(Re)defining Component Structures in Morphological Constructions: A Cognitive Grammar Perspective

Gerhard B van Huyssteen

1. Introduction

The study of words, parts of words, and word-formation is one of the fields of enquiry that has kept language philosophers and linguists busy for centuries. Despite this long tradition, the literature on word morphology sometimes remains rather imprecise, ambiguous and/or vague; for example, Tuggy (1992: 287) illustrates that "(d)efinitions, when given, are frequently circular to some degree or inexplicit at crucial points", while Bauer (2004: 70) writes about the term morpheme:

> Although the morpheme is the fundamental unit of morphology, it is given **a wide range of definitions**, some of which appear to be trying to capture the same concept, others of which clearly define a different unit. At the extremes, some scholars deny the validity of the notion completely. [My emphasis - GBVH]

Three reasons could probably be offered for this situation. Firstly, different theoretical frameworks call for different interpretations and definitions of certain concepts, notions, and terms. In this regard, Booij, Lehmann & Mugdan (2000: XXIV) rightly state that "(t)he diversity of theoretical approaches in linguistics is one of the reasons why morphological terminology is far from uniform". To give one example, in his Word Grammar theory Hudson (1984) rejects the Structuralist distinction between phonology/morphology and between morphology/syntax and, therefore, sees no importance to distinguish between word and morpheme: "…a word grammar will give a general definition for 'word', and also one for 'sound segment', but will make no generalizations about 'morpheme'" (Hudson 1984: 56).

Secondly, terminology is often selected and defined according to the nature of the specific language that is being described; for example, some languages have what Bauer (2003: 30) calls transfixes, others don't. As Gleason (1955: 59-60) aptly states:

> The broadest and most comprehensive classes of morphemes in English…are *roots* and *affixes*. … A definition of these two classes which would be universally applicable would be immensely complex and is probably unnecessary here. **A definition which will fit the needs of one specific language is commonly feasible.** [My emphasis – GBVH]

This viewpoint is reiterated by Plag (2003: 11):

> …these terms are not always clearly defined in the morphological literature and are therefore a potential source of confusion. One reason for this lamentable lack of clarity is that languages differ remarkably in their morphological make-up, so that **different terminologies reflect different organizational principles in the different languages**. [My emphasis – GBVH]

The last obvious reason why terminology in morphology literature often varies, even within the same theoretical framework and for the same language, is because of the context, purpose and scope of different scholarly works. For example, while Senekal, Ponelis & De Klerk (1972: 244-247) only distinguish between stems and affixes as morphemes in Afrikaans in their introductory grammar handbook, Combrink (1990) postulates a comprehensive, complex taxonomy of various morphemes (and subcategories of morphemes) in his landmark book *Afrikaanse Morfologie: Capita Exemplaria*.[1]

Therefore, if one wants to describe aspects of morphological constructions (or other morphological phenomena), one should be explicit in one's definition of notions and terminology about the theoretical framework (if any) of the description; the language(s) to which these notions and terms would apply; and what the context, purpose and scope of these would be.

For purposes of this article, I will define some basic concepts (specifically component structures) relevant to morphological constructions within a Cognitive Grammar framework. Langacker (amongst others 1987, 1991), Taylor (2002), Tuggy (1985, 1992, 2005) and Evans & Green (2006) cover various general morphological issues in their work and offer a good foundation for further theoretical and descriptive work.[2] This article is therefore aimed at extending and elaborating on these foundational works, by offering a more detailed characterization of component structures.

With regard to the language to which these notions apply, I will limit myself here to Afrikaans (a Low Franconian, West Germanic language, closely related to Dutch but with a somewhat simpler inflectional morphology); most of the definitions should, however, also be applicable to

other (at least closely-related) languages. The reason for choosing Afrikaans is that literature on Afrikaans morphology is even more so imprecise and confusing. For example, Van Schoor (1983: 353) distinguishes between dependent and independent stems, while De Klerk (1968: 170) labels the same distinction as stems and roots. Even worse is the confusing debate on which morphosyntactic categories of inflection should be recognized for Afrikaans, if any at all (Combrink 1990); many more examples abound. In short: theoretical Afrikaans morphology is rather "messy" and a need therefore exists to clarify some of the notions applicable to Afrikaans.

In terms of context, purpose and scope, this research sprouts from a very concrete and practical need. As part of a project to develop an automatic morphological parser for Afrikaans (see Section 4), it is imperative to have a clear and distinct understanding of the various component structures in complex Afrikaans words. To mention but two examples:

− The homonymous form *gas* (1. gas; 2. guest) in Afrikaans has two distinct plural forms, one for each homonym: *gasse* 'gasses' and *gaste* 'guests'. In computational morphological analysis (where no semantic or diachronic information is necessarily available), the question is whether these two plural forms should both be related to the stem *gas*, or rather to a stem *gas* (in the case of *gass-e*) and a root *gast-* (in the case of *gast-e*). Of course, the eventual analysis will depend on the purpose of the process, but the need remains to draw an explicit distinction between these two possible component structures in morphological constructions.

− Like in some other (Germanic) languages, Afrikaans compounding often employs a so-called linking morpheme to form a new compound. In a compound like *perdekar* 'horse cart' the question is whether it should be analyzed as *[[[perd]e] kar]* (i.e. where the *-e-* is considered a plural suffix), or as *[[perd] e [kar]]* (where the *-e-* is seen as a linking morpheme). To produce an unambiguous analysis of words like *perdekar*, one needs to have a comprehensive understanding of the possible component structures in complex words.

Given these practical problems and the fact that current theoretical Afrikaans morphology does not provide definite solutions or answers, the aim of this research is to postulate a theoretically unified, unambiguous categorization network (specifically a taxonomy) for component structures in Afrikaans morphology. The focus in this article will be the theoretical (linguistic) base of the above mentioned project; as will become clear,

however, the applied aims (such as details of the project and practical implementations) have to be kept in mind for making certain decisions.

Overall, this article aims to (re)define the component structures in Afrikaans morphological constructions from a Cognitive Grammar perspective. I will not pay attention to general issues in Cognitive Grammar, such as the status of morphology, morphological operations, constructions and construction schemas, composite structures, or other common cognitive processes; various standard works in Cognitive Grammar, as well as other articles in this current volume address some of these issues in more detail. Neither is it the intention, nor within the scope of this article to revisit all that has ever been written about component structures in morphology theory; it is assumed that the reader has sufficient background on issues pertaining to different interpretations and definitions of different notions.

In the next section, I will give a brief introduction to the concept "component structure" and will show why I choose to characterize component structures in terms of valence factors. Section 3 introduces the valence factors by first giving a concise overview of what it entails, before describing Afrikaans component structures in more detail. At the end of Section 3 a summary of the main characteristics is provided. In Section 4 a categorization network (i.e. taxonomy) based on three of the valence factors is presented and discussed and it is shown how this taxonomy is applied for purposes of annotating Afrikaans data manually. Section 5 concludes with a brief summary and an outlook on future work.

2. Background: Symbolic Units in Morphological Constructions

One of the very basic assumptions of Cognitive Grammar is that grammar is symbolic in nature. The grammar of a language is characterized as a "structured inventory of conventional units" (Langacker 1987: 57), where such units are symbolic assemblies – i.e. pairings of meaning (the semantic pole) and form (the phonological pole) – on any level of schematicity. If fully specified on both the semantic and phonological poles, it constitutes something like a morpheme (e.g. [[PL]/[s]] for the plural -*s* in Afrikaans), or a word (e.g. [[TAFEL]/[tafel]] 'table'); if it is underspecified/schematic, it comprises grammatical patterns/constructions/schemas (e.g. [[[THING]-[PL]]/[[…]-[s]]] for the common plural construction in Afrikaans). As such,

morphemes, words, and grammatical constructions are all seen as symbolic structures differing "not in kind, but only in degree of specificity" (Langacker 1987: 58). Let's look at each of these symbolic structures a bit more in detail.

A *morpheme* can be identified and defined as a simplex symbolic (i.e. (grammatically) meaningful) unit in the language system (Langacker 1987: 58/345, 1991: 105). It is simplex in the sense that it does not contain smaller symbolic units as subparts (Evans 2007: 197); smaller units of linguistic analysis (i.e. sounds like [s], [r] or [g]) normally don't have meanings associated with them, they don't "mean" anything. However, a morpheme, like the plural morpheme [[PL]/[s]], specifies through convention plurality, 'more than one'-ness on the semantic pole; hence the morpheme is a symbolic unit. Of course, this definition also implies that simplex words like *tafel* 'table', *groot* 'big' or *speel* 'play' are considered morphemes.

This implies that all morphemes constitute conventionalized form/meaning pairings, which might be schematic on various levels. Consider for instance a so called grammatical morpheme such as the partitive genitive morpheme in *iets groot-s* 'something big-PRTT.GEN'; this morpheme is highly specified on the phonological pole (the construction schema would specify in which phonological contexts, etc.), but is rather vague on the semantic pole in terms of its conceptual import and specification. This also applies to cran morphemes, like *aal-* in *aal-bessie* 'currant', or *boe-* in *boe-pens* 'potbelly'. On the other hand, zero morphemes such as the zero realization of the plural in *drie jaar-Ø* 'three year-PL' are specified on the semantic pole but highly schematic (i.e. to the extent where it's empty) on the phonological pole. Thus, all grammatical morphemes are also meaningful, "and function as they do in larger syntactic constructions because of their meanings" (Tuggy 2005: 235).

Note that no theoretical importance is given to the term morpheme – it is merely a label for the smallest symbolic unit in grammar. As such, the need to make a distinction between morphemes and morphs disappears, since morphemes are not seen as some abstract, theoretical construct that needs to be realized as morphs on some surface structure.

A *word* is defined as a simplex or complex symbolic unit in the language system, larger than a morpheme and smaller than a phrase (Evans & Green 2006: 485) and consists of a (relatively) stable, integral and promiscuous phonological structure associated with a (relatively) stable semantic structure (Taylor 2002: 173-175). Words can be simplex symbolic

structures, just like morphemes, or complex in that they could contain smaller symbolic assemblies as subparts (Evans 2007: 28). While a word like [[TAFEL]/[tafel]] 'table' can't be subdivided into smaller constituent and is thus a simplex symbolic unit, words like [[[TAFEL]/[tafel]] [[PL]/[s]]] 'tables', or [[[TAFEL]/[tafel]] [[POOT]/[poot]]] 'table leg' each consist of two simplex symbolic units and are thus complex.

A prototypical word like *brug* 'bridge' is phonologically stable in the sense that it retains its meaning BRUG whether it is pronounced as [brœx] or [brəx]; it has phonological integrity in that "the content cannot be broken up by intervening material, nor can the different parts appear in a different sequence" (Taylor 2002: 174); and it is promiscuous to the extent that it can co-occur in sequence (whether in compounds or in phrases) with a large variety of other words (in contrast to, for example, the plural suffix that can only combine on the right-hand side of count nouns). The word *brug* also means roughly the same in different contexts (either referring to a card game or to a building construction) and we can thus say that its meaning is stable and coherent.

Of course, as is generally accepted in Cognitive Grammar with regard to category membership, this definition sets scalar rather than absolute parameters – it aims to define parameters for prototypical members of the class WORD, while expecting variance on the scale of prototypicality. For example, if one would see the parameter of phonological integrity as absolute (and not as gradual), one would have to conclude that the compound *brugspeler* 'bridge player' is not a word since it allows for phenomena such as swear-word incorporation (e.g. *brug-fokken-speler* 'bridge-fucking-player'). The observation that *brugspeler* is in terms of these parameters a lesser prototypical word than *brug,* is unproblematic since it remains a word – just like an ostrich remains a bird, even though it is less prototypical than a sparrow. This characterization of the notion word is therefore an attempt to approximate the prototype, rather than to set absolute criteria.

In the context of Cognitive Grammar, the necessity to theoretically distinguish between notions such as lexeme, word-form, and grammatical/morphosyntactic word wanes since all symbolic units, irrespective of their complexity or specificity, are considered part of the structured inventory of conventionalized units (i.e. the grammar) and hence sanctioned by various schemas.

This brings us to *morphological constructions*, which are simply defined as composite symbolic structures (i.e. complexes) smaller than phrases,[3]

consisting of component structures (that could be simplexes or complexes) between which valence relations hold (Evans & Green 2006). Once again, since constructions are per definition symbolic, they can be specified on various levels of schematicity, allowing for high-level schemas (i.e. constructional schemas) and for very low, specific instantiations (e.g. as in the case of cran morpheme constructions, or exceptions such as unmotivated plural variants – compare for instance the two plural forms *kokk-e* 'cooks' and *kok-s* 'cooks' of the noun *kok* 'cook', where the first plural form is sanctioned by a construction subschema, but the second is a low-level instantiation/exception; cf. Langacker 1987: 409-410).

In this article, I am not focusing on composite structures per se, but rather on the component structures on a morphological level of composition, where a component structure is defined straightforwardly as a "structure that integrates with one or more other structures in a combinatory relationship (particularly a grammatical valence relation)" (Langacker 1987: 487). In the case of morphological constructions, these component structures are prototypically morphemes and/or words (as defined above), but could even include phrases in less prototypical instances (e.g. *traak-my-nie-agtige houding* 'couldn't-care-less attitude')*.

For purposes of this article, I will focus on those "basic" components that are generally defined in standard, introductory works on morphology (e.g. Aronoff & Fudeman 2005, Bauer 2003, Booij 2007, Haspelmath 2002, Katamba 1993, Matthews 1991). These include stems (or roots), a variety of affixes, zero morphemes, cran morphemes, and, specifically related to Afrikaans, linking morphemes.[4] Some of these distinctions are also often made in the Afrikaans literature, to greater or lesser degrees of success. If it is the aim to give a theoretical (re)definition of morphological component structures, as they apply to Afrikaans, the theory would have to be able to account for at least these concepts in some or other way.

Since valence plays such a central role in grammatical constructions, it makes sense to characterize component structures in terms of factors that determine the valence potential of two component structures to enter in a valence relation. This will be discussed in the next section.

3. Characterization of Component Structures

Taylor (2002: 229) defines valence simply as "the combinatorial possibilities of the unit". Based on a metaphor from chemistry, valence

implies that component structures (i.e. the "atoms") should be "understood
and explained with reference to their internal structure" (Langacker 1987:
277). As we have seen, the internal structure of component structures
consists of a paired phonological and semantic pole and our description of
the internal structure of component structures should thus pay attention to
both these levels of description.

    A variety of parameters can be identified in terms of which internal
structure of component structures should be explained. These include the
levels of specificity of component structures, the degree to which one
structure presupposes another structure (i.e. the autonomy and dependence
of structures), the ease with which one structure combines with a variety of
other structures (i.e. promiscuity), and the sequence in which structures
combine with each other (i.e. constituency) (cf. Langacker 1987, Taylor
2002, Tuggy 2005 for lengthy discussions of these valence factors).[5] Note
that all these parameters should be seen as scalar and not as clear-cut,
either-or distinctions (Taylor 2002: 324); as was the case above with the
notion word, we explicitly expect different degrees of prototypicality, as I
shall illustrate below. In this section, these valence factors will be used as
parameters in terms of which component structures in morphological
constructions can be characterized.

3.1 Schematicity

Schematicity pertains to the level of specificity on both the semantic and
phonological poles; Langacker (1991: 62) refers to the "grain" or
"resolution" of a structure. Component structures that are highly specified
on the semantic pole have a higher resolution (or finer grain) and can be
considered semantically "heavy" or "contentful". Conversely,
underspecified structures can be considered semantically "light" or
"empty". As the same applies to the phonological pole, we could
distinguish various combinations on these two scales. Taylor (2002: 327)
represents this situation in a two dimensional diagram, similar to the
diagram for prototypical Afrikaans component structures in Figure 1. Note
that the labels in the diagram represent relative positions (i.e. relative to
other labels) of prototypical instances; no absolute, or empirical claims are
therefore postulated by this diagram (or other similar diagrams in the
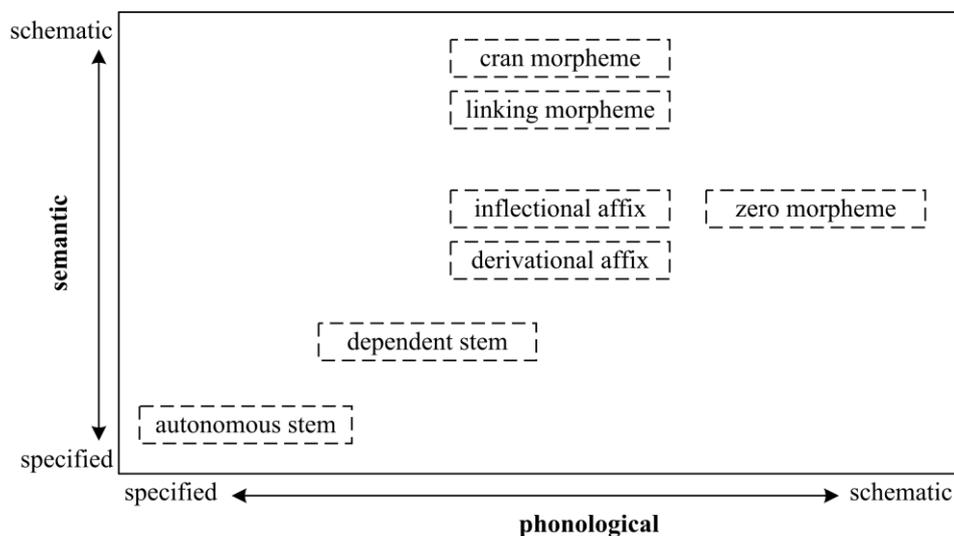remainder of this article).

Figure 1. Two-dimensional representation of Afrikaans component structures in terms of schematicity

All complex words have one thing in common: they contain at least one stem, whether this stem is autonomous or dependent. The reason for this is that stems are generally considered "the central meaningful element of the word" (Plag 2003: 10), that "the root generally carries the main component of meaning in a word" (Crystal 1997: 336).[6] Since autonomous stems are per definition words, we could expect their semantic structure to be relatively stable and to a great extent specified; of course, the semantic structures of content words (like *tafel* 'table' or *groot* 'big') are more specified than those of function words (like *en* 'and' or *op* 'on') and, thus, also more "contentful". Both function words and simplex content words (i.e. autonomous stems) are fully specified on the phonological pole and can therefore occur autonomously.

In the category of autonomous stems, we can distinguish between so-called full stems and reduced stems. While full stems represent the prototype (as explained in the previous paragraph), reduced stems are reduced forms of autonomous words or phrases, which came into existence through various reduction processes. For example, acronyms and letter-names are created through a process of reduction, resulting in autonomous stems that are available for morphological constructions (e.g. for

compounding, in a word like *CNN-nuus* 'CNN news', or for derivation in a word like *ANC-agtig* 'ANC-like'). Other examples include stems that were created through clipping (e.g. *admin < administrasie* 'administration', or *avocado < avokadopeer* 'avocado pear'), back-formation (e.g. *boskaas < boskasie* 'unkempt hair', or *dramaat < dramaties* 'dramatic'), and abbreviation (e.g. *km < kilometer* 'kilometer', or *ww. < werkwoord* 'verb'). Although these stems are in a sense phonologically less specified than their full-form counterparts, they are still fully specified symbolic units, available for morphological constructions.

Dependent stems are less specified than autonomous stems, but contain still more semantic content (i.e. they are more contentful) than most affixes. Compare for instance a, b, and c in Figure 2, where schematic content in the quasipictorial diagrams is indicated by cross-hatching. The autonomous stem *tafel* 'table' in Figure 2a is specified on both the semantic and phonological poles; the dependent stem *gast-* 'guest' in Figure 2b is specified on the semantic pole, with some schematicity on the phonological pole, while the plural suffix *-e* in Figure 2c is schematic on both the phonological and semantic poles (but still with some content).

Within the category of dependent stems, we also find various degrees of specificity. For example, so-called variant stems (e.g. *gast-* in *gast-e* 'guests', *leg-* in *leg-kaart* 'puzzle', or *aard-* in *aard-bol* 'globe'; Combrink 1990: 25) are more specified than non-variant stems (like *elektr-* in *elektr-isiteit* 'electricity', *kwot-* in *kwot-eer* 'quote' or *kwot-asie* 'quotation', or *sekret-* in *sekret-aris* 'secretary' or *sekret-ariaat* 'secretariat').[7] Yet another category would be so-called neo-classical stems or combining forms, such as *Anglo-* in *Anglo-Amerik-aans-e* 'Anglo-American', or *neuro-* in *neuro-linguistiek* 'neurolinguistics'. Variant stems are mostly from Germanic origin (even more specific, Dutch relics), they have all close word-correlates (e.g. *gast-/gas* 'guest', *leg-/lê* 'lay', or *aard-/aarde* 'earth'), and can therefore be specified much more easily and more precisely on both the phonological and semantic poles. Non-variant and neo-classical stems, on the other hand, don't have word-correlates, and their meanings seem to be understood or paraphrased most of the time only in terms of the paradigms in which they occur. Variant stems are in this sense much more like autonomous stems, while non-variant and neo-classical stems are more affix-like.
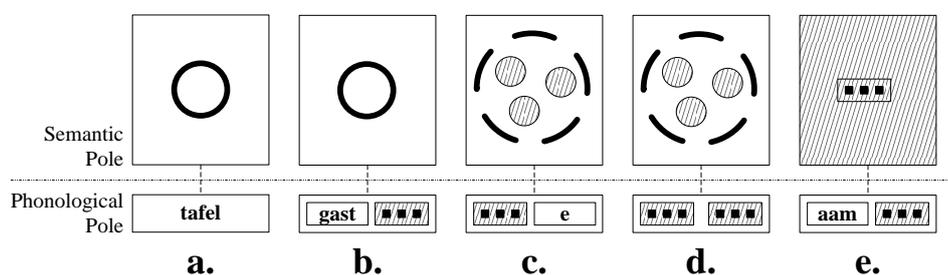
Figure 2. Quasipictorial representations of some component structures

Both derivational and inflectional affixes (and per implication also zero morphemes) fall somewhere in the middle of the continuum of schematicity. Consider, for example, the adjectivalizer *-agtig* as in *hond-agtig* 'dog-ADJR, dog-like', which is a derivational affix with a meaning that can be quite accurately defined as 'like X' (where X is the stem it attaches to). The meaning of the nominalizer *-e* as in *vertrou-e* 'trust-NR, trust', *bydra-e* 'contribute-NR, contribution', and *weet-e* 'know-NR, knowledge' is a bit more evasive and difficult to define (probably 'THING of X'), while the infinitive *-e* (as in *iets te drink-e* 'something to drink-INF'), or the partitive genitive *-s* (as in *iets groot-s* 'something big-PRTT.GEN') would be even more schematic (but still with semantic content). The *-e* suffix in attributive adjectival inflection in Afrikaans (e.g. *lelik-e meisie* 'ugly-e girl') is by and large phonologically determined, without any reference to definiteness, gender or number as is the case in Dutch; however, it is semantically specified to the extent that it presupposes both an adjective and noun (however schematic) on its semantic structure.

With regard to the phonological pole, derivational and inflectional affixes normally have phonological content, often specifying the contexts where they would occur; for example the plural *-e* (Figure 2c) makes reference on the phonological pole to monosyllabic stems that it attaches to prototypically, while the plural *-s* prototypically attaches to non-monosyllabic stems (of course, both have their exceptions). Zero morphemes (Figure 2d), on the other hand, are so highly schematic on the phonological pole, that they are realized as an empty string when they attach to stems; this is of course nothing strange and is accounted for well in a Cognitive Grammar framework (see for example Langacker 1987: 344-345, Tuggy 1992: 275-280).

Both linking and cran morphemes are affix-like with regard to their phonological specification, but their semantic content is even more schematic than those of derivational and inflectional affixes. In fact, their semantic specification is so schematic, that they actually seem to be semantically empty (Tuggy 1992: 280); it is impossible for us to give a characterization of the meaning of *aam-* in *aam-beeld* '*aam*-statue, anvil' or the -s- in *wild-s-bok* 'wild-*s*-buck, antelope' (Figure 2e).

In this sense, linking morphemes and cran morphemes are the opposites of zero morphemes: where zero morphemes have semantic content but no (or very schematic) phonological content, linking morphemes and cran morphemes have phonological content but no (or very schematic) semantic content. I will use the label 'paramorpheme'[8] to refer to these kinds of morphemes that extend so far away from the prototypical morpheme, that they almost seem non-morphemic. They are nonetheless still considered to be morphemes, since we could analyze them as symbolic units (see Tuggy 1992: 275-284 for a lengthy discussion of morphemes with minimal phonological and semantic weight).

3.2 Autonomy and Dependence

Closely related to schematicity, are the notions of autonomy and dependence, referring to the degree to which one structure presupposes at both the phonological and the semantic pole another structure for its manifestation (Langacker 1987: 486, 488, Taylor 2002: 327). The degree to which the internal structure of a component structure makes schematic reference to other structures for its manifestation (Langacker 1991: 7), determines the relative autonomy or dependence of a structure: if no internal specification is made to other structures, we say that the structure is autonomous (e.g. autonomous stems) and if it presupposes other structures, it is dependent (e.g. affixes). Langacker (1987: 300) formulates this relation as follows:

> One structure, *D*, is dependent on the other, *A*, to the extent that *A* constitutes an elaboration of a salient substructure within *D*.

Thus, dependent component structures (*D*) have some salient substructure that needs to be elaborated on by some other structure (*A*); or to use Tuggy's (1992: 242) metaphor: "dependent structures have holes, and their autonomous partners are spikes that fill the holes". These substructures that need to be filled, are called elaboration sites (or e-sites) and are indicated by cross-hatching in the quasipictorial diagrams in Cognitive Grammar (see

also Figure 2b-e). In Figure 3a the correspondence between the e-site and the structure that elaborates it is indicated with a dotted line, while the direction of elaboration is indicated with a line arrow.

E-sites are per definition schematic in their specification: the under- or unspecified substructure needs to be elaborated on by something that adds specificity, something that "fills" the "hole". Of course, in the most prototypical instances, autonomous structures elaborate these e-sites: stems fill the holes of affixes (without implying any directionality within this context). However, it does not have to be an autonomous structure that elaborates a dependent structure (Langacker 1987: 300, Tuggy 1992: 244); compare, for example, cases where neo-classical stems (e.g. *bio-* 'bio-', or *-logie* '-logy') combine to form neo-classical compounds (e.g. *bio-logie* 'technology'), or where two cran morphemes combine to form a word (e.g. *aam-* and *-bei* in *aam-bei* 'hemorrhoid').

Lastly, it is also important to notice that it is not only substructures that could be dependent but also component structures as a whole (see Figure 3b). Relational words (like prepositions or transitive verbs) are usually dependent on the semantic pole only (i.e. they don't depend phonologically on other words, but they require reference to other words for their meaning to be realized) while clitics, for example, are semantically autonomous but phonologically dependent (cf. Taylor 2002: 328). However, affixes are dependent on both the semantic and phonological poles (see Figure 2c) and this is "what makes affixation a morphological rather than a purely phonological or semantic phenomenon" (Tuggy 1992: 243).
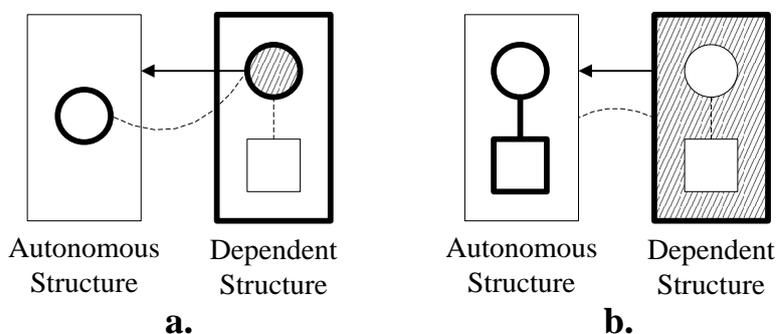

Figure 3. Elaboration of dependent structure

As was the case for specificity, we can distinguish various positions on the continuum of semantic and phonological autonomy and dependence.

Compare Figure 4 for a two dimensional representation of prototypical Afrikaans component structures.
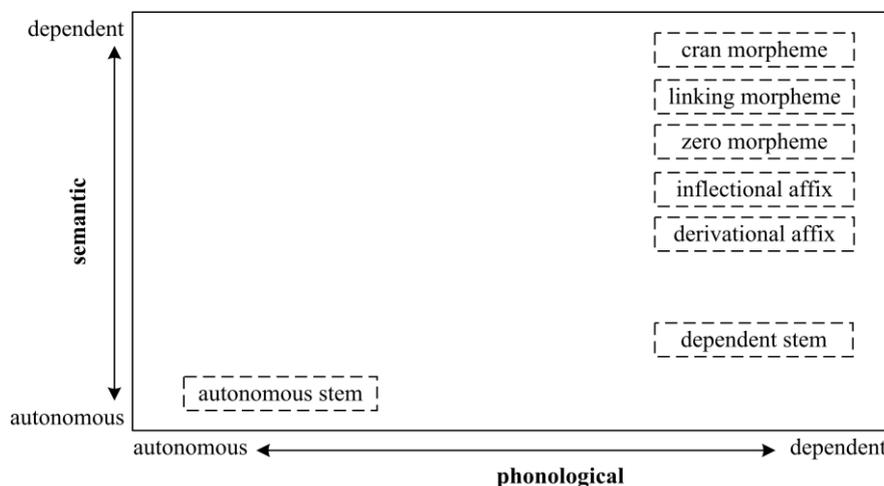

Figure 4. Two-dimensional representation of Afrikaans component structures in terms of autonomy and dependence

As we have already seen with regard to schematicity, autonomous stems are prototypically highly specified both semantically and phonologically and are therefore autonomous in terms of both poles. Even though reduced autonomous stems make internal reference to substructures that are more specified, we can still consider them to be fully autonomous since they can occur as autonomous words. Dependent stems, on the other hand, might be semantically rather autonomous (as is the case for variant dependent stems like *gast-*) but could also be semantically more dependent on elaboration from other component structures (as is the case with non-variant stems like *elektr-*, and neo-classical stems like *Afro-*). Nonetheless, the prototypical dependent stem is semantically more autonomous than affixes or paramorphemes since it is usually semantically more specified than these components. Hence, in Figure 4, dependent stems are indicated somewhat lower on the semantic dependence scale than affixes and paramorphemes.

All dependent stems, affixes, and paramorphemes are phonologically highly dependent, since none of them could occur on their own as words – they all require some other component structure to "fill" (i.e. complement) them phonologically. Hence, in their quasipictorial representations, all these component structures have e-sites on both their phonological and semantic poles (see Figure 2c-e).

3.3 Promiscuity

Taylor (2002: 328-329) defines promiscuity as "the extent to which a unit is free to combine with virtually anything as opposed to the requirement that it combines only with units of a specified kind". For instance, words are generally highly promiscuous, at least phonologically, since they can combine with many other component structures in morphological and syntactic contexts; as Taylor (2002: 174) puts it: "a word can occur adjacent to just about anything… [including] silence (or a hesitation pause)". In a two dimensional representation of the relative positions of Afrikaanse component structures on the promiscuity continuum (see Figure 4), we would then indicate autonomous stems as phonologically highly promiscuous.
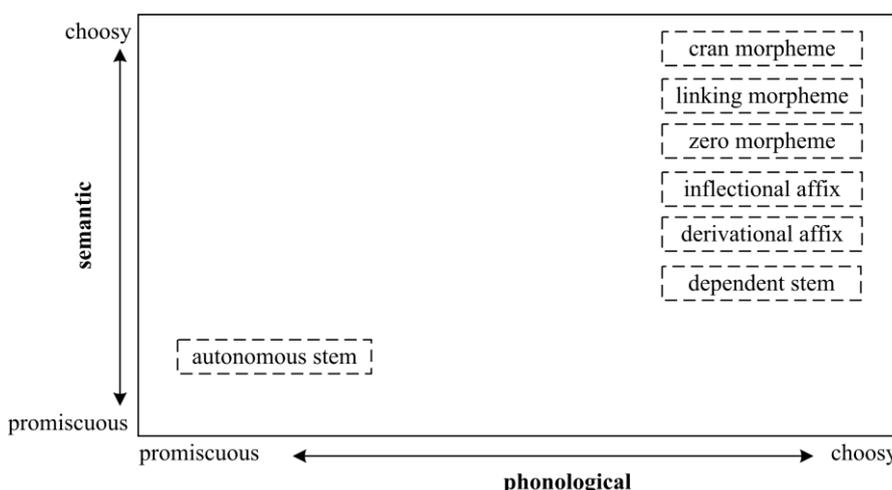

Figure 5. Two-dimensional representation of Afrikaans component structures in terms of promiscuity

Autonomous stems are semantically also rather promiscuous in the sense that they allow for semantic elaborations (e.g. polysemy) and extensions (e.g. homonymy), especially in compounding. However, it should be noted that this applies mainly to autonomous stems from open-classes (i.e. to nouns, verbs, adjectives, and partly to adverbs), while most autonomous stems from closed-classes (such as pronouns, determiners, conjunctives,

etc.) are not generally available for any morphological constructions. Likewise, reduced autonomous stems are also not that promiscuous, since they mostly only occur in compounds (and ever so often in adjectival constructions).

In Afrikaans, however, prepositions represent an exception to this rule, since prepositions occur frequently and productively in separable complex verbs, such as *op-staan* 'up-stand, stand up/rise', *onder-gaan* 'under-go, go down/suffer', or *af-takel* 'off-rig, unrig/dismantle'. In these contexts, they are often referred to as particles (i.e. parts of separable complex verbs). Booij (2002: 202-224) gives a thorough analysis of this phenomenon in Dutch, explaining how the peculiarities of separable complex verbs (such as the past tense of these verbs) should be understood in terms of what he calls constructional idioms. He argues that particles "can be seen as intermediate stages in the development of words into bound morphemes, in particular prefixes" (i.e. word > part of separable complex verb > prefix) and that this change "implies a loss of lexical meaning" (Booij 2002: 218). Compare for instance the homographic *voor-kom* 'before-come', which can mean (1) 'appear/occur to', or (2) 'prevent'. In its first usage, the preposition *voor-* has the status of a particle, since the past tense form of the verb is *voor-ge-kom* 'before-PST-come, appeared/occurred'. However, in the second instance *voor-* already has reached the state of a prefix in the grammaticalization process (with a meaning similar to 'pre-'); hence the past tense is realized by a zero morpheme (i.e. Ø-*voor-kom* 'PST-before-come, prevented'), as is the case in other verbs with the prefixes *be-* (as in *be-twis* 'contest'), *ge-* (as in *ge-dra* 'behave'), *er-* (as in *er-ken* 'recognize'), *her-* (as in *her-sien* 'revise'), *ont-* (as in *ont-slaan* 'dismiss'), and *ver-* (as in *ver-skoon* 'excuse') (Taalkommissie 2002: 189). So, we can conclude that prepositions, as items from a closed-class that are generally not promiscuous, should be analyzed as either particles (i.e. as autonomous stem-like components) or prefixes, and thereby explaining their higher degree of promiscuity.

All other component structures are both phonologically and semantically rather choosy in terms of their combinatory potential in morphological constructions. Affixes combine only in certain contexts with certain stems; for example, the Afrikaans partitive genitive morpheme is highly choosy (even more so than its Dutch counterpart), since it combines only with certain adjectives in highly specified syntactic context (i.e. only after indefinite pronouns). The same applies to all other affixes.

Linking morphemes could be considered choosy, since they don't occur systematically in various phonological contexts (Combrink 1990), neither do they occur systematically in all compounds. Likewise, the variety and occurrence of zero morphemes are highly limited and specified, while cran morphemes couldn't be considered productive at all.

One important aspect that should be kept in mind when considering valence of component structures in Afrikaans (as is the case in Dutch), is that Afrikaans component structures could, in broad terms, be divided in two strata: a native (Germanic) stratum and a nonnative (Romance) stratum. Examples of suffixes from the native stratum includes *-heid* as in *besig-heid* 'business', *-ing* as in *bots-ing* 'collision', or *-aard* as in *ryk-aard* 'wealthy person'; examples from the nonnative stratum include *-teit* as in *puber-teit* 'puberty', *-iet* as in *transvest-iet* 'transvestite', or *-eur* as in *kontrakt-eur* 'contractor'. With regard to promiscuity, it can be taken as a generalizing rule that nonnative suffixes combine only with nonnative stems, while native suffixes combine with both native and nonnative stems (Booij 2002: 95), thereby limiting the promiscuity of these component structures. There are of course, once again, various examples that extend from this pattern. For example, the nonnative adjectivalizer *-ief* not only combines with nonnative stems (as in *kongest-ief* 'congestive', or *abort-ief* 'abortive'), but also with native stems (as in *sport-ief* 'sportive', or *fout-ief* 'faulty'); whereas, contrary to the principle stated above, the native personifier *-aard* only combines with native stems (as in *woest-aard* 'savage' or *grys-aard* 'elder').

One should note that, when applying strata considerations to promiscuity, component structures could be hybrids or could show signs of homonymity. Booij (2002: 96) illustrates for example that in Dutch, complex units with the native suffix *-isch* (as in *afgod-isch* 'idolatrous', *fantast-isch* 'fantastic' or *Belg-isch* 'Belgian') have no marked status with regard to strata restrictions and *-isch* could therefore be considered a hybrid; the same would apply to the *-eer* suffix in Afrikaans, in examples such as *fout-eer* 'err' (combination with native stem) and *kommunik-eer* 'communicate' (combination with nonnative stem). The Afrikaans nominalizing suffix *-asie* (as in *kommunik-asie* 'communication' or *lekk-asie* 'leakage') is actually homonymous with different origins; in Dutch, this is still reflected in the orthography with two different suffixes, *-atie* (as in *communic-atie* 'communication) and *-age* (as in *lekk-age* 'leakage').

3.4 Constituency

Constituency is the last parameter in terms of which component structures could be characterized. It is defined as "the order in which component structures are successively combined to form progressively more elaborate composite structures" (Langacker 1987: 310). Since Langacker (1987: 317) explicitly states that constituency is a secondary valence factor, where the "choice of a particular constituency arrangement is often not critical" (Langacker 1987: 310, see also Tuggy 2005: 257), it will suffice here to be only concerned with constituency on the phonological pole – where do component structures attach to each other? This parameter is necessary to make a distinction between different affixes (i.e. prefixes, suffixes, ambifixes), including linking morphemes (as they tend to be affix-like).

Since our concern here is not with composite structures, I will not elaborate on how many of which component structures are possible in composite structures; I will rather assume the simplest forms of constructions with the minimal number of component structures on a first level of constituency (i.e. lowest level of combination). Hence, for purposes of this article, I assume that n component structures are available in n combinations on n levels of constituency; this is not reflected in Figure 6.
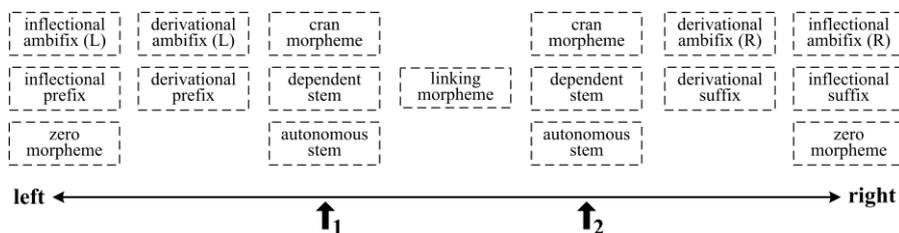

Figure 6. Relative positions of component structures in composite structures

As discussed before, stems are compulsory component structures in any morphological composite structure and are hence considered the nuclei in any complex morphological unit. The constituency of all other component structures could therefore be defined in terms of their position relative to stems. In Figure 6 autonomous and dependent stems are thus depicted as central on the diagram and indicated with two bold arrows (Arrow 1 being the main vantage point); two sets of stems are indicated, to illustrate the possibility of compounding.

Given the restricted valence of cran morphemes, they could also be considered central constituents in morphological constructions. Cran morphemes connect either to the right of an autonomous stem (such as *-ves* in *huis-ves* 'accommodate') or to the left of an autonomous stem (such as *stie-* in *stiebeuel* 'stirrup'). In some (debatable) occasions, two cran morphemes can combine to form a word; compare for instance *aam-* and *-bei* in *aam-bei* 'hemorrhoid'. Combrink (1990: 28) also indicates that cran morphemes can partake in affixation constructions, by either combining with prefixes (like *-rep* in *onge-rep* 'untouched') or with suffixes (like *hag-* in *haglik* 'precarious').[9] In Figure 6 cran morphemes are therefore indicated as central component structures, in line with stems.

Linking morphemes occur in the middle of two stems (e.g. *perd-e-kar* 'horse cart') with a strong phonological affinity with the left-hand stem (i.e. *perdekar* is hyphenated as *per-de-kar*). If we take Combrink's (1990) viewpoint on linking morphemes, we should also recognize that linking morphemes attach to the right of non-variant dependent stems to allow for further derivation (e.g. the *-ifi-* in *fals-ifi-eer* 'falsify', or *-at-* in *idiom-at-ies* 'idiomatic').

For purposes of this article, and without going into any detail, I ascribe to the viewpoint that inflectional affixes are peripheral to derivational affixes and that derivational affixes are therefore indicated closer to the nucleus in Figure 6. Suffice it to say that prefixes attach to the left of a stem, suffixes to the right, and ambifixes simultaneously to the left and right of a stem. Since zero morphemes are always affix-like, they are positioned in line with other inflectional affixes.

Note again that the representation in Figure 6 only pertains to a first level of constituency. On a next level of constituency it would be possible, for instance, to add autonomous and dependent stems to both sides of the diagram. Phenomena such as the past tense of separable complex verbs (e.g. *uit-ge-skop* 'out-PST-kick/kicked out') or compounding on already inflected forms (e.g. *kat-jie-kos* 'cat-DIM-food/kitten food') could thus be accounted for.

3.5 Summary

In this section we have characterized component structures in Afrikaans morphological constructions, using four valence factors as parameters, viz. schematicity, autonomy/dependence, promiscuity, and constituency. Based on the discussion above, we can summarize our discussion in Table 1,

where each component structure is described in terms of each of the parameters. Note that phonological and semantic dependence have been split in the table; this provides us with a more accurate description, especially of dependent stems. Also note that, as in the discussions, this table does not in any way aim to represent category membership as absolute; it is a mere visual representation and summary of the discussions above and, as in the discussions, it is based on prototypical instances of each category.

	Schematicity	Phon Dependence	Sem Dependence	Promiscuity	Constituency
Autonomous stem	Fully specified	Fully autonomous	Fully autonomous	Fully promiscuous	Nucleus
Full	Fully specified	Fully autonomous	Fully autonomous	Fully promiscuous	Nucleus
Reduced	Fully specified	Fully autonomous	Fully autonomous	Partially promiscuous	Nucleus
Dependent stem	Partially specified	Fully dependent	Partially dependent	Choosy	Nucleus
Variant	Partially specified	Fully dependent	Partially dependent	Choosy	Nucleus
Non-variant	Partially specified	Fully dependent	Partially dependent	Choosy	Nucleus
Neo-classical	Partially specified	Fully dependent	Partially dependent	Choosy	Nucleus
Affix	Partially schematic	Fully dependent	Fully dependent	Choosy	Peripheral
Prefix	Partially schematic	Fully dependent	Fully dependent	Choosy	Left
Suffix	Partially schematic	Fully dependent	Fully dependent	Choosy	Right
Ambifix	Partially schematic	Fully dependent	Fully dependent	Choosy	Both sides
Paramorpheme	Fully schematic	Fully dependent	Fully dependent	Choosy	Nucleus/Peripheral
Linking	Fully schematic	Fully dependent	Fully dependent	Choosy	Left
Zero	Fully schematic	Fully dependent	Fully dependent	Choosy	Left/right/both sides
Cran	Fully schematic	Fully dependent	Fully dependent	Choosy	Nucleus

Table 1. Summary of characteristics of component structures

Based on Table 1, we can now easily derive definitions for each of the component structures. For example, a reduced autonomous stem can be defined as a fully specified, fully autonomous, and partially promiscuous symbolic unit that appears in the nucleus of complex symbolic units while a linking morpheme would be defined as a fully schematic, fully dependent, and choosy symbolic unit that attaches to the left of the nucleus in a complex symbolic unit. Other definitions can also be derived in the same way.

In the next section I will illustrate how these definitions can be made operational by applying them in the context of a project to develop an automatic morphological parser for Afrikaans.

4. Application: Toward a Taxonomy of Component Structures

As explained in the introduction of this article, the outcomes of this research are being applied in a very concrete and practical field, namely the annotation of linguistic data. In a project to develop an automatic morphological parser for Afrikaans, using data-driven approaches (specifically machine learning), we need to annotate data for training the machine-learning algorithms. It is therefore quintessential (for us) to have clear-cut morphological categories for the accurate and detailed annotation of linguistic data.

However, one of the very basic assumptions in Cognitive Grammar is that "much in language is a matter of degree" (Langacker 1987: 14). In its very essence Cognitive Grammar defies the criteria-attribute and dichotomous models of categorization and instead ascribes to models of prototypes, where category members are described on continua in terms of parameters of gradation. The descriptions of component structures above serve as examples.

This assumption of Cognitive Grammar is therefore to a large extent at odds with the specific needs of this subpart of the project, where we have to operate with clear-cut, hard-and-fast categories for the annotation of data. So, the challenge is to translate the gradations and fuzzy categories of a Cognitive Grammar description into clear-cut, well-defined categories in order to make these descriptions more accessible and/or usable in an annotation environment. One way of doing this is by means of taxonomies (see Taylor 2002: 128-139).

Given the descriptions above (and specifically the representation in Table 1), we are able to "draw lines through" the identified characteristics in order to "translate" our descriptions in a taxonomy. Without disputing any of the assumptions of Cognitive Grammar, we are nonetheless forced to make hard-and-fast decisions about categories, in order to postulate a workable categorization network (i.e. taxonomy). The result of this process is presented in Figure 7.
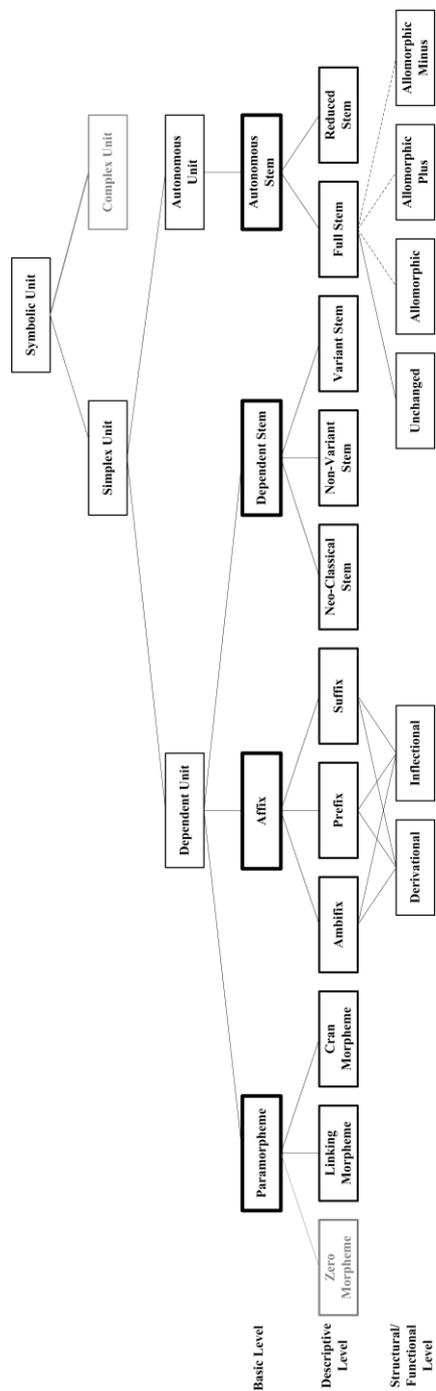
Figure 7. Taxonomy of Afrikaans component structures for morphological analysis

On the basic level of categorization, we distinguish between the four main categories of component structures in terms of schematicity (as labeled in Table 1), thus identifying autonomous stems, dependent stems, affixes, and paramorphemes as basic level categories (indicated in heavy-lined boxes in Figure 7). The alternative would have been to categorize only one category of stems (together with affixes and paramorphemes) on the basic level (with "simplex unit" as the superordinate) and instead distinguish between autonomous and dependent stems on a subordinate level of categorization. However, such a classification would have ignored the fact that affixes and paramorphemes are also categorized as dependent units; hence, the distinction between autonomous and dependence (see Table 1) is made on a higher, superordinate level.

On a more descriptive level (specifically categorized in terms of constituency, but also taking the other parameters into account) we identify the various subcategories for each basic level category, such as linking morpheme, prefix, variant stem, reduced stem, etc. Since these are the important labels for annotation of data (it is more important to know that a constituent is a prefix, than only to know that it is an affix), one could also argue that this is the basic level of categorization; however, for purposes of this project, this "theoretical"/"psychological" distinction is not that significant – it is only important to recognize these categories as finer-grained subcategories of the higher level categories. To indicate the salience of these categories, we also indicate them with a heavier-lined box in Figure 7.[10]

On lower levels of categorization, one can now add various subcategories as needed. For instance, we would like to distinguish between inflectional and derivational affixes; this is indicated in Figure 7 on the level of (what we call) structural and/or functional specification (e.g. the suffix *-e* functions sometimes as an inflectional suffix and sometimes as a nominalizer). For both these categories further/deeper levels of specification are indicated, the details of which are not important for purposes of this article (see Groenewald & Van Huyssteen 2008). Note that we don't take promiscuity (and information related to strata) into consideration in the postulated network; the reason being that we want the eventual classifier to discover this information automatically (unsupervised), based on pure statistics.

For the annotation of the data, we take an Item-and-Arrangement approach to analysis (Crystal 1997: 206) since the software that we use for annotation doesn't allow for annotation/indication of morphonological

processes. We therefore analyze complex units as "beads on a string": *elektrisiteit* 'electricity' is analyzed as *elektr+isi+teit* and *perdekar* 'horse cart' as *perd+e+kar*. However, one runs into trouble in cases where morphonological changes occur in words. Consider, for example, *katte* 'cats', whose correct analysis is *kat*-PL (where PL is the suffix *-e*). Should one now analyze it as *kat+te*, or rather as *katt+e*? (Note that we can't just ignore the second *t*, since all characters in a word need to be annotated.)

To solve this challenge, we identify on the level of structural/functional specification four "artificial", structural categories for purposes of this project. The category "unchanged" is used for autonomous stems where no morphonological processes have an effect, e.g. *hond-e* 'dog-PL, dogs' or *venster-tjie* 'window-DIM, small window'. The category "allomorph" (the term is here understood merely as a label and without any theoretical implications) is used to annotate stems with changes/alternations on the final consonants, e.g. *wolw-e* 'wolf-PL, wolfs' (stem is *wolf*), or *konink-ie* 'king-DIM, little king' (stem is *koning*). In cases where an identical consonant is added due to morphonological process, we use the category "allomorph plus", e.g. *katt-e* 'cat-PL, cats' (stem is *kat*) or *ball-etjie* 'ball-DIM, small ball' (stem is *bal*). The last category, "allomorph minus", is used in cases where an identical vowel has been deleted due to morphonological processes, as for example in *jar-e* 'year-PL, years' (stem is *jaar*) or *strep-ie* 'line-DIM, short line' (stem is *streep*).

The latter three subcategories actually represent mere extensions from the schema indicated on a higher level of specification and pose in so far no threat for the integrity of the theoretical basis of the categorization network. In accordance with conventions in Cognitive Grammar, these extensions of the network are indicated with dashed lines.

To illustrate the applicability of this categorization network, consider the following analysis of the word *elektrifikasieleidingsnetwerke* 'electrification pipe-line networks'.

Component	Categorization
elektr	non-variant dependent stem
ifik	linking morpheme
asie	nominalizing derivational suffix
leid	variant dependent stem
ing	nominalizing derivational suffix
s	linking morpheme
net	unchanged full autonomous stem
werk	unchanged full autonomous stem
e	plural inflectional suffix

Table 2. Analysis of *elektrifikasieleidingsnetwerke* 'electrification pipe-line networks'

For annotation purposes, all nodes in the taxonomy are converted to abbreviated tags (e.g. *NDS* for *non-variant dependent stem* or *PIS* for *plural inflectional suffix*), which are assigned automatically to component structures by the software during manual annotation. Thus far, circa 32,000 words have been manually annotated successfully, based on the taxonomy in Figure 7. In the next stage of the project, this data will be used as training data for a classifier that will automatically analyze (parse) new data.

5. Conclusion

This article aims at solving some practical issues related to the analysis of morphological data for a human language technology project (i.e. the development of a morphological parser for Afrikaans). These issues are approached theoretically from a Cognitive Grammar perspective in order to characterize component structures in Afrikaans morphological constructions.

Various component structures are described in terms of four basic valence factors, viz. schematicity, autonomy/dependence, promiscuity, and constituency. Based on these parameters, a theoretically unified, unambiguous categorization network (specifically a taxonomy) for component structures in Afrikaans morphology is postulated. It is indicated how this taxonomy is extended for specific purposes of the project and how it is implemented to manually annotate data.

Future work will include research along the same lines, but now focusing on a categorization network of composite morphological structures in Afrikaans. This will of course be a vast task, but nonetheless important for a better understanding of morphological constructions in general and specifically for Afrikaans.

Based on this work, one could also adapt the current categorization network for other languages – especially languages with a more complex morphology than Afrikaans. This could not only prove valuable for a better understanding of the specific languages, but also to provide new and deeper insights in Cognitive Grammar as a theoretical descriptive framework.

Notes

1. Combrink's 1990 book, *Afrikaanse Morfologie: Capita Exemplaria*, is the most comprehensive and also the most recent stand-alone publication on Afrikaans morphology. This work of Combrink is purely descriptive and aims to steer clear of any affiliation with any theoretical framework. As Combrink (1990: 9) states in the preface: "One way of avoiding theoretical linguistic "nearsightedness", is to analyze data by using mature theoretical linguistic notions and to adapt these notions as the data prescribes" [My translation – GBVH]. If any theoretical presuppositions slipped into his work, it is either idiosyncratic, or Structuralist in nature.

2. By saying this, I'm not negating the fact that various other Cognitive Grammar scholars, such as Eugene Casad, Kenneth W. Cook, Hans-Olav Enger, Laura Janda, Suzanne Kemmer, Tore Nesset, Johanna Rubba, Brygida Rudzka-Ostyn, and Arie Verhagen, have done valuable work in describing various morphological phenomena in various languages. Of course, the work by Ron Langacker, John Taylor and David Tuggy is fundamental to an understanding of morphology in Cognitive Grammar, while Vyv Evans and Melanie Green offer a valuable summary.

3. See Tuggy (2005: 259-260) for a discussion on the morphology-syntax boundary.

4. These component structures (such as the *-e-* in *perd-e-kar* 'horse cart') are also referred to as link phonemes, linking elements, interfixes, phonomorphemes, connecting morphemes, linkers, stem extenders, valence morphemes, etc.

5. Various other valence factors, such as correspondence, profile determinacy, analyzability (Langacker, 1987), coerciveness and bondedness (Taylor 2002: 330), internal and external complexity (Taylor 2002; Tuggy 1992), and entrenchment are often discussed in Cognitive Grammar literature. However, for purposes of a characterization of component structures (and not of

composite structures or constructions), the valence factors mentioned in this article will suffice.

6. Following the arguments of Tuggy (1992: 287), I also prefer to use the term "stem" instead of "root", and I, thus, also refrain from making a distinction between these two terms. Furthermore, this distinction is obsolete with regard to Afrikaans morphology; see also Combrink (1990). Note, however, that what I will call "non-variant dependent stems" (e.g. *elektr-* in *electricity* 'elektrisiteit'), are also often referred to as roots.

7. Under the heading "bound non-variant stems" Combrink (1990: 26) lists examples like *send-* in *sendeling* 'missionary' and *ren-* in *renperd* 'race horse'. Since both *send* 'send' and *ren* 'run' are autonomous words in Afrikaans (both these words appear as lemmas in standard Afrikaans dictionaries), this analysis of his is clearly incorrect. He categorizes stems like *elektr-* and *sekret-* as 'truncated free stems', which also seems to be a misnomer for these kinds of stems; for one, a stem like *sekretar-* combines with suffixes like *-is* or *-esse* (both recognized by Combrink (1990) as suffixes) to form *sekretaris* 'secretary' and *sekretaresse* 'female secretary', thereby voiding the interpretation of these stems being truncated.

8. To my knowledge, "paramorpheme" is not a widely used term; in fact, I found only two references to this notion. Trager (1953: 327) uses the term to name morphemes "that constitute a set of inflectional suffixes in a paradigm", while De Groot (1964: 127) uses it to denotes what I call here a linking morpheme (i.e. a morpheme that heightens the valence of two components to combine). In my terminology, I discard the notion as applied by Trager (1953), while extending the notion's meaning as used by De Groot (1964), to include other morphematic forms that extend far away from the prototype.

9. Diachronically speaking, none of the examples listed could be considered cran morphemes; synchronically speaking, they are. This raises the question of the difference between morphology and etymology. For purposes of this article, I support Plag (2003) when he states that "we have to be careful not to confuse morphology with etymology. Even though a morpheme may have had a certain meaning in the past, this does not entail that it still has this (or any other) meaning and can thus be considered a morpheme in today's language." (Plag 2003: 25).

10. Since zero morphemes have no realization in the orthography, it is impossible to annotate zero morphemes. For the sake of completeness, I include this node in grey in Figure 7, to indicate that it is not part of the practical implementation in this project. Likewise, I include "complex unit" for the sake of completeness, although this is not a label used in the data annotation project.

Acknowledgments

References

Aronoff, Mark, and Kirsten Fudeman
    2005   *What is Morphology?* Malden/Oxford/Victoria: Blackwell.
Bauer, Laurie
    2003   *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.
    2004   *A Glossary of Morphology.* Edinburgh: Edinburgh University Press.
Booij, Geert, Christian Lehmann, and Joachim Mugdan
    2000   *Preface.* In *Morphology: An International Handbook on Inflection and Word-Formation Vol. I*, Geert Booij, Christian Lehmann and Joachim Mugdan (eds.), xxi-xxviii. Berlin/New York: Walter de Gruyter.
    2002   *The Morphology of Dutch.* Oxford: Oxford University Press.
    2007   *The Grammar of Words: An Introduction to Morphology.* Oxford/New York: Oxford University Press.
Combrink, Johan G.H.
    1990   *Afrikaanse Morfologie: Capita Exemplaria* [Afrikaans Morphology: Capita Exemplaria]. Pretoria: Academica.
Crystal, David
    1997   *A Dictionary of Linguistics and Phonetics.* 4th Edition. Oxford: Blackwell.
De Groot, A.W.
    1964   *Inleiding tot de Algemene Taalwetenschap [Introduction to General Linguistics].* 2nd Revised Print. Groningen: J.B. Wolters.
De Klerk, Gert J.
    1968   Die Morfologie van Afrikaans [The Morphology of Afrikaans]. In *Afrikaans – Sy Aard en Ontwikkeling* [Afrikaans – Its Nature and Development], Van der Merwe, H.J.J.M. (ed.), 169-208. Pretoria: J.L. van Schaik.
De Villiers, Meyer

1983   *Afrikaanse Grammatika vir Volwassenes*. [Afrikaans Grammar for Adults]. Goodwood: Nasou.

Evans, Vyvyan
    2007   *A Glossary of Cognitive Linguistics.* Edinburgh: Edinburgh University Press.

Evans, Vyvyan, and Melanie Green
    2006   *Cognitive Linguistics: An Introduction.* Edinburgh: Edinburgh University Press.

Gleason, H.A.
    1955   *An Introduction to Descriptive Linguistics.* Revised Edition. New York: Holt, Rinehart and Winston.

Groenewald, Handré J, & Gerhard B Van Huyssteen
    2008   Outomatiese Lemma-identifisering vir Afrikaans [Automatic Lemmatization for Afrikaans]. *Literator.* 29(1): 65-91.

Haspelmath, Martin
    2002   *Understanding Morphology.* London: Arnold.

Hudson, Richard
    1984   *Word Grammar.* Oxford/New York: Basil Blackwell.

Katamba, Francis
    1993   *Morphology.* London: Macmillan.

Langacker, Ronald W.
    1987   *Foundations of Cognitive Grammar Volume I: Theoretical Prerequisites*. Stanford: Stanford University Press.
    1991   *Concept, Image, and Symbol: The Cognitive Basis of Grammar.* Berlin/New York: Mouton de Gruyter.

Matthews, P.H.
    1991   *Morphology.* 2d ed. Cambridge: Cambridge University Press.

Plag, Ingo.
    2003   *Word-Formation in English.* Cambridge: Cambridge University Press.

Senekal, H.E.J., Fritz A. Ponelis, and Willem J. de Klerk
    1972   *Die Patroon van Afrikaans* [The Pattern of Afrikaans]. Johannesburg: Afrikaanse Pers-Boekhandel.

Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns
    2002   *Afrikaanse Woordelys en Spelreëls [Afrikaans Word-list and Spelling Rules].* 9th Edition. Cape Town: Pharos.

Taylor, John R.
    2002   *Cognitive Grammar.* Oxford: Oxford University Press.

Trager, George
    1953   Russion Declensional Morphemes. *Language* 29, no. 3 (July-September 1953): 326-338.

Tuggy, David

1985   The Inflectional/Derivational Distinction. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session XXIX*. Grand Forks, N.Dak.: Summer Institute of Linguistics, University of North Dakota.

1992   The Affix-Stem Distinction: A Cognitive Grammar Analysis of Data from Orizaba Nahuatl. *Cognitive Linguistics* 3 (3): 237-300.

2005   Cognitive Approaches to Word-Formation. In *Handbook of Word-Formation*, Pavol Štekauer and Rochelle Lieber, (eds.), 233-265. Dordrecht: Springer.

Van Schoor, J.L.

1983   *Die Grammatika van Standaard-Afrikaans* [The Grammar of Standard Afrikaans]. Kaapstad: Lex Patria.