demonstrate that Bergenholtz and Agerbo (2014) seemingly confuse aspects of semiotic theory with aspects of linguistic theory by attempting to disprove the existence of the linguistic phenomena of polysemy and homonymy through arguments of semiotics relating to the concept of the sign. The apparent confusion results in a misapplication of the Saussurean model of the linguistic sign, which invalidates their lexicographic theory of the lemma as linguistic sign. Furthermore, it is shown that the theory of the lemma as linguistic sign is irrelevant, because the lemma does not function as linguistic sign in lexicographical communication. Consequently, the first premise for the model II solution fails.

The validity of the second premise is the focus of the next section.

## 3. Criticism and model II implementation

In this section the criticism on existing dictionary articles by Bergenholtz and Gouws (2017) is examined. The model will also be implemented hypothetically with regard to one actual dictionary article series in the *Oxford South African Concise Dictionary* in order to identify and evaluate salient implications.

### 3.1 Criticism on existing dictionary articles dealing with homonymy and polysemy

Bergenholtz and Gouws (2017) offer a comparative criticism of the treatment of polysemy in three Danish and six English dictionaries to motivate the model II proposal. The criticism can be summarised in the following points:

- crit<sub>1</sub> The numbering of polysemic values are sometimes done in a non-transparent way and therefore polysemic values are distinguished unsystematically.
- crit<sub>2</sub> Just as many "meaning gaps" can be detected in the dictionaries as lemma gaps.
- crit<sub>3</sub> Different dictionaries that have the same lemma have different (numbers of) polysemic values for that lemma.

- crit<sub>4</sub> The same polysemic values in different dictionaries are ordered differently.
- crit<sub>5</sub> It is often unclear how polysemic values are distinguished in the same and in different dictionaries.

The general conclusion is that there is often greater consistency in lemma selection but a "lack of consistency in polyseme selection" among the dictionaries (Bergenholtz and Gouws 2017: 124). The criticism acknowledges that different dictionaries have different purposes and serve different user sociologies, which would account for some discrepancies, but not for all.

With regard to homonymy, it is argued that the distinction of homonyms does not serve the user sociology of a dictionary with only a text reception function (Bergenholtz and Gouws 2017: 125).

In the following subsection an existing series of dictionary articles will be adapted to show how the implementation of the model II solution would impact presentation and lexicographical communication. This will be followed by combined comments in subsection 3.3 on both the hypothetical model II implementation and the above criticism.

### 3.2 Hypothetical implementation of the model II solution

Dictionary article series *das*<sub>1</sub> below, extracted from the *Oxford South African Concise Dictionary* (Van Niekerk and Wolvaardt 2010: 449), will be adapted to the model II solution and presented as dictionary article series *das*<sub>2</sub>.

*Oxford South African Concise Dictionary* article series *das*<sub>1</sub> = ⟨[flush<sup>1</sup>]<sub>da</sub> ... [flush<sup>4</sup>]<sub>da</sub>⟩:

- das*<sub>1</sub> **flush**<sup>1</sup> ■ v. 1 (of a person's skin or face) become red and hot, typically through illness or emotion. 2 cleanse (something, especially a toilet) by passing large quantities of water through it. ► remove or dispose in such a way. 3 drive (a bird or animal, especially a game bird) from cover. 4 (of a plant) send out fresh shoots. ■ n. 1 a reddening of the face or skin. ► an area of warm colour or light. 2 a sudden rush of intense emotion. ► a period of freshness and vigour: *the first flush of youth*. 3 an act of flushing. 4 a fresh growth of leaves, flowers or fruit.
- DERIVATIVES **flusher** n.
- flush**<sup>2</sup> ■ adj. 1 completely level or even with another surface. 2 *informal* having plenty of money. ■ v. fill in (a joint) level with a surface.
- DERIVATIVES **flushness** n.
- flush**<sup>3</sup> ■ n. (in poker or brag) a hand of cards all of the same suit.
- flush**<sup>4</sup> ■ n. *Ecology* a piece of wet ground over which water flows without being confined to a definite channel.

Model II dictionary article series  $das_2 = \langle [\text{flush}^1]_{da} \dots [\text{flushness}]_{da} \rangle$ :

- $das_2$  **flush**<sup>1</sup> v. (of a person's skin or face) become red and hot, typically through illness or emotion.  
**flush**<sup>2</sup> v. cleanse (something, especially a toilet) by passing large quantities of water through it.  
**flush**<sup>3</sup> v. remove or dispose by flushing (>**flush**<sup>2</sup>).  
**flush**<sup>4</sup> v. drive (a bird or animal, especially a game bird) from cover.  
**flush**<sup>5</sup> v. (of a plant) send out fresh shoots.  
**flush**<sup>6</sup> n. a reddening of the face or skin.  
**flush**<sup>7</sup> n. an area of warm colour or light.  
**flush**<sup>8</sup> n. a sudden rush of intense emotion.  
**flush**<sup>9</sup> n. a period of freshness and vigour: *the first flush of youth*.  
**flush**<sup>10</sup> n. (of a person's skin or face) an occurrence of becoming red and hot, typically through illness or emotion.  
**flush**<sup>11</sup> n. an act of cleansing (something, especially a toilet) by passing large quantities of water through it.  
**flush**<sup>12</sup> n. an act of removing or disposing by flushing (>**flush**<sup>2</sup>).  
**flush**<sup>13</sup> n. an act driving (a bird or animal, especially a game bird) from cover.  
**flush**<sup>14</sup> n. a fresh growth of leaves, flowers or fruit.  
**flush**<sup>15</sup> adj. completely level or even with another surface.  
**flush**<sup>16</sup> adj. *informal* having plenty of money.  
**flush**<sup>17</sup> v. fill in (a joint) level with a surface.  
**flush**<sup>18</sup> n. (in poker or brag) a hand of cards all of the same suit.  
**flush**<sup>19</sup> n. *Ecology* a piece of wet ground over which water flows without being confined to a definite channel.  
**flusher**<sup>1</sup> n. *informal* someone who easily becomes red in the face through emotion.  
**flusher**<sup>2</sup> n. someone who drives a bird or animal (especially a game bird) from cover.  
**flusher**<sup>3</sup> n. something that is used to drive a bird or animal (especially a game bird) from cover.  
**flushness** n. the state of being completely level or even with another surface.

### 3.3 Comments on Bergenholtz and Gouws's (2017) criticism and the model II implementation

Comments are presented in numbered paragraphs.

3.3.1. A total of 16 senses (including the subsenses introduced by | ► |) are presented in four dictionary articles in  $das_1$ . (Bergenholtz and Gouws (2017) treat subsenses as separate polysemic values.) The number of dictionary articles

have increased to 23 in  $das_2$ , representing an increase of 575%. This seems to contradict Bergenholtz and Gouws's (2017: 128) estimations that the number of dictionary articles would rise, "but not too much". It should be noted that the estimations are based on calculations involving the number of dictionary articles and polysemic values they represent in samples of the studied dictionaries (cf. Bergenholtz and Gouws 2017: 126-128). Therefore, it could be argued that either  $das_2$  represents a statistical exception, or that the samples are not representative of the populations involved. Nevertheless, if the variables used in the calculations are applied in adapting  $das_1$  to  $das_2$ , then no more than 16 dictionary articles should have resulted: one dictionary article for every sense in  $das_1$ . How, then, can the substantial surplus of seven dictionary articles (44%) be explained? To begin with, cognisance should be taken of the fact that the dictionary's target user group are mother tongue speakers of English. Firstly, derivatives are not lemmatised in  $das_1$ ; rather, they are listed as such without further treatment at the end of the articles representing their stems (cf.  $[\text{flush}^1]_{da}$  and  $[\text{flush}^2]_{da}$ ). This presentation is sufficient for target users engaged in text reception tasks. In  $das_2$  every derivative has to be lemmatised and treated in a separate article with regard to every relevant polysemic value of its stem. This accounts for the last four dictionary articles in  $das_2$ . Secondly, the remaining three surplus dictionary articles, i.e.  $[\text{flush}^{11}]_{da}$  to  $[\text{flush}^{13}]_{da}$ , are the result of the necessary deconstruction of the lexicographic definition |remove or dispose in such a way| of the subsense of polysemic value 2 in the dictionary article  $[\text{flush}^1]_{da}$  ( $das_1$ ). The reference of the phrase "in such a way" and textual cohesion is lost when each polysemic value is presented in a separate dictionary article, which necessitates the addition of an article and full lexicographic definition for every polysemic value which may be a referent of "such a way". The extent to which the loss of these two lexicographic strategies may cause an increase in dictionary articles are not accounted for by Bergenholtz and Gouws (2017), and they are possibly not the only potential causes, subject to the type of dictionary involved. This implies that the offered estimates of expected increases are not reliable.

3.3.2. In relation to the previous point, there are at least two ways of dealing with lexicographic definitions in  $das_2$  that might have been briefer in articles of polysemic lemmata thanks to the relatively easy establishment of textual cohesion, like in  $[\text{flush}^1]_{da}$  ( $das_1$ ). The first method is to employ cross-references, like in  $[\text{flush}^3]_{da}$  and  $[\text{flush}^{12}]_{da}$  ( $das_2$ ). This would require the numbering of lemma signs, for example as it is done in  $das_2$ , in order to disambiguate reference addresses. The clear disadvantage of this method is that the target user would not obtain instant access to all data relating to the lemma. The second method is to write full definitions, like in  $[\text{flusher}^1]_{da}$  to  $[\text{flusher}^3]_{da}$ . With regard to  $[\text{flusher}^2]_{da}$  and  $[\text{flusher}^3]_{da}$  the question might arise as to whether instead only one lemma sign could be listed with a lexicographic definition like |someone or something that drives a bird or animal (...) from cover| in order to avoid redundancy in the lexicographic definitions of two articles. The semiotic argu-

ment advanced by Bergenholtz and Agerbo (2014) would certainly oppose such a confluence, because clearly the linguistic sign represented by the lemma sign |**flusher**<sup>2</sup>| relates to a different signified (i.e. a person) than that represented by the lemma sign |**flusher**<sup>3</sup>| (i.e. something), requiring two linguistic signs which should each be represented by a separate lemma. Also compare the treatment of subsenses in the criticism, mentioned in paragraph 3.3.1. In this regard, Lyons (1977: 554) points out and demonstrates that "distinctions of sense [and therefore of separate linguistic signs and hence lemmata] can be multiplied indefinitely" and also result in "considerable redundancy in the dictionary", apparently contradicting the "not too much"-estimate in 3.3.1. If, on the other hand, the distinction between signifieds is regarded as not significant enough to warrant two dictionary articles and the semiotic requirement is consequently somewhat relaxed, the question soon arises as to when such types of distinction are to be regarded as significant, and when not. Different editorial teams would likely draw different conclusions, and the result would be that it is not always clear how different lemmata/articles are distinguished in the same and in different dictionaries. This state of affairs would attract the same type of criticism that is expressed in crit<sub>5</sub>, the only difference being that it would relate to a different lexicographic text structure. Once the semiotic requirement is relaxed, it is not a great cognitive step to ultimately reach a point where it is argued that all different senses of a lexeme could be grouped together in one article with a single lemma sign as guiding element, like in [**flush**<sup>1</sup>]<sub>da</sub> (das<sub>1</sub>).

3.3.3. In relation to the previous point, it is not axiomatic that the model II solution would offer easier access to sought data, and no proof to the contrary is provided by either Bergenholtz and Agerbo (2014) or Bergenholtz and Gouws (2017). Instead of having to navigate through a series of dictionary articles in order to find the (precise) relevant sense of a lexeme, it could very well be argued that the target user would find it more convenient to have to look up only one lemma sign and find all senses of the represented lexeme(s) in a single consolidated text. Access to data in single, multi-sense dictionary articles could be enhanced with a clearly differentiating *l*-morphology and smart microarchitectural design without having to resort to the model II solution. With regard to the favouring of model III by Bergenholtz and Agerbo (2014) on the grounds of user familiarity, Bergenholtz and Gouws (2017: 110) are doubtful: "Whether such an approach is convincing remains to be seen." Given the foregoing, the same can be said of the model II proposal.

3.3.4. As alluded to in paragraph 3.3.2, the implementation of the model II solution across dictionaries would not guarantee more uniform decision-making by different editorial teams or even members of the same editorial team than if model I were maintained. Therefore, much of Bergenholtz and Gouws's (2017) criticism of the treatment of polysemy in existing dictionaries would apply in equal measure to model II dictionaries, the only distinction



being that it would target different text structures: (i) It is clear that the dictionary articles in  $das_2$  are not ordered systematically. Which criteria of article ordering should be applied, and how would they differ from the criteria employed to order polysemic values in dictionary articles? If different dictionaries order polysemic values differently ( $<crit_4$ ), they will most likely also order articles differently in model II. (ii) Similarly, if different dictionaries display different (numbers of) polysemic values in articles of the same lemma ( $<crit_3$ ), they will most likely display different (numbers of) articles with identical lemma signs in model II. (iii) Similarly, "meaning gaps" in model I dictionaries ( $<crit_2$ ) will be manifested as article gaps in model II dictionaries. (iv) Only the strictest instance of the model II solution would fully address  $crit_1$ , and that would result in a presently unpredictable inflation of dictionary articles (cf. 3.3.2). Therefore, it is highly unlikely that the model II solution could be implemented without eventually some relaxation of the semiotic requirement. The risk of non-transparent and unsystematic distinctions between articles would be directly proportional to the extent to which the semiotic requirement would be relaxed, and it would be even greater across dictionaries.

3.3.5. Bergenholtz and Gouws's (2017: 125) argument that the distinction of homonyms does not serve the user sociology of a dictionary with only a text reception function is clearly valid. The model II solution successfully accommodates this issue.

### 3.4 Perspective

In this section it was shown that Bergenholtz and Gouws's (2017) criticism of the treatment of polysemy in existing model I dictionaries is hardly addressed by the model II solution, although it deals successfully with the question of homonymy. There are also potential quantitative consequences of the implementation of model II that have not been accounted for. Furthermore, it is highly unlikely that model II could be implemented without some eventual relaxation of the semiotic requirement, which would similarly have potential consequences that have not been considered and may be difficult to estimate. These undescribed and unidentified variables would be costly to the integrity of the model II theory, if it was otherwise in order. The conclusion is that the final premise for the model II solution is questionable at best.

In the following section the potential for an alternative to the model II solution is outlined. It is based on the practical treatment of homonymy and polysemy in Van Dale dictionaries.

## 4. A potential alternative to model II: *l*-polysemy and *l*-homonymy

Instead of arguing for the disposal of polysemy and homonymy in lexicography, the concepts could be adapted to lexicography so that they are not

limited to linguistic interpretation. This calls for the introduction of *l-polysemy* and *l-homonymy*. All senses that are allocated to one dictionary article and whose treatments are addressed at one lemma sign constitute *l-polysemy*, regardless of whether such senses represent linguistic polysemy. Similarly, when more than one formally identical lemma sign form, each with its separate dictionary article, is presented, those lemma sign forms are *l-homonyms* and constitute an instance of *l-homonymy*, regardless of whether they represent linguistic homonymy. Whereas linguistic polysemy and homonymy pertain to lexemes, *l-polysemy* and *l-homonymy* pertain to lemma sign forms. Lemma signs |flush<sup>1</sup>| to |flush<sup>19</sup>| in das<sub>2</sub> above (cf. 3.2), for example, constitute a paradigm of *l-homonyms*.

The application of *l-polysemy* and *l-homonymy* can be briefly illustrated by means of a set of articles from *Van Dale Online Gratis Woordenboek*<sup>6</sup>. In the interest of brevity, details and requirements of user sociology and dictionary purposes will not be accommodated here; the objective is to demonstrate the potential of the concepts and not to fully develop an alternative model to model II.

Consider the following dictionary article series, das<sub>3</sub>:

- |   |
|---|
| <p><sup>1</sup>as (<i>de; v(m)</i>; meervoud: <i>assen</i>) zie x-as, y-as</p> <ol style="list-style-type: none"><li>1. voorwerp waarom of waarmee iets ronddraait; = spil</li><li>2. denkbeeldige lijn door het middel van een voorwerp, ruimte of vlak: <i>de as van de aarde</i>; <i>de as Berlijn-Rome</i> het bondgenootschap tussen Duitsland en Italië van 1936 tot 1943</li><li>3. lijn die een lichaam in twee symmetrische helften verdeelt</li></ol> <p><sup>2</sup>as (<i>de; v(m)</i>; meervoud: <i>assen</i>)</p> <ol style="list-style-type: none"><li>1. overblijfsel bij verbranding: <i>een huis in de as leggen</i> verbranden</li></ol> |
|---|

**Dictionary article series das<sub>3</sub>** = ⟨[<sup>1</sup>as]<sub>da</sub>, [<sup>2</sup>as]<sub>da</sub>⟩ from *Van Dale Online Gratis Woordenboek* NL-NL

In das<sub>3</sub>, two linguistic homonyms are distinguished and presented as separate lemma signs, i.e. |<sup>1</sup>as| and |<sup>2</sup>as|. The first lemma is allocated three polysemic values, all relating to the semantic value 'axis'. The second lemma represents a monosemic lexeme with a lexicographic definition and cotext item signalling the semantic value 'ash'. In das<sub>3</sub> Van Dale applies a linguistic distinction between homonyms, i.e. two lexemes with identical form but unrelated semantic values. Here, *l-homonymy* corresponds to linguistic homonymy, and *l-polysemy* corresponds to linguistic polysemy. This is a typical application of model I.

In contrast, compare [as]<sub>da</sub> below:

- |  |
|--|
| <p><b>as</b></p> <ol style="list-style-type: none"><li>1. (<i>verbrande resten</i>) ashes, ash (<i>van sigaret</i>): <i>gloeiende as</i> (glowing) embers; <i>een stad in de as leggen</i> reduce a city to ashes</li><li>2. axle, (<i>drijf</i><i>as</i>) shaft</li><li>3. (<i>meetkunde</i>) axis: <i>om zijn as draaien</i> revolve on its axis</li><li>4. (<i>muziek</i>) A-flat</li></ol> |
|--|

**Dictionary article [as]<sub>da</sub>** in *Van Dale Online Gratis Woordenboek NL-EN*

In dictionary article [as]<sub>da</sub>, four senses are distinguished: The first sense is related to the homonym represented by the lemma sign |<sup>2</sup>as| in das<sub>3</sub>, senses 2 and 3 are polysemic values related to the homonym represented by the lemma sign |<sup>1</sup>as|, and sense 4 is related to a homonym not represented in das<sub>3</sub>. In this article, obviously, homonyms are not represented by separate lemma signs. Therefore, *l*-polysemy does not correspond to linguistic polysemy, although there is some overlap. Although linguistic homonymy could be said to be involved, it is not represented (by *l*-homonymy). In linguistic terms, lemma sign |as| represents three lexemes. In semiotic terms, it represents four linguistic signs (cf. 3.2.2).

Finally, compare the following dictionary article series, das<sub>4</sub>:

- |   |
|---|
| <p><sup>1</sup><b>dwaas</b> (<i>bijvoeglijk naamwoord, bijwoord</i>; vergrotende trap: <i>dwazer</i>, overtreffende trap: <i>dwaast</i>)</p> <ol style="list-style-type: none"><li>1. zot, gek</li></ol> <p><sup>2</sup><b>dwaas</b> (<i>de; m,v</i>; meervoud: <i>dwazen</i>)</p> <ol style="list-style-type: none"><li>2. gek, dwaas mens</li></ol> |
|---|

**Dictionary article series das<sub>4</sub>** = ⟨[<sup>1</sup>dwaas]<sub>da</sub>, [<sup>2</sup>dwaas]<sub>da</sub>⟩ in *Van Dale Online Gratis Woordenboek NL-NL*

In das<sub>4</sub>, two homonyms are distinguished and presented as separate lemma signs. From the paraphrases of meaning it is clear that both lemma signs represent lexemes with very closely related semantic values: [<sup>1</sup>dwaas]<sub>da</sub> (adj., adv.) the semantic value 'foolish', and [<sup>2</sup>dwaas]<sub>da</sub> (n.) the semantic value 'fool'. Here, *l*-homonymy is distinguished on the basis of lemma signs that represent formally identical lexemes from different parts of speech. If these lexemes are considered to be grammatical homonyms (cf. Carstens 2018: 116-117), then *l*-homonymy corresponds to a form of linguistic homonymy. If, instead, they are considered to represent an instance of part-of-speech multifunctionality (cf. Gouws 1989: 126-129), then *l*-homonymy does not correspond to linguistic homonymy.

In paragraphs 3.3.2 and 3.3.3 above it was argued that target users might

prefer senses to be grouped under one lemma sign for ease of access to the relevant data on the represented lexeme(s), instead of each sense being presented in a separate dictionary article to satisfy some extra-metalexigraphic requirement. The concepts of *l*-polysemy and *l*-homonymy provide the theoretical space to address the target user sociology without the obligation to conform to unduly restrictive elements of linguistic or semiotic theory. The terms have the added advantage that their denotations can vary according to the *l*-grammar in which they are applied, as demonstrated above. This does not imply, however, that they do not need to be applied systematically and be based in lexicographic theory.

The use of *l*-homonymy and *l*-polysemy in [as]<sub>da</sub> and das<sub>4</sub> yield similar results to model III. Yet, *l*-homonymy and *l*-polysemy represent a different model because it has a different theoretical base: Model III is predicated on the notion of polysemic and homonymic signifiers as defined by Bergenholtz and Agerbo (2014: 32) (although the notion of polysemic and homonymic relations between signifieds in fact defines linguistic polysemy and homonymy; cf. Hébert 2018), while *l*-homonymy and *l*-polysemy has the construct of an *l*-grammar as foundation. In lexicographic application, the flaws of the premises underlying model II also apply to model III (cf. 2).

## 5. Conclusion

This article has identified two main theoretical premises for Bergenholtz and Agerbo's (2014) and Bergenholtz and Gouws's (2017) model II solution to the treatment of polysemy and homonymy in dictionaries that have only a text reception function. Under examination, as reported in the foregoing sections, one of the premises have been proven invalid, and the second is only partially valid, inasmuch as it addresses homonymy. Both premises fail to support the proposed solution with regard to the question of polysemy in the dictionary type involved. This leaves only one argument cited in favour of the model II solution, namely that of data accessibility. However, the argument can equally well support a counter-model II conclusion, as shown in paragraphs 3.3.2 and 3.3.3, which can be theoretically defended by employing the notions of *l*-polysemy and *l*-homonymy in an *l*-grammar. Whether the model II solution or a solution involving *l*-polysemy and *l*-homonymy is the (more) valid one from a standpoint of practice, can only be proven by (independent) experimental user research based on a robust methodology. Even then, the general conclusion might entail that different target user groups prefer different solutions to the treatment of polysemy. Still, it is highly unlikely that a "pure" model II solution would be practicable.

During the course of the exposition in this article, a potential broad conceptual framework for the lexicographical communication theory was developed. In the same way that the well-established term *lemma* is used in meta-lexicography to distinguish a guiding element of a dictionary article from the

lexical item which it represents, the lexicographical communication theory introduces the notion of *l*-grammar (including *l*-polysemy and *l*-homonymy) parallel to linguistic grammar to distinguish lexicographic theory from linguistic theory, even while the former benefits from the latter.

## Endnotes

1. Although De Saussure (2013: 77) uses the term *sound pattern*, signifiers are "now commonly interpreted as the *material (or physical) form of the sign*" (Chandler 2007: 15); cf. 2.2.2.
2. Due to space considerations the principles of this constituent *l*-syntax (and the *l*-grammar) are not elaborated here. They will be explained in future work. However, it should be noted that the terms *comment* and *item* have different denotations from the formally identical terms in the lexicographic text theory.
3. The term *superfix* is introduced to refer to an *l*-affix that is superimposed onto another form instead of prefixed, suffixed, circumfixed or suprafixed to it. It is an affix because it is a dependent *l*-morpheme and it contributes to the construction of *l*-meaning.
4. The term *meaning* is not defined in either article despite evidently not sharing the denotation De Saussure assigns to it (cf. 2.1). If it is used as a synonym for *signified/content*, the problem is even more acute.
5. Morphological simplexes can be regarded as *simple linguistic signs*, and morphological complexes and syntagmata as *complex linguistic signs* (cf. Cruse 2011: 12-13).
6. The representation of the Van Dale dictionary articles in this section do not fully correspond to the actual articles' *l*-morphology and microarchitecture.

## References

### A. Primary literature (dictionary data)

- Reynolds, M. (Ed.). 2006. *South African Oxford Secondary School Dictionary*. Cape Town: Oxford University Press Southern Africa.
- Van Dale Online Gratis Woordenboek. Accessed at: <https://www.vandale.nl/opzoeken> [10 August 2018].
- Van Niekerk, T. and J. Wolvaardt (Eds.). 2010. *Oxford South African Concise Dictionary*. Second edition. Cape Town: Oxford University Press Southern Africa.

### B. Secondary literature

- Bergenholtz, H. and H. Agerbo. 2014. There is No Need for the Terms Polysemy and Homonymy in Lexicography. *Lexikos* 24: 27-35. DOI: <http://dx.doi.org/10.5788/24-1-1251>.
- Bergenholtz, H. and R.H. Gouws. 2017. Polyseme Selection, Lemma Selection and Article Selection. *Lexikos* 27: 107-131. DOI: <http://dx.doi.org/10.5788/27-1-1396>.
- Beyer, H.L. 2014. Explaining Dysfunctional Effects of Lexicographical Communication. *Lexikos* 24: 36-74. DOI: <http://dx.doi.org/10.5788/24-1-1252>.

- Beyer, H.L. and J. Augart.** 2017. From User Questions to a Basic Microstructure: Developing a Generative Communication Theory for a Namibian German Dictionary. *Journal for Studies in Humanities and Social Sciences* 6(2): 1-31.
- Bock, Z.** 2014. Introduction to Semiotics. Bock, Z. and G. Mheta (Eds.). 2014. *Language, Society and Communication: An Introduction*: 55-77. Pretoria: Van Schaik.
- Booij, G.** 2012. *The Grammar of Words. An Introduction to Linguistic Morphology*. Third edition. Oxford: Oxford University Press.
- Carstens, W.A.M.** 1997. *Afrikaanse tekstlinguistiek: 'n inleiding*. Pretoria: J.L. van Schaik.
- Carstens, W.A.M.** 2018. *Norme vir Afrikaans. Moderne Standaardafrikaans*. Sixth edition. Pretoria: Van Schaik.
- Chandler, D.** 2007. *Semiotics. The Basics*. Second edition. New York: Routledge.
- Clark, B.** 2013. *Relevance Theory*. Cambridge: Cambridge University Press.
- Cruse, A.C.** 2011. *Meaning in Language. An Introduction to Semantics and Pragmatics*. Third edition. Oxford: Oxford University Press.
- Danesi, M.** 2004. *Messages, Signs and Meanings. A Basic Textbook in Semiotics and Communication Theory*. Toronto: Canadian Scholars' Press.
- De Beaugrande, R. and W. Dressler.** 1981. *Introduction to Text Linguistics*. London/New York: Longman.
- De Saussure, F.** 2013. *Course in General Linguistics*. Translated and annotated by Roy Harris. London/New York: Bloomsbury Academic.
- Gallmann, P.** 1991. Wort, Lexem und Lemma. Augst, G. and B. Schaeder (Eds.). 1991. *Rechtschreibwörterbücher in der Diskussion. Geschichte — Analyse — Perspektiven*. Theorie und Vermittlung der Sprache 13: 261-280. Frankfurt am Main/Bern/New York/Paris: Peter Lang.
- Gouws, R.H.** 1989. *Leksikografie*. Pretoria/Cape Town: Academica.
- Gouws, R.H., U. Heid, W. Schweickard and H.E. Wiegand (Eds.)**. 2013. *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin/New York: De Gruyter Mouton.
- Grice, P.** 1991. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hébert, L.** 2018. Elements of Semiotics. *Signo. Theoretical Semiotics on the Web*. Accessed at: <http://www.signosemio.com/elements-of-semiotics.asp> [13 August 2018].
- Lyons, J.** 1977. *Semantics. Volume 2*. Cambridge: Cambridge University Press.
- Murphy, M.L.** 2010. *Lexical Meaning*. Cambridge: Cambridge University Press.
- Peirce, C.S.** 1985. Logic as Semiotic: The Theory of Signs. Innis, E. (Ed.). 1985. *Semiotics. An Introductory Anthology*: 4-23. London/Johannesburg: Hutchinson.
- Sebeok, T.A.** 2001. *Signs: An Introduction to Semiotics*. Second edition. Toronto/Buffalo/London: University of Toronto Press.
- Sperber, D. and D. Wilson.** 1995. *Relevance, Communication and Cognition*. Second edition. Oxford: Blackwell.
- Van Dijk, T.A.** 1972. *Some Aspects of Text Grammars. A Study in Theoretical Linguistics and Poetics*. Janua Linguarum: Studia Memoriae Nicolai Van Wijk Dedicata. Series Maior 63. The Hague: Mouton.
- Wiegand, H.E.** 1990. Printed Dictionaries and Their Parts as Texts. An Overview of More Recent Research as an Introduction. *Lexicographica* 6: 1-126.

- Wiegand, H.E.** 1996. A Theory of Lexicographic Texts: An Overview. *South African Journal of Linguistics* 14(4): 134-149.
- Wiegand, H.E.** 1996a. Textual Condensation in Printed Dictionaries. A Theoretical Draft. *Lexikos* 6: 133-158. DOI: <http://dx.doi.org/10.5788/6-1-1029>.
- Wiegand, H.E. and R.H. Gouws.** 2013. Addressing and Addressing Structures in Printed Dictionaries. Gouws, R.H., U. Heid, W. Schweickard and H.E. Wiegand (Eds.). 2013: 273-314.
- Zgusta, L.** 1971. *Manual of lexicography*. The Hague: Mouton.