Our starting point is that translation equivalents: true friends in their use as individual lexical items often become false friends in collocations. It is the duty of the lexicographer to guide the user, especially in learners’ dictionaries aimed at productive-encoding-use, in forming correct collocations and in warning the user of false friend cases. Our arguments are based on evidence from large newspaper corpora as well as on internet research. We will present several lexicographic presentation devices from printed dictionaries that allow lexicographers to warn users about false friend collocations. The study will be limited to false friend relations in general bilingual dictionaries, mainly for German, Dutch and Afrikaans. The compilation of dictionaries for false friends lies beyond the scope of this paper. We adopt a lexicographic notion of collocation, here, as used for example by the Oxford Collocation Dictionary for Students of English (2002). We use Hausmann’s (2004) terms-base and collocate-to denote the elements of collocations. Klégr (2006) transfers the notion of false friends from single words to collocations and classifies the relevant cases according to categories known from translation theory. We propose the following simple arrangement of false friend collocations, inspired by the concept’s basic principles:

- word combination-lexical (co-)selection: if true friend single word equivalents exist in a language pair, we consider collocations as false friends where the cooccurrence of the two single word true friends is impossible in a given language;
- morphosyntactic preferences: if true friend single word equivalents exist in a language pair, we consider collocations false friends where the languages differ with respect to morphosyntactic preferences, individual readings being equivalent;
- differences with respect to usage domains.

1. Introduction

1.1. Objectives

Our starting point is that translation equivalents which are true friends in their use as individual lexical items often become false friends in collocations. It is the duty of the lexicographer to guide the user, especially in learners’ dictionaries aimed at productive (encoding) use towards the formation of correct collocations and to warn the user in false friend cases. Our arguments are based upon evidence form large newspaper corpora and on internet searches. A few devices of lexicographic presentation in printed dictionaries that would allow lexicographers to warn users against collocational false friends will be presented. The study will be limited to false friend relations in bilingual general dictionaries, mainly for German, Dutch and Afrikaans. The compilation of dictionaries for false friends lies beyond the scope of this paper.

1.2. Corpus evidence used

The Afrikaans corpus utilised for this study is the Media 24 archive. This archive is estimated at 3,000,000,000 tokens. The Dutch corpora consulted are a section of the synchronic corpus of the VNC-project, “Convergentie en Divergentie in de Nederlandse woordenschat” jointly owned by the Universities of Leuven and Gent, V.U. Amsterdam and I.N.L. Leiden, and also the Twente
Nieuws Corpus. The German corpora consulted are a collection of newspaper texts from 
Stuttgarter Zeitung, Frankfurter Rundschau and die tageszeitung (1987 to 1993), as well as 
Handelsblatt, Frankfurter Allgemeine Zeitung, and dpa news services. The Dutch and German 
corpora contain more than 500,000,000 tokens of each language. Searches on the Internet using 
the Google search engine were also done.

2. State of the art

2.1. Collocation

We adopt a lexicographic notion of collocation here, as used for example by the Oxford 
and collocate) to denote the elements of collocations. Our view on the phenomena at stake is 
best summarized with the following quote from Bartsch (2004: 76):

Collocations are lexically and/or pragmatically constrained recurrent cooccurrences of at 
least two lexical items which are in a direct syntactic relation with each other.

In other words, the following criteria are particularly relevant for the notion of collocation 
adopted here:

collocations are a phenomenon of lexical selection: the base selects its collocate;

collocations show a grammatical relation between the two elements, e.g. verb + object, verb + 
subject, attributive adjective + noun, etc.;

collocations may show signs of pragmatic and/or (morpho-) syntactic idiomatization.

All these aspects are relevant within each single language. Each of them or several together may 
complicate translation.

2.2. False friends and dictionaries of false friends

Many translation equivalents between closely related languages are “true friends”, in the sense 
of having both a closely related form and a closely related meaning. True friends between 
Afrikaans and Dutch for example include words such as en (and), in (in) op (on) maar (but) all 
occurring thousands of times in both the Afrikaans and Dutch corpora utilised for this study. 
False friends, on the contrary, are words with closely related form, but diverging meaning.

A typical interpretation of the term false friends refers to two lexical items from different 
languages with the same form but different meanings. (Gouws, Prinsloo and De Schryver 
2004: 798).

False friends constitute pitfalls for language production as well as reception. We cannot go into 
exhaustive detail here, as far as ways of treating false friends in existing dictionaries are 
concerned, but we will in the following point to a few relevant devices currently used in 
lexicography.

The monolingual Cambridge International Dictionary of English (which has the comparative 
aspect already in its title) contains hints to false friends at lemma form vs. meaning level: it has 
symbols to indicate false friends in other languages. So, s.v. genie (“magical spirit”), it warns 
speakers of German and French (who have Genie, génie (“genius”)), without however 
indicating the words of the other languages.

Many bilingual dictionaries treat false friends in the same way as all other equivalents: they do 
not have specific devices to warn against false friends. Typical dictionaries of false friends tend 
to cover bilingual word pairs with different uses, often with translated example sentences. 
Dictionnaire des faux amis français-anglais has, for example, an entry for FR maniaque/EN 
maniac, with examples such as “Patricia est une maniaque de la propreté”/“Patricia is fanatical 
about cleanliness”. Such dictionaries focus on the meaning differences of single words, but 
rarely address collocations.
2.3. Collocational false friends

Klégr and Saldová (2006) transfers the notion of false friends from single words to collocations and classifies the relevant cases according to classes known from translation theory. Examples include transpositional multiword equivalent pairs (e.g. Verb → Verb + Adverb, srde vynechava ↔ the heart beats irregularly), closely related but not identical multiwords (protancovat boty ↔ dance the soles off (Czech lit: “dance the shoes off(“)) and differences in number (prolomit ledy (plural) ↔ break the ice (singular)). Nuccorini (2006) analyses verb-adverb collocations, focusing on adverbs like English (EN) totally, absolutely, entirely and their formally related Italian counterparts totalmente, assolutamente, interamente. She notes the differences in use between the adverbial equivalents which make, for example, interamente a mistranslation of entirely in negative contexts (entirely useless etc., p. 42).

3. A simple typology of collocational false friends

Both Klégr and Saldová, and Nuccorini address aspects of the false friend behaviour of collocations, but they do not, in our view, address all aspects, mainly because their focus is different from our one. We start from a lexicographic description of collocations as (second order) treatment units, i.e. an approach where collocations may receive the same detailed description as single words: they may require a definition or a meaning explanation (especially semantically opaque, idiomatized ones), and they have properties, such as morphosyntactic preferences (e.g. for number), specific syntactic (valency) constructions, or marks of style, register, regional or time-related usage specificities. And collocations may have relations with other lexical items, e.g. single word verbs of which a given collocation may be a synonym, an autonym or a hyponym.

The properties mentioned, as well as the basic aspects listed above (cf. section 2.1) are language specific. If two languages are compared, as it is the case in bilingual dictionaries, the comparison should account for each of the properties mentioned. A complete true friend translation relation is then one where not only the form of the lexical items involved in the collocations of both languages is closely related, but where also the collocational meanings and all relevant properties are parallel. Between related languages, such cases are not rare. Examples are given in table (1), below, for Afrikaans, Dutch and German.

As users of dictionaries for related languages may assume that this is the standard situation, it is all the more necessary to warn them, if two languages differ, in the equivalence of a given pair of collocations, with respect to one of the aspects mentioned above: lexical selection, (morpho)syntactic preferences, usage marks.

3.1. Overview

We thus propose the following simple classification of collocational false friends inspired by these basic aspects of collocations:

word combination (lexical (co-)selection): if true friend single word equivalents exist in a language pair, we consider collocations as false friends where the cooccurrence of the two single word true friends is impossible in a given language. This is Nuccorini’s case entirely useless ↔ *interamente inutile;

morphosyntactic preferences: if true friend single word equivalents exist in a language pair, we consider collocations as false friends where the languages differ with respect to morphosyntactic preferences, individual readings being equivalent: German den Rechtsweg einschlagen must be in the singular, whereas Italian prendere la vie legali has to be in the plural; cf. Klégr and Saldová’s example break the ice;

differences with respect to usage domains: if true friend single word equivalents exist in a language pair, we consider collocations as false friends where the languages differ with respect to usage: unmarked vs. marked, frequent vs. rare, general vs. domain-specific, etc. German (DE) Angst haben is unmarked and very frequent, but Afrikaans (AF) aang hê is extremely rare; on the other hand, Dutch (NL) angst ervaren (1330 Google matches) is not marked, whereas DE
Angst erfahren is mainly used in psychology texts (797 Google matches against more than 1 million for Angst haben).

We will now discuss each of the above three levels in more detail.

3.2. Collocational false friends at the lexical level

The most frequent case is that of false friends at the lexical level; this even occurs between very closely related languages or regional varieties, such as continental Portuguese and Brazilian Portuguese: apanhar um susto ("get a fright", literally: "pick a fright") is typical of Portugal (68 occurrences in 200 M words of the Cetempublico corpus), whereas this combination does not show up at all in our 14 M words Brazilian corpus. The Brazilian Portuguese collocation levar um susto (23 occ. in 14 M, lit. "raise a fright") is not at all found in Cetempublico (thanks to Heike Stadler for the data, p.c.).

As collocational combinatorics is often more a matter of preference than of categorical decisions, lexical false friendship is sometimes also a matter of preference: German has in Panik geraten ("to panic, to run scared, get into panic", 193 in 200 M words), and Dutch has a similar in paniek raken (233 in 63 M words of Volkskrant); in our Afrikaans corpus, in paniek raak (3 in 750 M) is absolutely marginal, against a frequent verb + adjective construction, “paniekerig raak” (782 in 750 M). Conversely, our Dutch corpora have no single occurrence of paniekerig raken, with a Google search rendering only 4 hits, even though both AF and NL have this adjective. A lexicographer who knows (or has the intuition) that AF in paniek raak is not excluded, should nevertheless account for the quantitative distribution. Table 1 contains more such cases.

<table>
<thead>
<tr>
<th>EN equivalent</th>
<th>AF (Media 24 Archive)/ Google</th>
<th>NL (Google search)</th>
<th>DE (Google search)</th>
</tr>
</thead>
<tbody>
<tr>
<td>report exists</td>
<td>berig aanwesig (0) // bericht (0)</td>
<td>aanwezig (1550**)</td>
<td>Bericht anwesend (9)</td>
</tr>
<tr>
<td>present, submit report</td>
<td>berig voorlê (0) // berig voorgêlê (0)</td>
<td>bericht (4) // bericht (29)</td>
<td>Bericht vorlegen (20,900) // Bericht liegt vor (4,570)</td>
</tr>
<tr>
<td>give a report through</td>
<td>berig deurgee (1)/(1) // bericht (13,800)</td>
<td>doorgeven (29) // Bericht durchgeben (48)</td>
<td></td>
</tr>
<tr>
<td>make a report</td>
<td>berig maak (0*)// bericht maken (10,500)</td>
<td>Bericht machan (911)</td>
<td></td>
</tr>
<tr>
<td>write a report</td>
<td>berig skryf (75**)/ bericht skrijven (1,720,000)</td>
<td>Bericht schreiben (2,200,000)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. True/false friend relations in terms of the collocative function of the translation equivalents for exists, give through, make

* context: make a report
** count also includes a few accidental combinations, e.g. “in het bericht aanwezig zijn” (be present in the report)

In table 1 the Afrikaans, Dutch and German words berig/bericht/Bericht “report”, aanwesig/aanwezig/anwesend “to be present”, deurgee/doorgeven/durchgeben and maak/maken/machen “make” are true friends in non-collocational contexts, but false friends in collocational relations. In Dutch bericht aanwezig in the sense of “a message/report exists” is frequent, but it is marginal in German and not found in the Afrikaans Media 24 archive. Consider also table 2:
Section 8. Phraseology and Collocation

<table>
<thead>
<tr>
<th>EN equivalent</th>
<th>AF (Media 24 Archive // Google search)</th>
<th>NL (Google search)</th>
<th>DE (Google search)</th>
</tr>
</thead>
<tbody>
<tr>
<td>feel fear</td>
<td>angs ervaar (47)/(70)</td>
<td>angst ervaren (1330)</td>
<td>Angst erfahren (797)</td>
</tr>
<tr>
<td>have fear</td>
<td>angs hè (2)/(5)</td>
<td>angst hebben (38,700)</td>
<td>Angst haben (1,150,000)</td>
</tr>
<tr>
<td>cause fear</td>
<td>angs maak (0)/(0)</td>
<td>angst maken (521)</td>
<td>Angst machen (250,000)</td>
</tr>
<tr>
<td>be filled with fear</td>
<td>angs bekom (0)/(0)</td>
<td>angst bekomen (147)</td>
<td>Angst bekomen (107,000)</td>
</tr>
<tr>
<td>get fear</td>
<td>angs kry (0) // (2)</td>
<td>angst krijgen (836)</td>
<td>Angst kriegen (12,000)</td>
</tr>
</tbody>
</table>

Table 2. True/false friend relations in terms of the collocative function of the translation equivalents for *experience, have, cause/make* and *angs/angst/Angst*

The collocations *Angst bekomen/kriegen* “become anxious” in German and its Dutch equivalent *angst bekomen/krijgen* occur frequently in the respective corpora but their Afrikaans equivalents *angs bekom/kry* were almost non-existent in either the Media 24 archive or Google searches. The same holds true for Dutch *angst maken* and German *Angst machen* versus Afrikaans *angs maak*. From table 2, it is clear that a network of true/false relations exists in terms of the collocative function of the translation equivalents for *experience, have, cause/make*, etc., in German, Dutch and Afrikaans.

Especially dictionaries for productive use should clearly indicate these relations to the respective target users. So, for example should the user of an Afrikaans - Dutch – German dictionary be warned that Afrikaans *maak* and *bekom* are false friends of Dutch *maken* and *bekomen* and German *machen* and *bekommen* if used in collocation with *angs/angst*. This of course does not mean that *maken* and *bekomen* are false friends with their Afrikaans equivalents in other collocational relations. Numerous examples exist in the corpora of Afrikaans and Dutch reflecting true friend collocational relations as in example (1):

1. **Dutch** Afrikaans  
   a. maken maak  
   Optimaal Gebruik Maken Van Google optimaal gebruik maak van Google  
   Agenda Agenda  
   (optimal use of Google Agenda)  
   www.slideshare.net/antoinevandinter/optimaal-gebruik-maken-van-google-agenda  
   kettingen maken kettings maak  
   (make chains)  
   b. bekomen bekom  
   vulkanisch materiaal te bekomen vulkaniese materiaal te bekom  
   (to obtain volcanic material)  
   adres van het ... BTW-controlekantoor te adres van die BTW-kontrolekantoor te  
   bekomen bekom  
   (to obtain the address of the VAT Control Office)  

### 3.3. Collocational false friends at the morphosyntactic level

False friends at the morphosyntactic level are a problem for foreign language production. For example, most German collocations with *Schreck* (“fright”) take an article: *mit dem Schrecken*


davonkommen ("get a fright (and that is all)"; definite), einen Schreck bekommen, kriegen ("get a fright"; indefinite), jemandem fährt der Schreck in die Knochen (semi-idiomatic: "someone gets quite a scare"; definite). Most Afrikaans collocations with skrik, however, have no article.

Similar situations occur with respect to number (see above Klégr and Saldová’s example break the ice) or to the syntactic subcategorization of a collocation or of the noun it contains.

3.4. Colloccational false friends at the level of marked vocabulary

False friends with respect to register or other diasystematic marks, are less frequent, but equally problematic for a foreign language user. AF voordrag aanbied ("give a talk") is unmarked, its German literal equivalent einen Vortrag anbieten either has another meaning or is ironical; DE Vortrag halten, unmarked equivalent of the AF collocation has however no frequent literal translation into Afrikaans.

4. Presenting collocational false friends in dictionaries

In the presentation of collocational equivalents, in dictionary articles, an explicit marking of a false friend danger would be useful. At the lexical level, Martin and Gouws (2000: 789) propose a non-typographical structural marker “@”. The device proposed by Gouws/Prinsloo/De Schryver (2004) for an Afrikaans - Dutch dictionary is “vv” (valse vriende / valse vrienden (false friends)). If the false friend relation is between frequently used words in the source- and/or the target language, they suggest the addition of an exclamation mark “!” i.e. “vv!” to mark it as a dangerous false friend relation. In examples 2 - 5 we show sample entries which make use of these devices including FF in case of false friends entries for German.

(2) (Gouws/Prinsloo/De Schryver (2004: 798))

bakkie s.nw. (-s) 1 v.v.! Ligte vragmotor: Ons laai die tuinvullis sommer self op die ~ en neem dit na die stortingsterrein. 2 ‘n Klein houer: Ma sit my toebroodjes soggens in ‘n plastiese ~. Verkleinwoord van bak.

(Ligte vragmotor “light pickup truck”; ‘n Klein houer “a small container”)

In respect of sense 2, the meaning “small container”, Afrikaans and Dutch are true friends. However, the very specific Afrikaans sense bakkie “pickup truck” (not any other kind of vehicle) is a false friend. The fact that the false friend, based upon corpus data, constitutes 67% of the use of bakkie in Afrikaans justifies the inclusion of an exclamation mark.

We propose to use this device also in the equivalence indications of collocations.

The same device can also be used to warn against morphosyntactic differences or marked equivalents. If needed, an additional marker could be used in such cases, to clarify the property which is at stake, e.g. number or determination.

(3) voordrag ... Vortrag; ... ~ aanbied: Vortrag halten; vv Vortrag anbieten

Rechtsweg ... den ~ einschlagen FF prendere le vie legali (plural!)

An Afrikaans - German dictionary should reflect the false friend relation for the case of Angst erfahren, which involves domain marking; we suggest the treatment given in (4) below

(4) angst ... Angst;

~ ervaar Angst hebben, vv Angst erfahren (PSYCH)

Finally, consider the simplified treatment of selected sections of report in relation to the translation equivalent berig (AF), bericht (NL) and Bericht (DE). In this example the focus is on the use of vv to warn against combinations which are not commonly used in one or more of the languages. Naturally an article of this nature should also contain entries of the correct equivalent collocations in each of the languages.

(5) report (n) berig, verslag … (AF); exposé, melding, rapport, verslag, bericht, … (NL); Bericht, Protokoll, Report … (DE); write a ~ berig skryf (AF); bericht schrijven (NL); Bericht schreiben (DE); make a ~ bericht maken … (NL); Bericht machen … (DE); vv! berig maak …
For the English lemma *report*, the user should be able to comprehend from the translation equivalent paradigm that *berig*, *bericht* and *Bericht* are possible translation equivalents and true friends of the English lemma *report* and from the collocations following *write a report* that *skryf/schrijven/schreiben* are also true friends in the collocational relations *berig skryf* (AF)/*bericht schrijven* (NL)/*Bericht schreiben* (DE). However (s)he should notice that although the collocations *bericht maken* (NL) and *Bericht machen* (DE) are true friends, they are false friends with the Afrikaans *berig maak*. Finally (s)he should be warned that the collocation *Bericht vorlegen* (DE) stands in a false friend relation to both *berig voorlê* (AF) and *bericht voorleggen* (NL).

5. Conclusion, future work

The few examples discussed in this abstract should have shown the need for corpus based studies in the field of true vs. false friends in collocations, as an input to bilingual lexicography. Obviously, large enough monolingual corpora are needed. The lexicographer then has to compare the collocations of different languages with respect to the three levels mentioned above: lexical, morphosyntactic and diasystematic.

We have proposed simple devices to include explicit warnings against collocational false friends in dictionaries, which allow the lexicographer to clearly indicate the differences at stake.
References


http://www.google.com [Access date: 29 March 2008]


www.slideshare.net/ [Access date: 29 March 2008]