ANNA: A Dictionary with a Name (and what Lies Behind it)[*]

W. Martin, *Department of Language and Communication, VU University Amsterdam, The Netherlands (w.j.r.martin@vu.nl)*

Abstract: In 2011 the *Groot Woordenboek Afrikaans en Nederlands (Large Dictionary Afrikaans and Dutch),* commonly known as ANNA, appeared. Contrary to so-called *difference dictionaries,* bilingual dictionaries of narrowly related languages which describe *only differences* between the two languages, ANNA describes both differences and similarities between Afrikaans and Dutch, not only on the semantic level but on the combinatorial and pragmatic level as well. In this sense ANNA is a unique project, based on an original *amalgamation* model. In this article first some background information will be given about the ANNA project and its results, followed by a presentation of the underlying model and an evaluation of it.

Keywords: AMALGAMATED DICTIONARY, BILINGUAL DICTIONARY, AMALGAMATION MODEL, AFRIKAANS, DUTCH, ANNA, COGNATES, FALSE FRIENDS, CONTRASTIVE DICTIONARY, READING DICTIONARY, TRANSLATION DICTIONARY

Samenvatting: ANNA: een Woordenboek met een Naam (en wat er achter steekt). In 2011 verscheen het *Groot Woordenboek Afrikaans en Nederlands,* met als roepnaam *ANNA.* In tegenstelling tot een traditioneel tweetalig woordenboek beschrijft ANNA de twee talen (Afrikaans en Nederlands) als één taal en behandelt zij niet alleen de onderlinge verschillen maar ook de (vaak vermeende) gelijkenissen op het vlak van betekenis, combinatoriek en pragmatiek. Dit maakt ANNA tot het eerste ge-amalgameerde tweetalige woordenboek gebaseerd op een origineel *amalgamatiemodel.* In dit artikel wordt allereerst achtergrondinformatie gegeven over het ANNA-project en zijn resultaten. Daarna wordt het onderliggende *amalgamatiemodel* voorgesteld en geëvalueerd.

Sleutelwoorden: AMALGAMATIEMODEL, GEAMALGAMEERD WOORDENBOEK, TWEETALIG WOORDENBOEK, AFRIKAANS, NEDERLANDS, ANNA, COGNATES, VALSE VRIENDEN, CONTRASTIEF WOORDENBOEK, LEESWOORDENBOEK, VERTAALWOORDENBOEK

0. Introduction

In March 2011, the *Groot Woordenboek Afrikaans en Nederlands* (Large Dictionary Afrikaans and Dutch, 2228 pgs.), known as ANNA, the *first amalgamated (bilingual) dictionary* ever, appeared. Since then I have given many lectures, seminars

---

[*] Two other articles where ANNA has been presented extensively are: Martin 2012 (in Genis et al. 2012) and Martin (to appear in *Tydskrif vir Nederlands en Afrikaans* in 2012).

and presentations on ANNA. In this contribution then, I will proceed as I have done in most of those speeches up till now: starting with some background and inside information about ANNA which, otherwise, is difficult for readers to get (section 1, *factual* part), to continue with a more *reflective*, *evaluative part* in which most attention will be paid to the model underlying the dictionary (sections 2 and 3).

1.   ANNA: Background Information

1.1   What's in a name?

ANNA is a translation or bilingual dictionary Afrikaans–Dutch v.v. which, contrary to what is generally the case, does not treat the two languages as different ones but as the *same.* From this point of view the acronym AN-NA is rather misleading as it suggests that the dictionary consists of two parts, one part Afrikaans–Dutch (AN), — N standing for Nederlands (=Dutch) — and one part Dutch–Afrikaans (NA). ANNA, however, contains only one part, implying that both Afrikaans and Dutch can function in one and the same volume as a source language.

In a traditional bilingual dictionary Afrikaans–Dutch v.v. an entry such as *robot*, for instance, would look as follows:

In the Dutch–Afrikaans part:

robot
(automaat die arbeid verricht) [automaton carrying out work]   *robot, blikman*

In the Afrikaans–Dutch part:

robot
(paal met lig wat verkeer reël) [pole with light regulating traffic]   *stoplicht, verkeerslicht*
(outomaat wat werk verrig) [automaton carrying out work]   *robot*

In ANNA the information from the two parts is brought together, *amalgamated*, resulting in an entry like this:

robot/*robot*
A/N (automaat die arbeid verricht)   *robot, blikman*
A   *(paal met lig wat verkeer reël)*   stoplicht, verkeerslicht
[in ANNA roman font is used for Dutch; italics for Afrikaans]

As I will deal more extensively with amalgamation in the next part (sections 2 and following), it may suffice for the moment to draw the attention to the fact that *amalgamation* has the following effects:

—   First of all, it leads to a (more) direct comparison of the two languages:

one does not have to look up the entry *robot*, for instance, in both parts to get a complete picture of both the differences and the similarities between this word in both Afrikaans and Dutch.

— Secondly, amalgamation also leads to a decrease of redundancy in the description of the two languages: in the case of *robot*, for instance, the common meaning 'automaton' need not be repeated in both parts, one mention now suffices.

— Finally, since amalgamation is a novel approach in lexicography[1], ANNA, as the first amalgamated bilingual dictionary ever, can serve as a kind of litmus test or touchstone for the evaluation of the model itself.

1.2   Facts and Figures

1.2.1   Temporal and financial aspects

The (editing) work on ANNA started in January 2000 and was completed in March 2011. Before that, a pilot-study was undertaken (see Martin, Gouws and Renders 1999) in order to define the project and to investigate its feasibility. Time was underestimated in the pilot as only a period of six years was foreseen. On the other hand, the financial prognosis proved to be rather realistic: the budget was overrun by only 25%. All in all, ANNA was a rather 'cheap' project as the project costs did not exceed 400,000 euros[2]. This financial asset was mainly due to the fact that use could be made of two invaluable pieces of *lexicographical infrastructure,* namely an existing Dutch Database (the *RBN* = Referentie Bestand Nederlands = Reference Database of Dutch) and an editor with reversal function at semantic level, namely *OMBI* (= editor for OMkeerbare BIlinguale Bestanden, editor for Reversible Bilingual Databases)[3]. Indeed, the fact that the RBN could be used provided us already with the macro-and the microstructure of the Dutch side of the N-A part, and, while semantically linking Dutch to Afrikaans by means of OMBI, the A-N part was being constructed for the greater part (see below).

However, the fact that this project left the down-trodden lexicographical paths to follow a model of its own (the amalgamation model) lead to the usual teething troubles and entailed some delay. In fact the project has been developed in four steps/phases:

— step one: the elaboration of the Dutch–Afrikaans part
This part of the project took most of the time (4,5 years: from beginning 2000 to mid-2004), among others, because of the teething problems just mentioned. So, for instance, it took some time before the editors got acquainted with the new model and with working with OMBI. The work was carried out at the University of Stellenbosch in close collaboration with the *VU* University of Amsterdam**.**

— step two: the elaboration of the Afrikaans–Dutch part
 The work for this part was done at Port Elizabeth in close collaboration with Amsterdam. It took 3,5 years (from mid-2004 to end 2007).

— step three: the amalgamation and its editing
 This step involved the amalgamation of the two previous parts and its editing; it was mainly carried out in Amsterdam in close collaboration with Port Elizabeth. It took 2 years (2008–2009) to finalise this phase.

— step four: final correction and production phase
 The overall correction took one full year (2010) and was mainly carried out in Amsterdam and in Houten near Utrecht (publisher's place).

A couple of words of comments on these aspects may be in order here:

1. Work on this project has been carried out at locations often at a distance of several thousands of miles from each other. Thanks to modern ICT this has not been a major problem.

2. This project has been developed *stepwise.* In doing so, it was important to provide in step 1 for the information needed for steps 2 and 3. In the case of 'robot', for instance, it was necessary to indicate (by means of the (non)-appearance of the marker c2) that items were cognates or not:

 N. robot (= automaat) = A. robot (= outomaat) = c2
 N. stoplicht (= verkeerslicht) = A. robot (= verkeerslig)

Mainly on the basis of this information, OMBI and the amalgamation program, the data needed for step 2 (Afrikaans–Dutch) and step 3 (amalgamation) could be *automatically derived*, so to come to:

 (step 2)
 robot
 1. (verkeerslig) = stoplicht (verkeerslicht)
 2. (outomaat) = robot (automaat) = c2

 (step 3)
 robot/*robot*
 A/N   1.(automaat) *robot*
 A   2.(*verkeerslig)* stoplicht

See section 2 for more details on the amalgamation.

1.2.2 ANNA: a multinational/multilingual project

ANNA, as is the case with most bilingual dictionary projects, is a *multinational*

project, involving a small team of editors spread over three countries: South Africa, the Netherlands and Flanders. The fact that the editorial staff was relatively small was not a matter of principle but simply a pragmatic fact: it was not easy to find people with the necessary qualifications willing to engage in a long-running project of this nature. On the other hand, the small size of the team (ten members) offered the advantage of greater consistency. The project was lead by an editorial board consisting of five people[4] in which I functioned as the project leader.

ANNA is not only a multinational but also a *multilingual* project. At first sight, there is nothing remarkable in that, given the fact that all bilingual projects are, by definition, concerned with more than one language. However, ANNA, right from the start, has been set up as a project that should supersede the two languages any bilingual dictionary project involves, and in this sense ANNA is more than just *bi-lingual.* This has to do with the infrastructure developed for this project and which is meant to be (re-)usable for other languages than those involved in the language pair at stake (Afrikaans and Dutch). I refer to the section 1.3.4 (on the reasons to start up this project) for further details.

1.2.3   ANNA in numbers

ANNA counts 2228 pages and has a weight of 2,2 kg. Next to this 'heavy' print version there is also a 'light' version of ANNA in the form of a CD-ROM[5]. The latter, in fact, does not differ from the former as to contents, but offers, of course, the usual facilities and advantages electronic dictionaries have over their print companions.

The book contains 59051 dictionary entries among which there are:

—   29840 *cognates* or *combination words*
      These are words which have the same form in Dutch and Afrikaans and, at least one common meaning (for 'sameness' of form see below, section 2.2).
      Examples: tafel/*tafel* **[E. table], ontsnappen/***ontsnap* **[E. escape], gel/***jel* **[E. gel].**

—   15011 *unique Afrikaans items*
      These are words that only occur in Afrikaans. As a rule they have in Dutch a non-cognate translation equivalent or a paraphrase.
      Examples: *trapsoetjies* **[N. kameleon, E. chameleon],** *verkleurmannetjie* **[N. kameleon, E. chameleon],** *suurlemoen* **[N. citroen, E. lemon].**

—   14200 *unique Dutch items*
      These are items that only occur in Dutch, the counterparts of the unique Afrikaans items.
      Examples: citroen [A. *suurlemoen***], kameleon [A.** *verkleurmannetjie, trapsoetjies***], giraf [A.** *kameelperd,* **E. giraffe].**

As one will have noticed, on a total number of about 45000 items both in Afrikaans and in Dutch, nearly two thirds of them are *cognates,* which means that both languages have a large common core and two exclusive parts which balance each other. However, this common core of cognates does not exclude mutual differences. 'Robot', for instance, is both similar and different in Afrikaans and Dutch. Such cognates are called *partial cognates* or *partial false friends:* they share the same form, have at least one common meaning but also at least one different one.

Absolute false friends, which share form only but no meanings, such as 'amper' [A. = E. nearly ; N. = E. hardly] or neuken/*neuk* [A. = E. beat; N. = E. fuck] also occur. Although there are not many of them (about 500), they are, indeed rather striking because they lead to misunderstandings, therefore in ANNA they have a special marker (!!).

The number of meanings or senses in ANNA amounts to 73619 of which there are:

— 36311 cognate or common ones

— 18049 unique Afrikaans ones

— 19259 unique Dutch ones

The above numbers confirm the image of a large degree of formal and semantic cognateness (2/3) between Afrikaans and Dutch. The greater amount in forms in Afrikaans is balanced by the greater amount of unique meanings in Dutch.

Finally, ANNA contains 90008 examples/combinations from which there are:

— 64196 contrastively relevant

— 25812 not contrastively relevant
  Combinations were considered contrastively relevant

— either when there was a difference in the combination words in both languages;

— or when the translation of the entry word differed from the translation equivalents given;

— or when both these criteria were met.

This number should be interpreted against a functional background: not contrastively relevant combinations are only taken up when they have an illustrative and/or discriminatory function (which is the case with polysemes for instance), contrastively combinations, on the other hand, are always taken up because of their role in the understanding or production process of a foreign language. For examples see the appendix where contrastive combinations are marked by ≠, whereas non contrastive ones are marked by =.

1.3   To start or not to start (a bilingual dictionary project)

An important question at the beginning of any dictionary project and, in particular, a long term and rather expensive one, is whether there are good reasons to start up the project at all.

Understandably, this question was raised at the beginning of the ANNA project and in the pilot study mentioned above. There we came to the conclusion that, as a rule, for a bilingual dictionary project four arguments could be taken into account to start or not to start the project. In the next sections these arguments will be dealt with.

1.3.1   The communicative argument

If (many) speakers of two different language communities for economical, cultural, political or any other reason, often come into contact with each other, then a bilingual dictionary can be called upon as an interlingual instrument[6] to facilitate the communication between the two groups. However, given the fact that speakers of Afrikaans and Dutch can communicate with each other, each of them using his/her mother tongue, be it with the inevitable misunderstandings, miscommunications and problems (see, among others, the false-friends-cases), the communicative argument in itself is *not a sufficient argument*, in the case of Afrikaans and Dutch, to start up such a project[7]. On the other hand, there are at least *three good reasons* that do apply in the case of Afrikaans and Dutch: the functional, the descriptive and the infrastructural argument.

1.3.2   The functional argument

If one accepts that language is a vehicle not only for basic communication, but also one to properly and fully express oneself in, be it in literature, in science or in everyday situations, then a bilingual dictionary is an important instrument to understand/express the subtleties and nuances of the other language, the other culture. Before the appearance of ANNA however, only small or mini-dictionaries existed between Afrikaans and Dutch such as: Dekker and Paarde-kooper (1990), Prisma Miniwoordenboek (2004) and Veltkamp-Visser (1998, 5th edition). So, in order to carry out important language functions such as speaking, understanding, translating and learning the foreign language at an advanced level, a large dictionary was needed and ANNA could fill that gap.

1.3.3   The descriptive argument

With most bilingual dictionary projects the 'gap-in-the-market'-argument suffices. In the case of Afrikaans and Dutch one very specific other argument can be added. Because of the fact that Afrikaans is strongly indebted to Dutch —

up till 1925 Dutch was still an official language in South Africa — an empirically based *confrontation/comparison* with Dutch could *free* Afrikaans and Afrikaans monolingual dictionaries, further from Dutch and Dutchisms. With this I mean words and meanings that still apply in Dutch but no longer do in Afrikaans, such as, for instance, the word 'aardig' that in HAT4 (*Verklarende Handwoordeboek van die Afrikaanse Taal*, Explanatory Desk Dictionary for Afrikaans) still gets as one of its meanings 'aantrekkelik, aangenaam' [E. nice], which has since long passed out of use in Afrikaans where it has been replaced by 'gaaf' or 'nice'.

The paradox of ANNA is that she can make speakers of Afrikaans, by confronting them with Dutch, much more conscious of the existence of 'Dutchisms' (also see Houwelingen and Carstens 1998).

1.3.4   The infrastructural argument

From the beginning ANNA has been set up as a project that should supersede the two languages in question. In other words, the aim was not only to produce a contrastive dictionary Afrikaans–Dutch, but also to lay the foundation for an *exportable model,* one that could be used for other closely related languages, such as the 'black' languages in South-Africa: Xhosa and Zulu, and North-Sotho, South-Sotho and Tswana etc.

The infrastructural argument, i.e. the wish to provide for *dictionary technology* that could lead to a 'new lexicography in South-Africa' (see Martin 2005), thus played an important role in setting up the ANNA project, and within this setup the amalgamation model in its turn took a central position. In the next section we will focus on this aspect by trying to give an answer to the following questions:

What is the amalgamation model?

When can it be used?

What does it look like in actual practice?

2.   The amalgamation model

2.1   What is the amalgamation model?

The amalgamation model is not an explanatory model but a descriptive one. It aims to describe the lexemes of two languages in a bilingual dictionary in a *directly contrastive way*, contrary to the *indirect way* found in a traditional bilingual dictionary. In this sense the model is innovative, a real novelty in lexicography. As a rule however, there is no innovation at will: I had to see a couple of hundreds of Dutch–Afrikaans items made in a traditional way first, to fully understand that this was not the way to proceed and only thereafter I was able

to design the model. Moreover, when I designed it, it was my aim not to draft a model that could apply for Afrikaans and Dutch only, but one that would be *generic*, at least generic for any pair of closely related languages (also see Martin and Gouws 2000 where the model is mentioned for the first time). What this means is explained in the following section.

2.2   Scope of the amalgamation model

As stated above, although the amalgamation model is not bound to the language pair Afrikaans and Dutch, it can only be applied to closely related languages. In order to successfully qualify as a pair of closely related languages the following 'operational' qualitative and quantitative criteria need to be met:

— Qualitative aspects: both the 'form' of the words (spelling) needs to be the 'same' and at least one of the meanings.

— Quantitative aspects: not all words from the two languages need to show the above characteristics; however, in order for the model to be applied successfully, there has to be a sufficient critical mass.

In what follows these characteristics are dealt with in more detail.

A.   'Sameness' of form

'Same' items in this context cover three groups:

— Items with a fully identical spelling form both in Afrikaans and Dutch such as 'tafel' [E. table].

— Items with a small, systematic spelling or morphological difference in the two languages such as N. 'zalm'/A. 'salm' [E. salmon] or N. 'ontsnappen'/A. 'ontsnap' [E. escape].

— Items with a bigger, non-systematic difference between Afrikaans and Dutch, but which are still recognizably similar in form, such as N. 'gel'/A. 'jel' [E. gel] or N. 'pompoen'/A. 'pampoen' [E. pumpkin] or N. 'pinguïn'/A. 'pikkewyn' [E. penguin].

Although this last group may give rise to different interpretations, the pragmatic approach that was followed in ANNA did not lead to dramatic difficulties (see further under References 3.2.1).

B.   'Cognates' versus 'non-cognates'

Words that are considered the same are amalgamated, i.e. treated together in

one entry. They are called *cognates. Non-cognates* on the other hand, such as N. 'kameleon'/A. 'verkleurmannetjie' [E. chameleon] are treated as in any traditional bilingual dictionary, which means that they are treated separately, in two entries.

C.   Sharing (at least one) meaning

In order to qualify as a cognate, and so to be treated in one entry, the items in question must also meet one semantic criterion: they need to share at least one meaning.

In other words, sameness/similarity of form is a necessary, but not a sufficient feature to qualify for amalgamation. Consequently, absolute false friends, which only share form, no meaning, such as 'amper', meaning 'hardly' in Dutch and 'almost' in Afrikaans, will be two entries. 'Robot' on the other hand, with one common meaning and one which applies to Afrikaans only (see above), will be treated as a cognate and will, as such, be dealt with in one entry.

D.   Critical mass

To apply the amalgamation model 'successfully' there needs to be a sufficient critical mass of cognates. Of course, 'success' is a relative concept. Yet one could argue that the degree of (relative) success correlates directly with the number of cognates there is to be found between the two languages. If this number is (much) smaller than half of the total number of words to be described, the degree of amalgamation will be smaller than 0,5 too and the smaller this coefficient the less the dictionary will differ from a 'normal', traditional, dictionary.

In ANNA the cognates/non-cognates ratio is about 50% as there are:

> 59051 dictionary entries, of which there are
> 29840 cognates
> 15011 unique Afrikaans items
> 14200 unique Dutch items

The above numbers make clear that there is an overlap between Afrikaans and Dutch of about 2/3. This seems to satisfy sufficiently the 'critical mass' criterion.

2.3   Illustration of the model: macro and micro

Instead of the traditional double macrostructure distributed over two volumes (A-B, B-A), the *macrostructures* of the two languages in an amalgamated bilingual dictionary (A *and* B) become *unified,* combined as one whole. So, for instance, in ANNA the section 'lekker-lekkerte' (E. tasty-tastiness) looks as fol-

lows (Dutch items are in roman, Afrikaans ones in italics, followed by English translations between brackets):

lekker, *lekker*	(nice, tasty)
lekkerbek, *lekkerbek*	(gourmet)
lekkerbekje, *lekkerbek*	(fried fillet of haddock)
lekkerbekkig	(finicky)
lekkergoed	(sweets)
lekkerig	(who likes to lick)
lekkerkry	(pleasure)
lekkerlyf	(squiffy)
lekkernij, *lekkerny*	(delicacy)
lekkerruik	(scented)
lekkers	(sweets)
lekkerte	(nice thing)

This unification allows not only for a direct illustration of morphological similarities and differences but also for the generation of hypotheses at this level.

The micro-amalgamation reveals itself in the cognates or amalgamated items. To clearly indicate both differences and similarities at the semantic as well as at the combinatorial level, use is made of the following markers:

A/N	indicates a common meaning
N	indicates a meaning which occurs in Dutch only
A	indicates a meaning which occurs in Afrikaans only
≠	marks a combination or example which shows contrast
=	marks a combination or example which shows no contrast

The following somewhat simplified ANNA-entry can illustrate how this system works.

WERF

A/N (werkplaats voor schepen) *skeepswerf,* [m.g.] *werf*
=   op de werf *op die skeepswerf* ≠ een schip van de werf laten lopen *'n skip te water laat*
N   ([BN] bouwterrein) *bouterrein*
A   (*oop stuk grond rondom die huis*)   erf
≠   *'n pragtig geleë plaashuis met werf en tuin* een fraai gelegen woonboerderij met erf en tuin; *'n motor het die werf opgery* een auto reed het erf op; (fig.) *elke mens moet sy eie werf skoonhou* ieder moet zijn eigen tuintje wieden

[The three meanings are: shipyard; building site; yard. The translation of the examples reads: on the shipyard; launch a ship; a nice farmstead with yard and garden; a car drove into the premises; everyone has to clean up one's own backyard.]

For more elaborated examples of entry words I refer to the *appendix*. From

the data presented there it becomes clear that, actually, there are three kinds of cognates: *absolute cognates*, *absolute cognates with form difference* and *partial cognates.*

Absolute cognates are words which are completely identical in Afrikaans and Dutch, both in form and in meaning, such as is the case with, for instance, opwinding/*opwinding.* However, as one can observe, this does not mean that these items are used in completely the same way, neither collocationally nor pragmatically.

Absolute cognates with form difference are, in fact, a subset of the absolute cognate-class. They have the same meaning(s) in both languages and the form difference they show is small enough to be (easily) recognizable. See, for example, hartinfarkt/*hartinfark,* stikken/*stik.*

Partial cognates are words which share the 'same' form having at least one meaning in common and one meaning that differs. See, for example, taai/*taai* and geil/*geil.*

In a way these *partial cognates* are also *partial false friends* (also see above, section 1.2.3). Contrary to *absolute false friends* such as stoep1 and *stoep2* or *lemoen* and limoen, which only share forms, no meanings, they share forms and *part of their meanings.*

A last type of items is the *non-cognates:* they have the same meaning but a form which is clearly different. *Non-cognates*, as in any traditional bilingual dictionary, are treated as separate entries. See, for instance, appelsien, sinaas-appel, and *lemoen.*

3.   Amalgamation: pros, cons and pitfalls

3.1   Advantages

Two of the advantages of the model have already been mentioned in section 1, I briefly repeat them here.

a.   The amalgamation facilitates the direct comparison between the two languages.

Differences and similarities at the level of orthography, morphology, semantics, pragmatics and combinatorics become, so to speak, clear immediately and not through the combination of two volumes.

In the above case of 'werf', for instance, all meanings are brought together, yielding a global overview, whereas in a traditional treatment they would be distributed over the two volumes as follows:

—   meanings 1 and 2 would appear in the Dutch–Afrikaans part
—   meanings 1 and 3 would appear in the Afrikaans–Dutch part

In this respect it is interesting to refer to Reneé Marais, who, in a review of ANNA, points at the advantages this approach offers in the context of lan-

guage learning. Marais writes the following:

> The fact that the two languages in this dictionary are mixed is very convenient. The contrastive approach yields many useful data. One learns a non-related language from scratch, from *tabula rasa,* from point zero. Although learning a closely related language also implies that one simply has to learn certain grammatical aspects and unknown lexical items, one, especially, has to be conscious of similarities, small differences and false friends. (Marais 2011: 191; my English)

b.   The amalgamation *reduces redundancy*.

In the case of 'werf', for instance, the first meaning, shared by both languages, needs only to be mentioned once.

In addition to these 'known' advantages, I will in what follows, elaborate on two other advantages:

c.   In an amalgamated bilingual dictionary it is easier to detect *contrastive patterns* (at morphological, semantic or combinatorial level) than in a traditional bilingual dictionary. These patterns then can subsequently serve as *hypothesis generators.*

One can, for instance, wonder whether there are different systematic patterns for meaning extensions for certain classes of words. Compare, for instance, 'N. woest' to 'A. woest' where the 'desolate' meaning applicable to landscapes in both languages, 'extends' in Dutch to an attribute applicable to 'people', whereas in Afrikaans it applies to 'situations'. Or think of differences in figurative extensions. In this respect one can test more general hypotheses concerning stereotypes or opinions about the two languages. Dutch speakers, for instance, consider Afrikaans as a cute, charming language, richer in imagery than Dutch. However, speakers of Afrikaans, when asked about the 'richness' of Dutch, also consider Dutch very rich in metaphors and images. The fact that figurative expressions are marked for both languages and brought together can help in 'objectifying' these opinions. In the entry 'muur' (E. wall), for instance, figurative expressions in Dutch appear (such as: 'geld uit de muur halen', literally: get money out of the wall, meaning: 'get money from a vending machine') which have no figurative counterpart in Afrikaans and vice versa (for instance: '*oor die muur wees'*, literally*:* to be beyond the wall, meaning: to be worn-out) (on figurative expressions in Afrikaans and Dutch, also see Swanepoel 1997).

In other words, often, by leaving one's own frame of reference, one can better see oneself. ANNA, in confronting the two languages from both angles, can certainly offer data to test certain claims and beliefs in more detail.

d.   As always, however, the proof of the pudding is in the eating. The fact that there is now an example of amalgamation makes it possible to *evaluate the model*. This not only allows us to stress the advantages of the

model but also to detect certain side effects and/or pitfalls, two phenomena I will deal with in the next section.

3.2   System side effects

Just like any system, the amalgamation model shows certain side effects due to the system itself. These, however, need not be harmful to the user.

3.2.1   References

The fact that cognates are treated together raises the question where to deal with those cognates which show a form difference. In other words: in which entry will the user have access to items such as gel/*jel* or *pikkewyn*/pinguïn? Will he/she find this information under 'gel' or under 'jel', under 'pikkewyn' or under 'pinguïn'?

As these are combined entries or cognates with difference in form, there must be a *preferred* first item to guarantee access. This problem is typical for paper dictionaries as in an electronic version (e.g. the CD-ROM version of ANNA) the user can access the data by means of both language forms.

In the paper version we have chosen a representation in which the Dutch form is followed by the Afrikaans one, for instance:

drijven, *dryf*	[float]
drijvend, *drywend*	[floating]
drijver, *drywer*	[driver, drover]
drillen, *dril*	[drill]
dringen, *dring*	[push] etc.

To help the user who wants to consult these cognate entries via the Afrikaans form, a pragmatically functioning reference system had to be worked out. To put it briefly: if, in a purely alphabetical ordering, the Afrikaans item would precede or follow the Dutch item immediately or occur in the immediate neighbourhood, then no reference is included. If, however, the Afrikaans and Dutch item would be separated by more than seven items, then a reference is included, leading the user from the Afrikaans to the Dutch form. As a consequence, no reference is made from '*dring*' to 'dringen', nor from '*dril*' to 'drillen' (those items would follow each other). Reference is, however, made from '*dryf*' to 'drijven', from '*drywend*' to 'drijvend' and from '*drywer*' to 'drijver', as these items are too far removed from each other.

ANNA contains about 5000 such references, always going from Afrikaans to Dutch. In this way users who want to look up a cognate word using the Afrikaans form will find what they are looking for, not being hindered by too many references.

3.2.2   Meaning order

With polysemous cognates the meaning order is not defined by, for instance, frequency of occurrence of meanings, but by the 'system' itself. Indeed it is the amalgamation system that in the case of polysemous cognates 'dictates' the following order:

> The meanings Afrikaans and Dutch have in common precede
> those which are exclusively Dutch, if any, which in their turn precede
> those which are exclusively Afrikaans, if any.

See, for instance, the 'werf'-example in section 2.3 and the 'taai'-example in the appendix.

Such a 'predefined' path is needed if one wants to guide the user through the semantic wood of an amalgamated dictionary. Of course, this path will not necessarily coincide with the meaning order in an (Afrikaans) monolingual dictionary based on frequency. For instance, in ANNA, one will find that in the entry 'pad' [E. path] the 'small road'-reading will precede the 'road'-reading (which is by far the most frequent in Afrikaans) because of the fact that the former is the reading Dutch and Afrikaans share. Other examples are: 'robot', where the 'automaton'-reading precedes the 'traffic light'-reading, 'kar', where the 'vehicle on two wheels'- precedes the 'car'-reading, 'brander', where the 'apparatus'- precedes the 'wave'-reading etc., etc.

Consequently, the ANNA-user has to be conscious of the fact that he/she is consulting a bilingual contrastive dictionary with a logic of its own regarding the order in polysemous cognates.

3.2.3   The place of cognate equivalents in macro and micro

If an Afrikaans item has a Dutch cognate equivalent, then, even if there is another more usual equivalent in Afrikaans, it is the cognate equivalent that will be put in the macro. For instance, one will find as a macro-entry the 'head' taxateur/*taksateur* although in Afrikaans the more common word for 'taxateur' [E. appraiser, valuer] is *waardeerder.* However, in the microstructure the most frequent equivalent will be put first or, alternatively, the less frequently used (but cognate) equivalent will be given a restrictive label, such as, for instance, [formal], in the case of *taksateur.*

3.2.4   Pitfalls

More serious perhaps than the system side effects just mentioned, which, after all, can be 'overcome' within the system, are, what I will call, pitfalls. What I am referring to are cases where there is a meaning overlap or minor meaning/ usage differences between Afrikaans and Dutch. As a rule, in ANNA, we have

chosen to abstract away from these differences, in other words, to 'lump', not to 'split', quite in line with the amalgamation approach itself. This does not mean, however, that these differences have been ignored: although these 'differences' are treated under 'common' meanings, use is made of examples, pragmatic labels and/or comments to make these often more subtle differences clear.

In ANNA, for instance, the Dutch item 'tergen' with the meaning 'sarren' [E. provoke] and the Afrikaans item 'terg' with the meaning 'pla' [E. tease] are considered cognates rather than false friends, and, consequently, they are treated under one common *A*/N-meaning and not in separate entries. However, the following comments are added (in Afrikaans, I translate them here into English):

> Although 'terg' in Afrikaans can also mean 'provoke', compared to Dutch, its meaning is less strong mostly having a jocular rather than a nasty connotation.

The difference between the two items is also made clear by means of examples. For instance: N. iemand op alle mogelijke manieren tergen [E. provoke s.o. in all possible ways] A. *iemand op alle moontlike maniere tart;* A. *graag terg* [E. like to tease] N. graag plagen.

Some other examples:

— rukken/*ruk*
   In Afrikaans as well as in Dutch this verb can be used both transitively and intransitively. In the latter case, Dutch needs to use a prepositional object as a complement. For instance: 'aan iets rukken' [E. to pull at something]. In Afrikaans this needs not be the case. The verb *ruk* can also be used without a prepositional object, meaning 'to make a jerk, moving with jolts, as if pulled'. This specific Afrikaans usage is not treated in a different entry nor as a different sense, but by means of the example *die vliegtuig ruk* (N. het vliegtuig schokt, E. the air plane is shaking/jolting).

— aanbieder/*aanbieder*
   Afrikaans and Dutch share the common general meaning 'someone who offers something, a provider'. However, depending on the context, this word can get a more specific semantic load in Afrikaans. For instance: the *aanbieder* of a TV-programme is called a 'presentator' in Dutch [E. presenter, host], the *aanbieder* of a course is a 'cursusleider' in Dutch [E. course instructor]. As one can observe, here too, the more specific usages in Afrikaans are accommodated under the more general, common meaning and the differences between Afrikaans and Dutch are made clear by means of examples.

— ontkennen/*ontken*
   In this case, depending on the language, different semantic actants are used with the same item, without changing the basic meaning:

	Affected Object
N. ontkennen	iets
A. *ontken*	*iets, iemand*

If the Affected Object is [-human], then the semantic load is 'deny'.
If the Affected Object is [+human], then the semantic load is 'deny/ refuse to recognise'.

Onthouden/*onthou* is a similar case: here again the general meaning of 'keep in memory' if applied to a human object is specified to 'remember someone'. Such cases are 'lumped' in ANNA, the difference made clear by means of examples.

	Affected Object
N. onthouden	iets
A. *onthou*	*iets, iemand*

To conclude: it is important for users to be aware of the fact that the amalgamation approach entails the lumping of meanings rather than the splitting of them when small, more subtle differences are at stake. In ANNA we have tried to deal with these cases in a sensible way, although it is quite well possible that in this approach some differences have been overlooked or could be dealt with in a more explicit way.

4.   Conclusion

In this article I have presented ANNA and the underlying amalgamation model. In addition, I have provided a first evaluation of the model. As for its applicability to other languages, within the ANNA-project itself the focus was on the application possibilities for so-called 'black' languages in South Africa, where it seems to be promising.

Whether the model is also applicable to other 'black' languages such as, for instance, Kirundi and Kinyarwanda spoken in, respectively Burundi and Ruanda, or Scandinavian and Slavic ones, like Danish and Norwegian or Russian and Byelorussian, is up to the specialists in the field to find out, although it seems at first sight, that his could be the case[8].

If this proves true then between these languages a new type of *contrastive dictionary,* comparable to ANNA, could emerge.

However, there is more. As I pointed out in the *Introduction* to the ANNA dictionary (p. 11, my translation into English):

> Last but not least, an amalgamated dictionary is also a *reading dictionary:* a dictionary in which one cannot only look for translations or contrasts, but one in which one can also read and browse. Anyone looking in ANNA for N. 'bloot' [E. naked] for instance, will not only find that the most common equivalent in Afrikaans is *kaal*, but ANNA will take him further from 'bloot' to N. 'blootshoofds' [A. *kaalkop*, E. bald] and from N.' blootshoofds' to A. *bloots ry* [N. het (iem.) lastig

maken, E. give (s.o.) a dressing-down], and from A. *bloots ry* to N. 'blootsvoets' [A. *kaalvoet,* E. bare-footed] etc. In other words, before you know it, ANNA takes you with her from word to word to word and that is exactly where she has quite a lot to offer.

It therefore seems justified to conclude that if ANNA could inspire others, to new, more adequate, types of bilingual dictionaries, be it amalgamated, contrastive, reading or some other type, I would feel more than rewarded for the effort made by the ANNA-team in realizing this project.

Notes

1. In a sense monolingual diachronic dictionaries also show a kind of amalgamation as they bring 'same items' from different time periods together. However, here all similarity with the amalgamation model stops as the latter focuses on synchronic similarities and differences of two different languages and not on the evolution of lexical items through time.

2. ANNA has been financed mainly by private sponsors, the main sponsor being the ZASM foundation (ZASM = Zuid-Afrikaanse Spoorwegmaatschappij (South African Railway Company)). Other sponsors were the PUK Vice-Chancellor's Trust of the North West University (Potchefstroom), the University of Stellenbosch, The Dutch Language Union, The department of Foreign Affairs of Flanders, the Van den Bergh van Heemstede Foundation, the Joan Louw Trust, The LW Hiemstra Trust and the Prince Bernhard Cultural Foundation.

3. For more information on the RBN, see Van der Vliet 2007; for OMBI, see Maks 2007.

4. The ANNA-team consisted of nine members. The editorial team itself consisted of five members: E. Boekkooi (Port Elizabeth), R. Gouws (Stellenbosch), I. Maks (Amsterdam), L. Renders (Hasselt) and myself acting as an editor-in-chief and project leader.

5. The CD-ROM version has been produced by Pharos Publishers, who also distribute the South African version of ANNA.

6. Of course also a 'foreign' language can be used as an 'interlingua'.

7. It goes without saying that any project can be started up if there are no financial constraints. In most cases however one has to prioritise because of financial limitations.

8. Notice that next to a certain degree of 'sameness', the two languages should also show a certain degree of 'difference'. Cognates should, for instance, show sufficient differences at combinatorial, pragmatic and/or semantic level.

References

Dekker, L. en P.C. Paardekooper. 1990. *Nederlands–Afrikaanse Woordeboek.* Pretoria: J.L. van Schaik.

Groot Woordenboek Afrikaans en Nederlands. 2011. Martin, W. (Editor-in-chief). 2011. *Groot Woordenboek Afrikaans en Nederlands*. Houten/Antwerpen: Het Spectrum B.V.

HAT. 2000. Odendal, F.F. and R.H. Gouws. 2000. *HAT: Verklarende Handwoordeboek van die Afrikaanse Taal*. Cape Town: Perskor.

Maks, I. 2007. OMBI: The Practice of Reversing Dictionaries. *International Journal of Lexicography* 20(3): 259-274.

Marais, R. 2011. Een mooie dikke dame. "ANNA", het Groot Woordenboek Afrikaans en Neder-lands. *Ons Erfdeel*, nr. 1: 190-192.

Martin, W. 2005. ANNA and a 'New' Lexicography in South Africa. Schutte, G.J. and H. Wels (Eds.). 2005. *The Vrije Universiteit and South Africa. From 1880 to the Present and Towards the Future: Images, Practice and Policies*: 49-57. Amsterdam: Rozenberg Publishers.

Martin, W. 2012. Amalgamated Bilingual Dictionaries. Genis, René et al. (Eds.). 2012. *Between East and West. Festschrift for Wim Honselaar on the Occasion of his 65th Birthday. Pegasus Oost-Euro-pese Studies* 20*: 437-449. Amsterdam: Uitgeverij Pegasus.

Martin, W. To appear. ANNA:een woordenboek met een naam en wat er achter steekt. *Tydskrif vir Nederlands en Afrikaans*. 2012.

Martin, W., R. Gouws and L. Renders. 1999. *Haalbaarheids- en Definitiestudie voor een Woordenboek Afrikaans–Nederlands/Nederlands–Afrikaans.* Amsterdam: Vrije Universiteit.

Martin, W. and R. Gouws. 2000. A New Dictionary Model for Closely Related Languages: The Dutch–Afrikaans Dictionary Project as a Case-in-point. Heid, Ulrich et al. (Eds.). 2000. *Pro-ceedings of the Ninth EURALEX International Congress. EURALEX 2000. Stuttgart, Germany, August 8th–12th, 2000*: 783-792. Stuttgart: Stuttgart University.

Prisma Miniwoordenboek Afrikaans–Nederlands/Nederlands–Afrikaans. 2004. Utrecht: Het Spectrum.

Swanepoel, P. 1997. Idiomatiese taalgebruik en die afstand tussen Afrikaans en Nederlands. *Tyd-skrif vir Nederlands en Afrikaans* 4(1): 128-161.

Van der Vliet, H. 2007. The Referentiebestand Nederlands as a Multi-purpose Lexical Database. *International Journal of Lexicography* 20(3): 239-257.

Van Houwelingen, F. and A. Carstens. 1998. "Nederlandismes" in HAT. *Literator* 19(2): 1-12.

Veltkamp-Visser, S. 1998. *Afrikaans op reis:* taalgids voor de Nederlandssprekende toerist in Zuid-Afrika. Fifth Edition. Amsterdam: NZAV.

Appendix: ANNA

appelsien [nw.] [de; mv: -en]

N   (<BN> sinaasappel) *lemoen*

geil [bnw.] , ***geil*** [b.nw.] {*Opm.: In Afr. het 'geil' 'n ruimer betekenis as in Ned. en beteken nie alleen paarlustig nie.*}

A/N   (<inf.> met een zeer sterke geslachtsdrift) *jags, bronstig, katools, geil* = geile gedachten
       jagse gedagtes ≠ een geile bok *'n bronstige bokram, 'n geil bokram*; geil zijn op iemand
       katools wees oor iemand; een geile griet *'n katoolse meisiekind*

A   (*welig, vrugbaar*) welig, vruchtbaar ≠ *geil grond* vruchtbare/vette grond; *geil groei* welig
       tieren; *'n geil jaar* een jaar met een goede oogst

hartinfarct [nw.] [het; mv: -en] , ***hartinfark*** [nw.] [mv: -e] {*Opm.: In Afr. kom 'hartinfark' uitsluitlik in die mediese register voor.*}

A/N   (het afsterven v.d. hartspier) *hartversaking*, <form.> *hartinfark* ≠ roken verhoogt ook de
       kans op een hartinfarct *rook verhoog ook die kans op hartversaking*; een hartinfarct krijgen
       hartversaking kry

lemoen [nw.] [mv: -e] {*!! Opm.: Sien ook Ned. 'limoen'.*}

A   (*oranje/goudgeel sitrusvrug*) sinaasappel, <BN> appelsien ≠ *'n lemoen skil/eet* een
       sinaasappel pellen/eten ≠ <fig.> *weggegooi word soos 'n uitgesuigde lemoen* wegge-
       worpen worden als een uitgeknepen citroen; <fig.> *iemand uitsuig soos 'n lemoen* iemand
       als een citroen uitknijpen

opwinding [nw.] [de; mv: -] , ***opwinding*** [nw.] [mv: -]

A/N   (spanning) *opwinding, opgewondenheid* = er heerst (grote) opwinding *daar heers (groot)
       opwinding*; zorgen voor (de nodige) opwinding *sorg vir (die nodige) opwinding*; grote/
       enorme opwinding *groot/enorme opwinding* ≠ van opwinding niet meer kunnen slapen *van
       opwinding nie meer kan slaap nie*; er was veel opwinding om de nieuwe auto *daar was
       groot opwinding oor die nuwe motor*

sinaasappel [nw.] [de; mv: -s, sinaasappelen]

N   (sappige zuidvrucht) *lemoen* ≠ een sinaasappel pellen/eten *'n lemoen skil/eet*

stik zie stikken

stikken₁ [ww.intr.] , ***stik*** [ww.intr.]

A/N   (door ademnood sterven) *stik* = <fig., inf.> stikken van de warmte/het lachen/woede/jaloe-
       zie *stik van die warmte/die lag/woede/jaloesie* ≠ <inf.> stik! *gaan bars!, bokker jou!*; stik-
       ken door de rook *stik van die rook*; <inf.> iemand laten stikken *iemand kwaai in die steek
       laat* ≠ <fig., inf.> stikken in het geld *stink van die geld*; <fig., inf.> het stikt hier van de sol-
       daten/muggen/kroegen *dit wemel hier van die soldate/muskiete/kroeë*

stikken₂ [ww.tr.] , *stik* [ww.tr.]

*A*N   (naaien met eenvoudige steek) *stik* = een zoom in een broekspijp stikken *'n soom in 'n broekspyp stik*

stoep₁ [nw.] [de; mv: -en] {*!!*}

N   (strook langs weg voor voetgangers) *sypaadjie* ≠ denk erom, op de stoep blijven met je step! *onthou, bly op die sypaadjie met jou skopfiets!* ≠ <fig.> bij iemand op de stoep staan *voor iemand se deur staan*; <fig.> stoepetje/stoepje spelen *'n tipe balspeletjie op straat speel*

stoep₂ [nw.] [mv: -e] {*!!*}

A   (*verhoogde vloerarea buite huis, veranda*) veranda ≠ *aan die voorkant was 'n stoep wat met wingerd omrank was waar 'n mens heerlik kon sit* aan de voorzijde was een met wingerd omrankte veranda waar het heerlijk zitten was ≠ <fig.> *(eers) voor jou eie stoep vee* (eerst) je eigen straatje schoonvegen

taai [bnw.] , *taai* [b.nw.]

*A*N   1 (stug) *taai* = dat vlees is bijzonder taai *daardie vleis is besonder taai* ≠ <fig.> zo taai als een schoenzool zijn *so taai soos (skoen)leer wees* ≠ <fig.> *'n (taai) turksvy* een heet hangijzer; <fig.> *taai in die bek wees* hardleers zijn, weerspannig zijn; <fig.> *taai gesprekke/onderhandelinge* moeizame gesprekken/onderhandelingen 2 (volhardend, sterk) *taai* = je taai houden *jou taai hou* ≠ een taaie oude man *'n taai ou man*, *'n seningtaai ou man*; houd je taai! *uithou en aanhou!, vasbyt!* ≠ *jou taai hou* je kras houden ≠ <fig.> *'n taai klap* een flinke optater; <fig.> *so taai soos 'n ratel wees* uitzonderlijk taai zijn

N   (vervelend) *moeilik* = taaie leerstof *moeilike leerstof*; een taaie roman *'n moeilike roman* ≠ een taaie beklimming *'n moeilike klim*

A   (*klewerig*) kleverig, plakkerig, klef = *taai deeg* plakkerig deeg; *taai vingers* plakkerige vingers; *taai hande* kleverige handen, kleffe handen; *taai brood/gebak* klef brood/gebak ≠ *'n taai spul(letjie)* een kleverig goedje