1 Introduction

Translation of lexical combinations into foreign languages is often unpredictable, word for word translation would give the wrong result in the following Dutch examples and their English and Spanish translations.

<table>
<thead>
<tr>
<th>Dutch</th>
<th>Lit. (Transl.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>een voorstel</td>
<td>do (make) a proposal</td>
</tr>
<tr>
<td>doen</td>
<td></td>
</tr>
<tr>
<td>een salto</td>
<td>make (do) a somersault</td>
</tr>
<tr>
<td>maken</td>
<td></td>
</tr>
<tr>
<td>een bad</td>
<td>take (have) a bath</td>
</tr>
<tr>
<td>nemen</td>
<td></td>
</tr>
<tr>
<td>afwassen</td>
<td>wash (do) the dishes</td>
</tr>
<tr>
<td>rode wijn</td>
<td>vino * rojo (tinto)</td>
</tr>
</tbody>
</table>

In ordinary mono- and bilingual dictionaries one can find all sorts of information about lexical items, e.g. syntactic, semantic, morphological and phonological information. If one combined these lexical items freely according to the grammar rules, it would produce a lot of grammatical combinations but a lot of ungrammatical ones too. Lexical substitution is restricted, this is what is called productivity.
Some information about lexical combinations is present in the dictionary, but it is often given in an implicit way and it is not sufficient. Take for example the Dutch words *moord* (murder) and *zelfmoord* (suicide). One can say in Dutch *een moord plegen* (to commit (a) murder) or *zelfmoord plegen* (to commit suicide). The translation of *plegen* in English is *to commit* or *to perpetrate*. It gives you the idea that one can say in English *to perpetrate (a) murder* and *to perpetrate suicide*. It happens to be the case that one can say *to perpetrate* for *murder*, but not for *suicide*. This kind of information is not made explicit in the dictionary. So a language learner might make the mistake to use the wrong expression, a machine translator will most certainly generate the wrong solution too.

It can be concluded that it would be helpful for language learners to have lexical combinatory information whereas for machine translation it is essential. The examples of lexical combinations we have given are considered 'collocations'. The purpose of this paper is to find out how collocational information could be represented. The outline of this paper is as follows. In the first section we discuss problems that occur with the characterization of collocations. At the end of this section we come up with a table that shows the distinction of collocations from other lexical combinations, i.e. *idioms* and *free combinations*. We found the lexical functions that Mel'čuk et al. suggested for lexical combinations helpful as possible representation of collocations. These lexical functions are discussed in section two. Finally, in section three, we try to make a proposal for implementation of collocations in a Categorial Grammar like formalism, in which our ideas are incorporated.

2 Characteristics of lexical combinations

In this section we discuss some essential characteristics of lexical combinations. It should help us to identify collocations. The characteristics are *cooccurrence, productivity, compositionality, metaphorical interpretation, internal modification, quantification* and *sensitivity to transformations*.

We distinguish the following kind of lexical combinations:

1. Idioms

2. Syntactically versatile idioms

3. Collocations

4. Free combinations

2.1 Idioms

Wood (1986) claims that a true idiom is a construction that is totally non-compositional and non-productive. Non-compositional means that the meaning
of the construction cannot be made up by composing the meanings of the individual words. Furthermore it can be said that their meaning is based on metaphors. Productivity is used in terms of tolerance of lexical substitution.

An example of a true idiom would be *de pijp aan Maarten geven* (lit: *to give the pipe to Maarten*, cf. *to kick the bucket*). The meaning of this expression is *to die*, a metaphorical meaning which cannot be made up of the parts of the expression.

Lexical substitution is not possible for any part of the idiom, without losing the interpretation to die. E.g. *de pijp aan Maarten overhandigen* (lit: *to hand over the pipe to Maarten*) and *to give the bucket a kick* are unambiguous, as are *de pijp aan Jan geven* (lit: *to give the pipe to John*) or *het rookartikel aan Maarten geven* and *to kick the pail*.

Idioms do not allow internal modification, quantification, passivization, topicalization, *wh*-movement, raising or VP ellipsis, without losing the idiomatic interpretation, as will be shown.

**Internal modification** This is not possible for any part of the idiom.

* de grote pijp aan Maarten geven  
  * to kick the great bucket  
  (lit: *to give the large pipe to Maarten*)

* de pijp aan kleine Maarten geven  
  * to give the pipe to little Maarten*)

**Quantification** Similarly:

* alle pijpen aan Maarten geven  
  (lit: *to give all the pipes to Maarten*)

* to kick all the buckets

**Passivization** * De pijp werd aan Maarten gegeven  
  (lit: The pipe was given to Maarten)  
  * The bucket was kicked

**Topicalization** * De pijp gaf hij aan Maarten  
  (lit: The pipe he gave to Maarten)  
  * The bucket he kicked

**Wh-movement** * wat gaf hij aan Maarten  
  (what did he give to Maarten)  
  * what did he kick  
  * aan wie gaf hij de pijp  
  (to whom did he give the pipe)
raising * de pijp schijnt aan Maarten gegeven te zijn
   (lit: the pipe seems to have been given to Maarten)
   * the bucket seems to have been kicked

VP-ellipsis * hij zei dat hij de pijp aan Maarten gaf, maar dat deed hij niet
   (lit: he said he gave the pipe to Maarten, but he didn’t)
   * he said he kicked the bucket, but he didn’t

2.2 Syntactically versatile idioms

In GPSG (1985), GKPS discuss a group of lexical combinations that are consid-
ered traditionally as idioms, because they are totally non-productive and they
look semantically non-compositional. Examples of these idioms are:
to keep tabs on someone
to pull strings to get the job

They are non-productive as no element can be substituted:
to *maintain/hold tabs on someone
to keep tabs *against/for someone
to *yank/tugg at strings to get the job
to pull *twine/rope to get the job

The parts of those idioms are syntactically mobile, i.e. which appear in passive
or raising constructions:
“Tabs were kept on Terry by the KGB”
“He believed strings to have been pulled in Sandy’s getting the job”

A Dutch example would be:
“De teugels van het bewind werden in handen genomen door het leger”
(the reins of government were taken up by the army)

GKPS claim that there are reasons to believe that the parts of such idioms
should be assigned interpretations: They allow semantic internal modification,
parts of these idioms may be quantified, emphasized through topicalization, and
can be omitted in VP ellipsis. Wh-movement of the object is not allowed.

1. Parts of idioms may be modified internally by either adjectives or relative
   clauses: Close tabs were kept on Terry by the KGB, or She got the job by
   pulling strings that weren’t available to anyone else.

2. Parts of idioms may be quantified: He would pull a string or two to get
   the job
3. Parts of idioms may be emphasized through topicalization: *They might keep tabs on us, but close tabs, they'll never keep on us.*

4. Wh-movement on the object is not allowed: *What did they keep on us.*

5. Parts of idioms can be omitted in VP-Ellipsis: *I said close tabs would be kept on her, but they weren't.*

On their analysis the notion of a partial function is exploited so that principles of semantic interpretation apply to these expressions in the normal manner to assign an interpretation (e.g. the interpretation of tabs is undefined unless it is an argument of the interpretation of keep). In support of this compositional treatment, they suggest that the syntactically versatile idioms all have a metaphorical basis.

### 2.3 Collocations

**Frequency of cooccurrence** People like Benson (1986) and Heid and Raab (1980) have tried to identify collocations on the basis of large text corpora, making use of the notion *frequency of occurrence*. This notion is either too small, as not all collocations occur frequently in texts, or it is too wide, as it also includes *idiomatic expressions* and a lot of *free combinations*, which are fully *compositional* and do not pose any problem in translation presumably, e.g. *een huis/auto/boek/kado kopen/verkopen* (lit. to buy/sell a house/car/book/present).

**Productivity** Amongst others Dillon (1977) and Benson (1986) see as a distinctive feature of collocations the restricted commutability of one of the component elements. For example the only synonym for the verb *plegen* (to commit) in *een moord plegen* (to commit (a) murder) seems to be *begaan* (to commit/perpetrate). In this respect collocations are more free than idioms, but less free than *free combinations*. Productivity is not an entirely reliable notion to identify collocations, as some of them seem unproductive and others seem rather productive. The following lexical combinations seem rather unproductive:

- *schrall/severe troost/? (poor/cold comfort)*
- *schril/grof tegenstelling/ contrast (violent/great contrast)*
- *zelfmoord/? plegen/ *begaan* (to commit suicide)*
- *zwaar/? ademhalen/respireren ? (to breath/respire heavily)*
- *een moord/? plegen/begaan* (to commit/perpetrate a murder)

There are combinations that are more productive.

*baarlijk/volslagen/klinkklare onzin/flauwekul/nonsense*  
*(sheer/downright/rank/arrant rubbish/codswallop/nonsense)*
compositionality A very distinctive property of collocations as opposed to free combinations is the fact that they are not entirely compositional, the meaning carried by one or more of the constituent elements of a collocation is different from its meaning in more 'neutral' contexts. E.g.:

1. het boek is zwaar (the book is heavy)
2. zwaar ademhalen (to breath heavily)
3. een zware taak (a heavy task)

In (1) the most 'neutral' context, zwaar (heavy) has to be interpreted as a lot of weight. In (2) zwaar (heavy) can only be understood in combination with ademhalen (to breathe) as 'to breathe deeply with some noise', there is no such meaning of zwaar independently. In (3) again zwaar has to be interpreted differently from its neutral sense, together with taak (task) it means difficult, taking a lot of work.

Collocations are considered to differ from idioms in the sense that they exhibit some semantic compositionality, whereas idioms do not. So one of the element retains its normal sense. een moord plegen (to commit a murder) means (over)moorden (to murder).

To decide which elements of a lexical combination have a literal meaning and which ones have not L. Verstraten proposes a test for which she has used the notion of entailment of Keenan & Faltz (1985) "e entails e' iff e contains all the information in e'(and possibly more)". It means that 'less informative' is included by 'more informative'. The test says that if a lexical element in a lexical combination has a literal meaning it should be possible to leave this element unchanged and replace the other elements without problems by elements that are less informative.

<table>
<thead>
<tr>
<th>enthalment</th>
<th>(bitter cold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bittere kou</td>
<td>(bitter cold)</td>
</tr>
<tr>
<td>* bittere temperatuur/</td>
<td>bitter temperature/</td>
</tr>
<tr>
<td>weersteldeheid</td>
<td>weather condition</td>
</tr>
<tr>
<td>onprettige kou</td>
<td>unpleasant cold</td>
</tr>
</tbody>
</table>

Only cold is used in its normal sense as it is the only element that has the same interpretation if it is modified by a less informative element than bitter. Bitter does not have an interpretation within this lexical combination that it can have with a hyperonym of cold. With a drijvende doodskist (a floating coffin) it is the adjective that retains its semantic value and not the noun, cause it is not some kind of coffin, but a not seaworthy ship. Note that coffin has a metaphorical meaning here. For Verb Adverb combinations it is difficult to find hyperonyms. We think that the ideas of GKPS about syntactic versatile idioms might help. They find these constructions essentially compositional, because they allow a.o. semantic internal modification. This property can be
used to prove that a Verb-Adverb collocation cannot be considered semantically compositional.

If the adverb is semantically empty it should not be possible to modify it internally, e.g. *zaad ademen* vs. *loodszaad ademen*, or *hoog weglopen met* vs. *torenhoog weglopen met*.

Verstraten mentions that for Noun Verb combinations with a semantically empty verb it is not possible to find a verb that it less informative. In his thesis on Support Verb Constructions (SVC) M. Verhagen (90) tries to prove the semantic emptiness of the Support Verb in SVC’s, which are a subset of collocations. He has used a test to distinguish them from ordinary Noun-Verb constructions.

*Adverb-Adjective alternation*

An adverb which is morphologically related to an adjective Adj and which modifies an SVCphrase can play the same role as the adjective Adj which modifies the Npred in the SVCphrase.

Compare *Zij deden snel een aanval* (They carried out quickly an attack) and *Zij deden een snelle aanval* (They carried out a quick attack). The semantic emptiness of Vsup shows itself here as well. Intuitively, the adverb does not modify the Vsup but rather the SVC as a whole. And since the Noun bears the largest part of the meaning of the complete SVC there is little difference between modifying the Noun and modifying the Vsup.

Generally, this alternation is not possible with ordinary verbs: *Zij beschreef snel een aanval* (She quickly described an attack) vs. *Zij beschreef een snelle aanval* (She described a quick attack).

To show that collocations differ from syntactic versatile idioms in this respect also, see the example *zij namen snel de teugels in handen* (they took up the reins quickly), which doesn’t mean the same as *zij namen de snelle teugels in handen* (they took up the quick reins).

So ‘Noun-Verb collocations’ where the verb does not maintain its normal sense could be identified with this test. There happen to be ‘Noun-Verb’ collocations where it is the noun that has a nonneutral meaning, e.g. *tegen de bierknaai vechten* (fighting a losing battle). Again bierknaai is used metaphorically.

Internal modification of the Noun in Noun-Verb collocations is possible if the noun is the semantic head of the combination: *een bloedige moord plegen* (to commit a bloody murder) vs. *tegen de harde bierknaai vechten* (to fight a heavy loosing battle).

For Noun-Adj collocations it seems to be possible to modify the head of the collocation, although it is not clear whether internal modification is possible.

---

1The compound adjectives *loodszaad* and *torenhoog* are are magnifications, meaning very heavy/high.
een zeer bittere kou       a very bitter cold  
*een galbittere kou       a cold bitter as gall  
?een wegrijvende doodskist a coffin floating away  
?een drijvende witte doodskist a floating white coffin

We have already discussed internal modification for Verb-Adverb collocations. Quantification is allowed: *hij heeft alle moorden gepleegd (he committed all the murders) vs. * tegen alle bierkaaien vechten (to fight all losing battles). The same holds for Adverb-Noun combinations: *alle strenge winters (all severe winters) vs. *alle drijvende doodskisten (all floating coffins).

Noun-Verb collocations, with the noun as the semantic head allow topicalization: *de moord plegde hij (the murder he committed). Compare *de bierkaai vocht hij tegen (the losing battle he fought against).

With Noun-Adj and Verb-Adverb collocations topicalization is possible: *de winter was stren (the winter, that was severe) vs. zwaar was de ademhaling heavy was the breathing).

Passivization only holds for Noun-Verb collocations if the noun is the semantic head: *de moord werd gepleegd (the murder was committed) vs. *er werd tegen de bierkaai gevochten (a losing battle was fought).

Wh-movement gives the following result:

<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wat werd er gepleegd?</td>
<td>Een moord</td>
</tr>
<tr>
<td>(What was committed?)</td>
<td>A murder )</td>
</tr>
<tr>
<td>Waartegen werd er gevochten?</td>
<td>* De bierkaai</td>
</tr>
<tr>
<td>(What was fought against?)</td>
<td>The losing battle)</td>
</tr>
<tr>
<td>Hoe was de winter?</td>
<td>Strenge</td>
</tr>
<tr>
<td>(How was the winter?)</td>
<td>Severe )</td>
</tr>
<tr>
<td>Wat voor een doodskist?</td>
<td>* Een vliegende</td>
</tr>
<tr>
<td>(What kind of coffin?)</td>
<td>A flying one)</td>
</tr>
<tr>
<td>Hoe werd er geadaand?</td>
<td>Zwaar</td>
</tr>
<tr>
<td>Hoe werd er gedaan?</td>
<td>* Moeilijk</td>
</tr>
</tbody>
</table>

Raising and VP-Ellipsis only hold for Noun-Verb collocations, if the noun is the semantic head it is possible: *Hij zei dat de moord gepleegd scheen te zijn door Jan (He said the murder seems to have been committed by John). But: * Hij zei dat er tegen de bierkaai bleek gevochten te zijn (he said a losing battle seemed to have been fought). Furthermore: *Hij zei dat hij de moord gepleegd had, maar dat had hij niet (He said he had committed the murder, but he hadn’t) vs. *Hij zei dat hij tegen de bierkaai vocht, maar dat deed hij niet (he said he fought a losing battle, but he didn’t).

2.4 Free combinations

Free combinations do not pose any problem in translation. They are totally semantically compositional in the sense that each element has its semantic independence and the meaning of the lexical combination can be composed by the meaning of the parts. Free combinations are totally productive in the sense that
each element could be substituted without affecting the meaning of the other element or elements. E.g.: een huis/schip/brug bouwen/vernietigen/bekijken (a house/ship/bridge) (to build/destroy/look at).

The verb piegen (to commit) can only combine with a restricted number of nouns, e.g. with misdaad (crime), overvrij (adultery), fraude (fraud) and some other nouns that have the meaning of "something wrong or unacceptable". Whereas e.g. the verb to buy can combine with almost any noun.

Free combinations allow all the following properties; internal modification, quantification, passivization, topicalization, wh-movement, raising and VP-Ellipsis.

We have gathered all the characteristics that we discussed for lexical combinations in the following table, where the columns are 1 idioms, 2 versatile idioms, 3 heads and 4 collocates in collocations and 5 free combinations.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>coocurrence</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/</td>
</tr>
<tr>
<td>productivity</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>metaphorical</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>compositional</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>modification</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>?/-</td>
<td>-</td>
</tr>
<tr>
<td>quantification</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>?/+</td>
<td>+</td>
</tr>
<tr>
<td>passivization</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>topicalization</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>wh-movement</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>raising</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>VP-ellipsis</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

As can be seen from this table idioms have a negative value for all the characteristics except for metaphorical and coocurrence. Free combinations are positive to all characteristics except for metaphorical and w.r.t. coocurrence it is not possible to classify them positive or negative as some free combinations might occur very frequently. The syntactically mobile idioms differ only from free combinations in the aspects productivity and metaphorical. We have divided the collocations into heads and collocates as they display different values for the properties. The heads of the collocations differ from idioms' in productivity and metaphorical. They display the same values for all the properties as free combinations. The values for the collocates of the collocations look very similar to the values of idioms. Internal modification and quantification are problematic in the respect that it accounts for some collocations, but not for all. Collocates may be metaphorical, whereas idioms always are. If a collocate is not metaphorical, it still has a meaning that is different from its 'neutral' sense or it is semantically empty.
3 Lexical Functions

In their Explanatory and Combinatory Dictionary (ECD), Mel'uk et al. have attempted to handle collocations using a relatively small collection of lexical functions (LF's) covering both Adj-Noun and Verb-Noun collocations. We believe these functions to be useful as they might express what we have called so far semantic emptiness. In this section we take a closer look at these functions.

It has to be understood that within Verb-Noun and Adj-Noun collocations Mel'uk et al. consider the noun to be the head of the collocation. The LF concerned is then applied to this head and gives the appropriate collocate which can be an adjective or a verb. So for example, the LF Magn(X) picks out the collocate for 'high degree':

Magn(winter) = streng (severe)
Magn(kou (Eng. cold) = bitter

The LF OPER means something like activate, it results in the following:

OPER(aandacht) = geven (Eng. to pay attention)
OPER(afwas) = doen (Eng. doing/washing the dishes)

It is possible to combine several LF's, for example the LF's INCEP, CONT, FIN: to start, to continue and to finish are used the most in combination with other LF's.

INCEP.OPER(aanval) = inzetten / openen op (launch / mount an attack)
CONT.OPER(geduld) = bewaren (to keep one's patience)
FIN.OPER(geduld) = verliezen (to lose one's patience)

4 Implementation

In this section we want to examine how Categorial Grammar (CG) can serve as a computational vehicle for implementing ideas about collocations. We argue that lexical combinations can be treated as special cases of partial execution of the lexicon.

To start, we have to be clear about our notation of categories. The category 'X/Y' will denote a functor that combines with an argument Y to the right to form an X, and 'Y\X' will denote a functor that combines with an argument Y to the left to form an X.

In the lexicon, every word will be linked to a syntactic category and a meaning symbol. For clarity, we will only use the combinatory rule of application:

Forward Application \[X/Y : f + Y : a \Rightarrow X : f(a)\]
Backward Application \[Y : a + Y\X : f \Rightarrow X : f(a)\]
An example of a lexicon:

```
<table>
<thead>
<tr>
<th>&quot;John&quot;</th>
<th>&quot;sleeps&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>syntax: NP</td>
<td>syntax: NP\S</td>
</tr>
<tr>
<td>semantics: JOHNN</td>
<td>semantics: SLEEP</td>
</tr>
</tbody>
</table>
```

We will employ a format for derivation trees where indentation corresponds to branching in the tree and top-down order corresponds to left-right order, for example:

```
Appl[
  S : SLEEP(JOHN)
  NP : JOHN
  NP\S : SLEEP
]
```

**Partial execution** By partial execution we mean that we calculate a semantical expression for each of the atomic categories that a complex category consists of and store the results. One way of doing this is via proofnets (for discussion of proofnets, see Moortgat 90). Since this is not a discussion of proofnets, we will use a notation where the calculated semantical expressions are paired with the atomic syntactic categories.

For example, if we start with this lexical entry of the adjective "severe":

```
"severe"
syntax: N/N
semantics: SEVERE
```

Then the result of partial execution would look like this — with 'x' denoting a variable:

```
"severe"
(N : SEVERE(x))/N : x
```

Of course the Application rules must now be adjusted:

```
Appl /: (X : T)/(Y : U) + Y : V ⇒ X : T (U = V)
Appl \: Y : V + (Y : U)\(X : T) ⇒ X : T (U = V )
```

Here 'U = V' means: unify U and V.

The next step is to produce extra reading distinctions by executing collocations. In the case of the collocation "severe winter", the semantical argument of "severe" gets bound to WINTER and the semantics of the whole expression becomes SEVERE(WINTER). Since in this collocation "winter" retains its original meaning while "severe" behaves like a lexical function, we replace SEVERE with the lexical function Magn. The resultant extra lexical entry is the following:

```
"severe"
(N : Magn(WINTER))/(N : WINTER)
```
Also, idioms can be treated this way. For example, if we execute “kicks the bucket” and replace the composition of the meanings of “kicks” and “the bucket” with a single meaning atom DIE, an extra lexical entry for “kicks” can be the following:

“kicks”

\(((\text{NP} : x))(((\text{S} : \text{DIE}(x)))/(\text{NP} : \text{THE(BUCKET)}))\)

A derivation for “john kicks the bucket” would in this approach look like this:

\[
\begin{align*}
\text{Appl}[] & \quad \text{S} : \text{DIE(JOHN)} \\
& \quad \text{NP} : \text{JOHN} \\
\text{Appl}[] & \quad (\text{NP} : x) (((\text{S} : \text{DIE}(x)))/(\text{NP} : \text{THE(BUCKET)})) \\
\text{Appl}[] & \quad \text{NP} : \text{THE(BUCKET)} \\
& \quad (\text{NP} : \text{THE(y)})/(\text{N} : y) \\
& \quad \text{NP} : \text{BUCKET}
\end{align*}
\]

This is reminiscent of the way this idiom is treated in the Utrecht MiMo2 translation system.

We now have a way of making a rough distinction between collocations and idioms: with collocations only a part of the semantics of the lexical combination gets reinterpreted after partial execution — e.g. “severe” in “severe winter” —, while with idioms the reinterpretation is total. Of course finer distinctions based on the tests that we have been discussing in this paper can only be made in a categorial grammar formalism that can handle the phenomena listed in the table in 2.4.

**Lexical functions** If we examine the relationship between LF’s and CG, we find that the head in a LF corresponds to the argument category in CG and the collocate in a LF corresponds to the functor category in CG, as in these examples — with ‘()’ meaning head:

<table>
<thead>
<tr>
<th>CG</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>(NP\S)/NP + NP</td>
<td>turn + (a somersault)</td>
</tr>
<tr>
<td>N/N + N</td>
<td>severe + (winter)</td>
</tr>
<tr>
<td>NP + NP\S</td>
<td>(the alarm) + goes off</td>
</tr>
<tr>
<td>N + N\N</td>
<td>(war) + of agression</td>
</tr>
<tr>
<td>(N/N)/(N/N) + N/N</td>
<td>firmly + (shut)</td>
</tr>
<tr>
<td>NP\S + (NP\S)/(NP\S)</td>
<td>(sleep) + deeply</td>
</tr>
</tbody>
</table>

This suggests a way of constructing extra lexical entries for collocations from lexical functions such as used in Melčuk and Apresjan’s Explanatory and Combinatory Dictionary. Take for example the following lexical function:
Magn("winter") = "severe"

We can execute the collocation "severe winter" as outlined above, and use this equation to determine that SEVERE should be changed to Magn in the creation of a new lexical entry.

References
