CHAPTER 8: FUNGRAMKB

3. The Grammaticon and ARTEMIS

The Grammaticon constitutes the grammatical module within the linguistic level of FunGramKB. As pointed out in the overview of FunGramKB, the Grammaticon has inherited its structure from the four levels of the LCM. Thus, the Grammaticon in FunGramKB has four Constructicons that roughly correspond to the four layers of the LCM. The Grammaticon contains argument-structure constructions (L1-Constructicon), implicational constructions (L2-Constructicon), illocutionary constructions (L3-Constructicon), and discourse constructions (L4-Constructicon). The Grammaticon has thus been conceived as a repository of constructional schemata in a given language.

The representation of constructional schemata in the FunGramKB Grammaticon was at earlier stages restricted to argument-structure constructions. However, ARTEMIS is now able to deal with non-propositional meaning. There are substantial differences between the ways in which argument and non-argument constructions are represented in the Grammaticon. Let us outline the main aspects related to the representation of each type of constructional schemata.

(i) The representation of argument-structure constructions in FunGramKB is realized through *Attribute Value Matrixes* (AVMs). AVMs include a number of descriptors and constraints that regulate the unification of elements in such a way that the semantic plausibility of the result is guaranteed. The result of this unification process is the CLS. The completion of each AVM and the automatic generation of the CLS requires the collaboration oh three types of production rules: (a) lexical rules, which specify the properties of lexical entries by retrieving information from the Lexicon and the Ontology; (b) constructional rules, which feed from the Grammaticon in order to determine the properties of constructional schemata; (c) syntactic rules, which do not draw knowledge automatically from FunGramKB, but need to be constructed by elaborating the layered structure of the sentence in the form of a syntactic tree.

Once the CLS has been created, ARTEMIS automatically generates the COREL scheme, which constitutes the input for the reasoning engine. The CLS includes the morpho-syntactic features of the construction (i.e. Aktionsart, number and type of variables, thematic roles, etc.), while the COREL representation provides a conceptual semantic description, which is language independent. For a more detailed discussion of the functioning of ARTEMIS at the level of argument-structure constructions, we refer the reader to Periñán (in press), and Periñán and Arcas (in press).

(ii) Idiomatic constructions are also expressed in terms of AVMs. However, the representation of these constructions does not include a CLS. Non-argument constructions have fixed and variable elements. The latter are represented by the letters X, Y, and Z, and may be parametrized in different ways. For instance, the X and the Y in the level-2 construction *It Wouldn't Kill X to Y* usually represent a volitional entity (X) and an action (Y), as in *It wouldn't kill you to have a haircut*. Once the text is introduced and the variables have been filled in, ARTEMIS retrieves the necessary information about these variable units from the Lexicon or from the L1-Constructicon. In our example, the information related to the X ('you') would be contained within the Lexicon, while the conceptual construct related to the Y ('have a haircut') needs to be drawn from the L1-Constructicon. The fixed part of the construction is matched to the corresponding construction in the L1-Constructicon, from which ARTEMIS imports the CLS, which is not altered. Also, the fixed part of the construction is a keyword that serves as an activator for the identification of the construction as idiomatic. The idiomatic construction contributes the pragmatic reading of the sentence. In our example, the level-2 construction provides the implicated meaning that the speaker disapproves of some situation and wants it to be changed. This semantic description is represented in COREL. Next section provides a more detailed account of this process, along with examples from the implicational, illocutionary and discourse levels of FunGramKB.

4. The representation of constructional schemata at idiomatic levels of the Grammaticon

The Grammaticon in FunGramKB presents the following interface:

Figure 3. The interface of the Grammaticon in FunGramKB

In this section we address in detail the representation of idiomatic constructions in the Grammaticon. In order to do so, we will present three constructions for each of the idiomatic levels of the Grammaticon. Furthermore, these constructions will serve as illustration of the functioning of ARTEMIS at these levels as sketched in the previous section.

4.1. Level 2-Constructicon

The constructional schemata to be stored within the level 2-Constructicon are those that represent implicational constructions, that is, those that correspond to level 2 in the LCM. The following subsections offer an account of the representation of the following constructions in the level 2-Constructicon: *Why Does X Have To Y?, Do I Look Like I X?,* and *It Wouldn't Kill X To Y*. Let us address each of them in turn.

4.1.1. The Why Does X Have To Y? *construction*

This constructional schema is generally used in contexts in which the speaker is upset about a given state of affairs. An instantiation of this construction is the sentence *Why do things have to be so complicated?* By uttering this sentence, the speaker is not enquiring about the reason that makes things complicated, but is rather expressing his discomfort about it. In much the same vein, the speaker of a sentence such as *Why does he have to call you every five minutes* is not concerned about what causes a third person to call the hearer every five minutes. Instead, he is making evident that he is bothered by the frequent calls. Figure 4 below captures the interface of this construction in the Grammaticon:

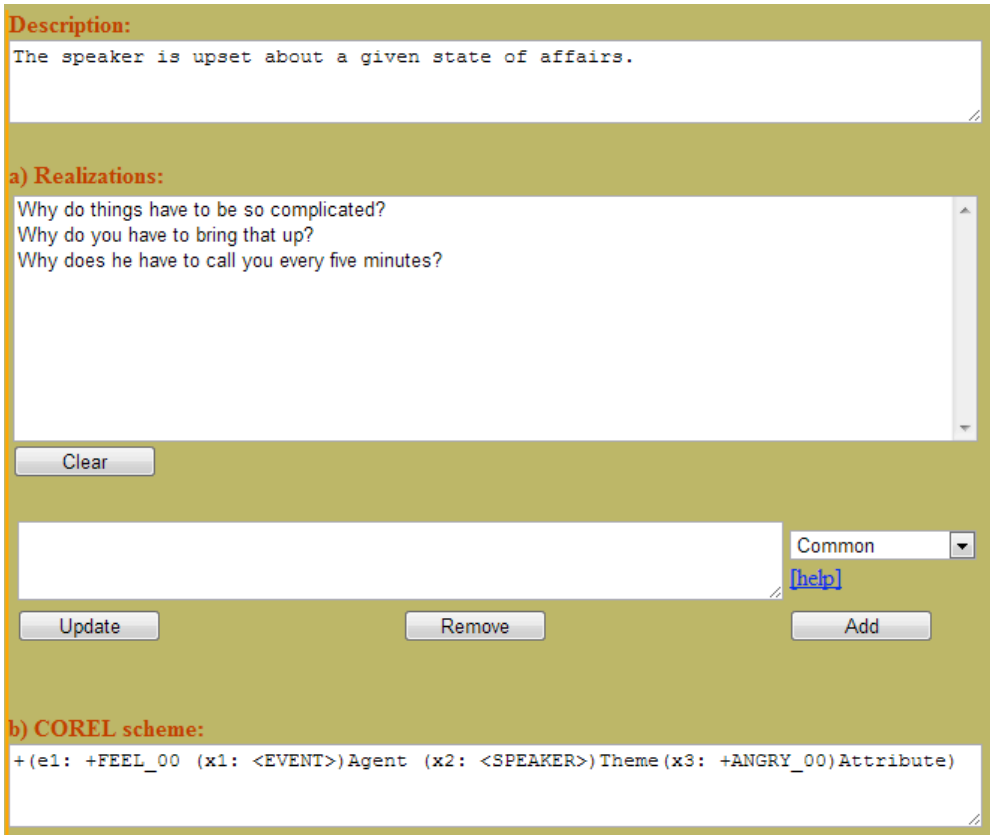

Figure 4. The interface of the *Why Does X Have to Y?* construction

As it is shown in figure 4, the interface for idiomatic constructions in FunGramKB has three parts: description, realizations and the COREL scheme. The description of the construction is a statement that captures the non-propositional meaning of the construction, in the present case, the idea that the speaker is upset. The realization box includes real instances of the construction as used by speakers of English. In turn, the COREL scheme comprises semantic information that matches the linguistic description provided for the construction. In other words, the COREL scheme is a conceptual representation of the non-propositional meaning of the construction:

(1) +(e1: +FEEL_00 (x1: <EVENT>) Agent (x2: <SPEAKER>) Theme (x3: +ANGRY_00)Attribute)

The COREL proposition in (1) encodes the following information: an event (x1) causes the speaker (x2) to be angry (x3). This conceptual representation enriches the information contained within the COREL scheme of the level-1 construction from which ARTEMIS draws the CLS. Thus, the CLS of the text is not modified, but the COREL scheme is extended. The fixed part of the construction ('Why Does/Have to') allows ARTEMIS to (i) identify the construction as idiomatic, and (ii) identify, through pattern matching, the level-1 construction from which the CLS is to be retrieved. In turn, the X and Y elements are to be filled in by items in the Lexicon and the L1-Constructicon. Let us take the sentence *Why does Tom have to be aggressive?* The variables X and Y have been parametrized by 'Tom' and 'be aggressive' respectively. ARTEMIS will retrieve the information related to 'Tom' from the Onomasticon, which stores conceptual information about actual entities and events. In turn, 'be' and 'aggressive', which are lexical entries in the Lexicon, are connected to the concepts BE_01 and VIOLENT_00 in the Ontology. ARTEMIS draws the semantic information from the meaning postulates of these basic concepts.

4.1.2. The Do I Look Like I X? *construction*

In the linguistic analysis based on cognitive operations offered in Chapter 7, section 4.3, we contended that the use of this constructional pattern implicates that (i) the hearer presupposes the speaker's involvement in a given state of affairs, and (ii) the speaker is upset about the hearer's presupposition, because it is evident that the content of such a presupposition is not the case. These two statements are included in the description of the construction in the interface of the Grammaticon:

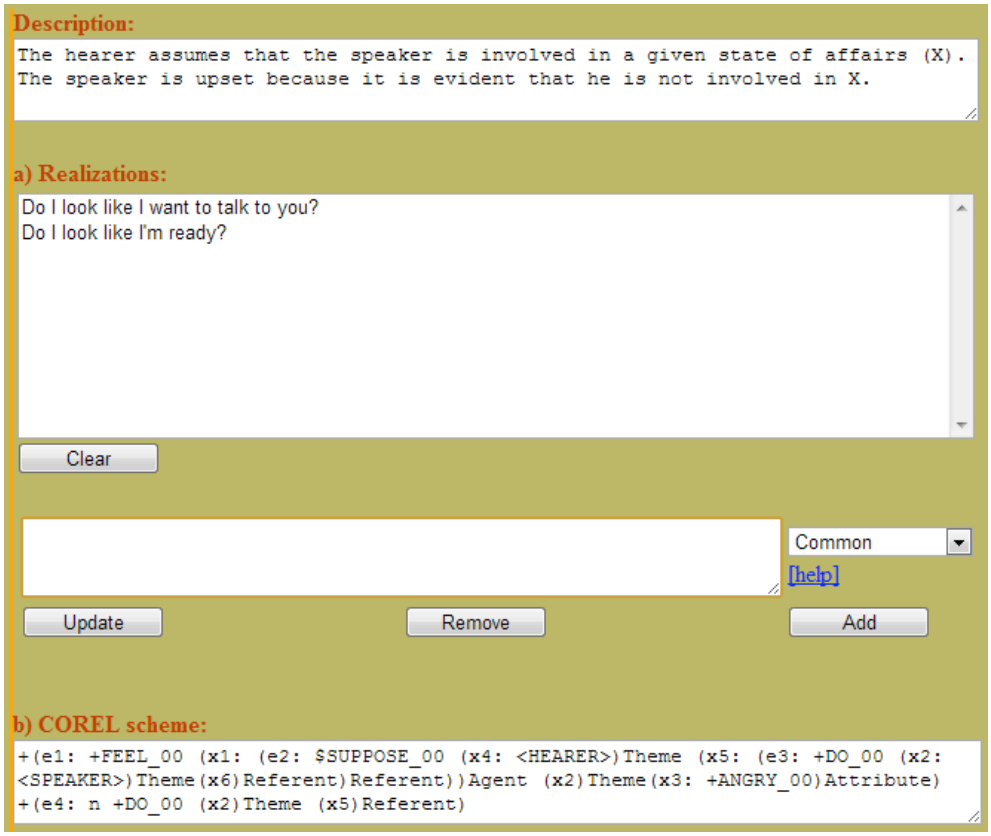

Figure 5. The interface of the *Do I Look Like I X?* construction

These meaning implications, which go beyond propositional meaning, are comprised within the COREL scheme of this level-2 construction. Once ARTEMIS has identified that this is a level-2 construction, the implications conveyed by its COREL scheme add up to the propositional meaning obtained from the level-1 construction that matches the fixed elements. In this case, more that one predication is needed:

(2) +(e1: +FEEL_00 (x1: (e2: $SUPPOSE_00 (x4: <HEARER>)Theme (x5: (e3: +DO_00 (x2: <SPEAKER>)Theme(x6)Referent)Referent))Agent (x2)Theme(x3: +ANGRY_00)Attribute)
+(e4: n +DO_00 (x2)Theme (x5)Referent)

Two subordinate predications (e2 and e3) are embedded within the first predication. This predication encodes the conceptual information related to the first part of the linguistic description: what makes the speaker feel angry is the fact that the hearer presupposes that he is involved in the course of a given event. The last predication (e4) means that the speaker is not engaged in such a state of affairs.

4.1.3. The It Wouldn't Kill X To Y *construction*

As we advanced in section 3 above, the implicated meaning of this construction is that (i) the speaker thinks that it would be good that somebody (X) would carry out a given action (Y), and (ii) the speaker wants this person to

do it. This construction, which is a hyperbolic variant of the *It Wouldn't Harm X To Y* construction, was analyzed as an instance of strengthening underlying hyperbolic litotes at the implicational level (see Chapter 7, section 5.2).

By using this construction, the speaker chooses an indirect way of expressing his approval for certain action, and his wish for this action to be carried out. These implications are captured in the COREL scheme, as we can see in Figure 6:

Figure 6. The interface of the *It Wouldn't Kill X To Y* construction

The different realizations of this construction show that the X variable may be filled in by other entities other than the hearer. However, the parametrization of

this variable is restricted to volitional agents. Variable Y may be realized by any action that the agent (X) is able to perform.

As was the case in the *Do I Look Like I X* construction, the COREL representation of the non-propositional meaning of this construction requires several predications, as shown in (3):

(3) +(e1: +SAY_00 (x1)Theme (x4: (e2: +BE_01 (x2)Theme (x5: +GOOD_00 | +RIGHT_00)Attribute))Referent (x3)Goal)
+(e3: +WANT_00 (x1)Theme (x6: (e4: +DO_00 (x7)Theme (x2)Referent))Referent)

The first predication e1, in which predication e2 is embedded, can be translated into natural language by saying that the speaker (x1) tells the hearer (x3) that something (x2) is good or right. Predication e3, complemented by predication e4, captures the information that the speaker (x1) wants someone (x7) to do something (x2).

Before we move on to the level 3-Constructicon, we would like to present a problematic case of an implicational construction that cannot be processed by ARTEMIS. It is the case of the *Don't X Me* construction (see Chapter X, section X). Recall from our linguistic analysis of this construction in terms of cognitive operations that the implicated meaning arises from the parametrization of the X variable through an echo of a previously uttered expression. When the X slot of the construction is filled in by a verbal

predicate (which is generally the case), we have instantiations of a level-1 construction (e.g. *Don't leave me, Don't speak to me*, etc.). In other words, the *Don't X Me* construction only qualifies as a level-2 construction if the X is realized by the repetition of previous discourse (e.g. *Husband: But if this were Syracuse, we would be in the air, honey. / Wife: Don't 'honey' me*). This peculiarity poses a problem for ARTEMIS, which, to the present, is only able to process one sentence at a time. Therefore, for now, it is not viable to make the program recognize which realizations of the X variable are the repetition of a previous utterance. This means that, for ARTEMIS, the *Don't X Me* construction is always to be processed as a level-1 construction, so the potential non-propositional meaning of certain instantiations of the construction would be missed.

4.2. Level 3-Constructicon

At this level, constructions have been grouped according to their illocutionary meaning. Consider, for instance, the fact that every constructional realization that conveys a request (e.g. *Can You X?, Will You X?, Could You X?*, etc), shares the same non-propositional meaning, that is, the speaker asks somebody to do something. Therefore, the COREL scheme for all these level-3 constructions is the same. As the reader may be aware by now, the COREL representation is the most relevant part of the representation of idiomatic constructions in the Grammaticon. It would thus be unwise to create different

entries for each construction. Rather, they are comprised within one single entry named after the illocutionary meaning shared by the constructions. The fixed part of the constructional realizations serves ARTEMIS in the identification of these patterns as illocutionary. Also, as with implicational constructions, the fixed part allows ARTEMIS to import the CLS from the corresponding level-1 construction, while the variables are to be filled in by retrieving information from the Lexicon of the level 1-Constructicon.

Let us now see the representation of illocutionary constructions for offering, promising and requesting in the Grammaticon.

4.2.1. Offering

The interface for Offering constructions (and every illocutionary construction) is the same as in implicational constructions:

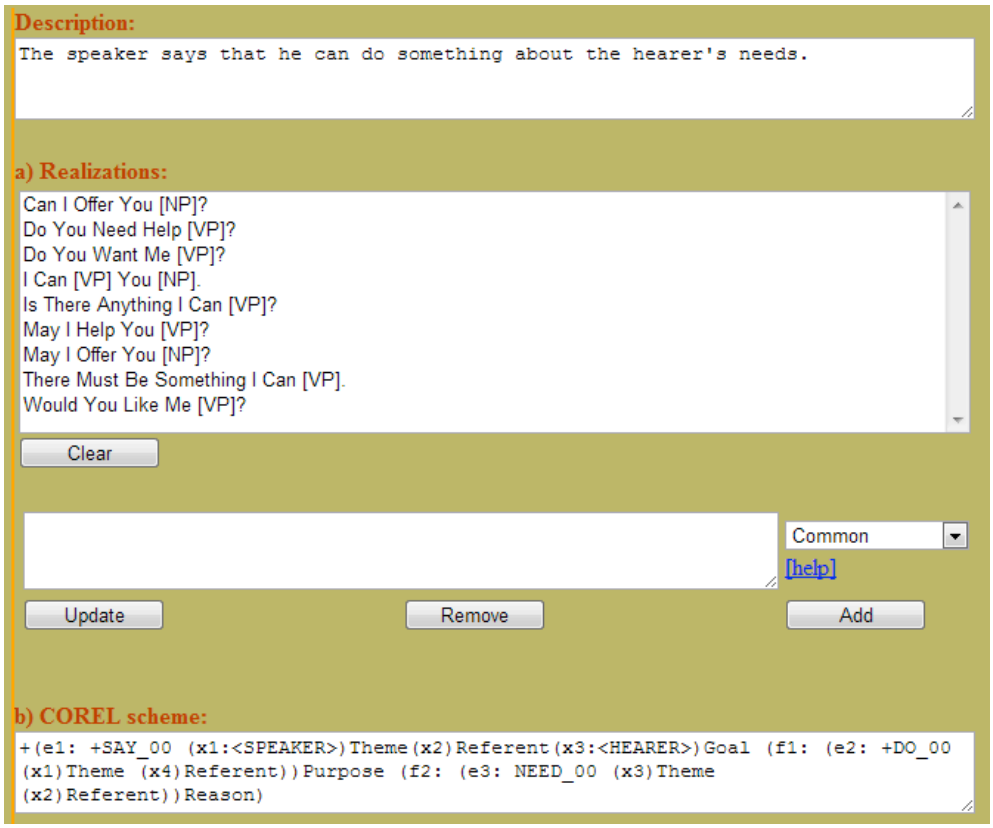

Figure 7. The interface of Offering constructions

The non-propositional meaning of Offering constructions, which is to be drawn from the COREL description shown in figure 7 above, is that the speaker tells the hearer that he (the speaker) can do something about the hearer's needs. The different constructions in the realization box convey the same illocutionary force, and therefore share the same COREL representation:

(4)+(e1:+SAY_00(x1:<SPEAKER>)Theme(x2)Referent(x3:<HEARER>)Goal (f1: (e2: +DO_00 (x1)Theme (x4)Referent))Purpose (f2: (e3: NEED_00 (x3)Theme (x2)Referent))Reason)

The predication in (4) can be translated as follows: the speaker (x1) tells something (x2) to the hearer (x3) with the purpose of doing something (x4) because the hearer is in need.

In the same way as with implicational constructions, the fixed part of these illocutionary constructions (e.g. 'Can I Offer You', 'Do You Need Help', etc.) constitutes the trigger that indicates that these are illocutionary constructions, so that ARTEMIS can recover the information from the COREL description. Furthermore, through pattern matching, ARTEMIS identifies the level-1 construction from which the CLS is to be imported. The information related to the realizations of the variable part is to be retrieved from the Lexicon or from the Level 1-Constructicon. Consider, for instance, the sentence *Can I offer you a drink?* In this case, ARTEMIS needs to go to the Lexicon in order to recover the necessary information about the element that has parametrized the X variable, i.e. 'a drink'. This information is contained within the lexical entry 'drink'. The lexical entry is linked to the basic concept +BEVERAGE_00 in the Ontology, from whose thematic frame ARTEMIS retrieves the conceptual information.

4.2.2. Promising

The constructions gathered under the label of Promising share the following description: the speaker says that he will do something for the hearer in the future. Find in figure 8 below the interface of Promising constructions in the Grammaticon, which may have a wide range of realizations:

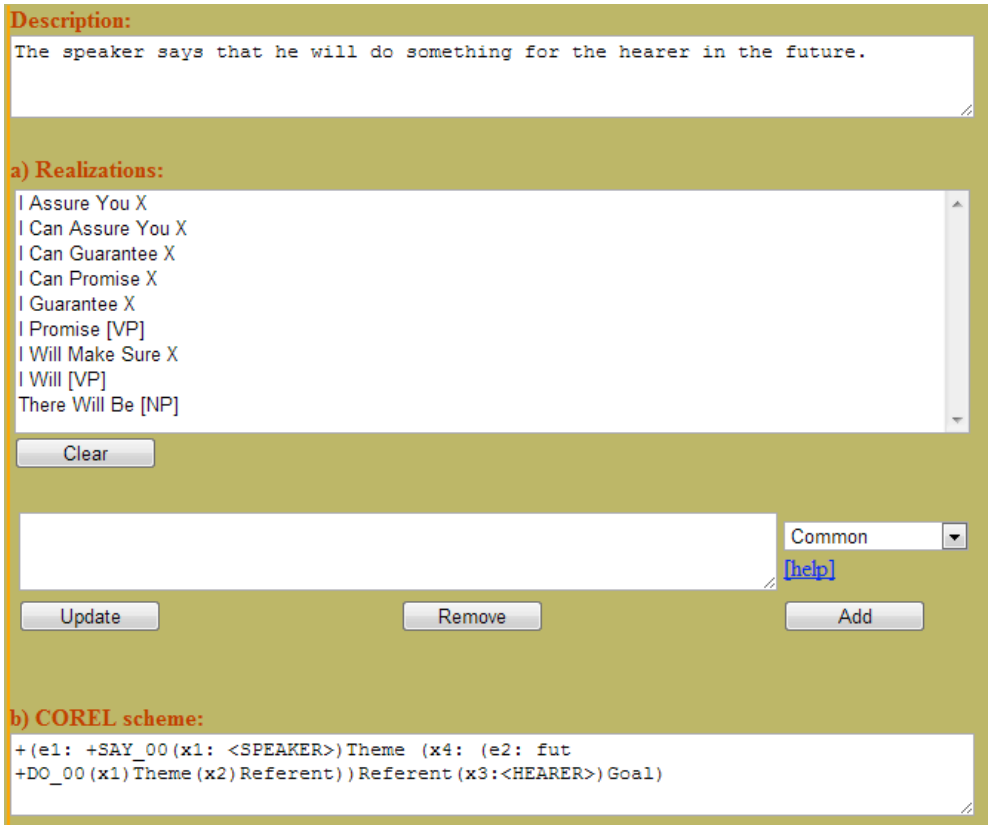

Figure 8. The interface of Promising constructions

Each of the instantiations of Promising constructions in the realizations box carries the semantic load represented in the following COREL schema:

(5) +(e1: +SAY_00(x1: <SPEAKER>)Theme (x4: (e2: fut +DO_00(x1)Theme(x2)Referent))Referent(x3:<HEARER>)Goal)

The translation of this complex predication is that the speaker tells the hearer that he will do something in the future. It goes without saying that the COREL description does not cover the whole range of meaning implications that we

can derive from the statement of a promise. For instance, the moral obligation of the hearer to comply with his promise is missed out. As we advanced at the beginning of this chapter, the conceptual information represented by COREL descriptions is far from equalling the conceptual information available in a person's brain. However, the essence of the meaning implication is captured in the semantic description.

4.2.3. Requesting

Requesting constructions consist in the speaker asking the hearer to do something for him. See figure 9 below, which presents the interface of Requesting constructions:

Figure 9. The interface of Requesting constructions

All the realizations shown above share the same semantic information, captured by the following COREL scheme:

(6) +(e1: +SAY_00 (x1:<SPEAKER>)Theme(x2)Referent(x3:<HEARER>)Goal (f1: (e2: +DO_00 (x3)Theme (x4)Referent))Purpose (f2: (e3: NEED_00 (x1)Theme (x2)Referent))Reason)

This COREL representation means that the speaker (x1) tells something (x2) to the hearer (x3) with the purpose of getting the hearer to do something (x4) because the speaker is in need.

As we pointed out in our discussion of Promising constructions, not every detail can be encoded into the COREL scheme. Some information, mostly related to social conventions, is left out. For instance, the degree of politeness of the request is not reflected in the COREL representation. The degree of politeness is what differentiates, for example, the *Can You X* from the *Could You Please X* construction. However, we believe that, in order to keep a balance between the degree of refinement of the linguistic description and the requirements imposed by the program, the difference would not justify the creation of different entries for each constructional realization.

4.3. Level 4-Constructicon

So-called discourse constructions are stored at this level of the Grammaticon. Discourse constructions enclose some sort of semantic connection between two utterances.

The representation of constructions within the Level 4-Constructicon follows the same pattern as the rest of idiomatic constructions. Here, the word(s) that link the two sentences forming the discourse construction constitute(s) the indicator for ARTEMIS to catalogue it as a level-4 construction. As with illocutionary constructions, discourse constructions have been gathered into groups in the Grammaticon following the same criteria as in illocutionary constructions; that is, constructions that share their non-propositional meaning (and thus share the same COREL scheme) are to be part of the same entry. In this case, the entries have been labelled according to the clausal relation that holds between the utterances, i.e. cause, condition, and consequence.

4.3.1. Cause

The most typical instantiations of cause constructions are the *X Because Y* and the *X Because Of Y* constructions, as shown in the realization box of the interface in figure 10 below:

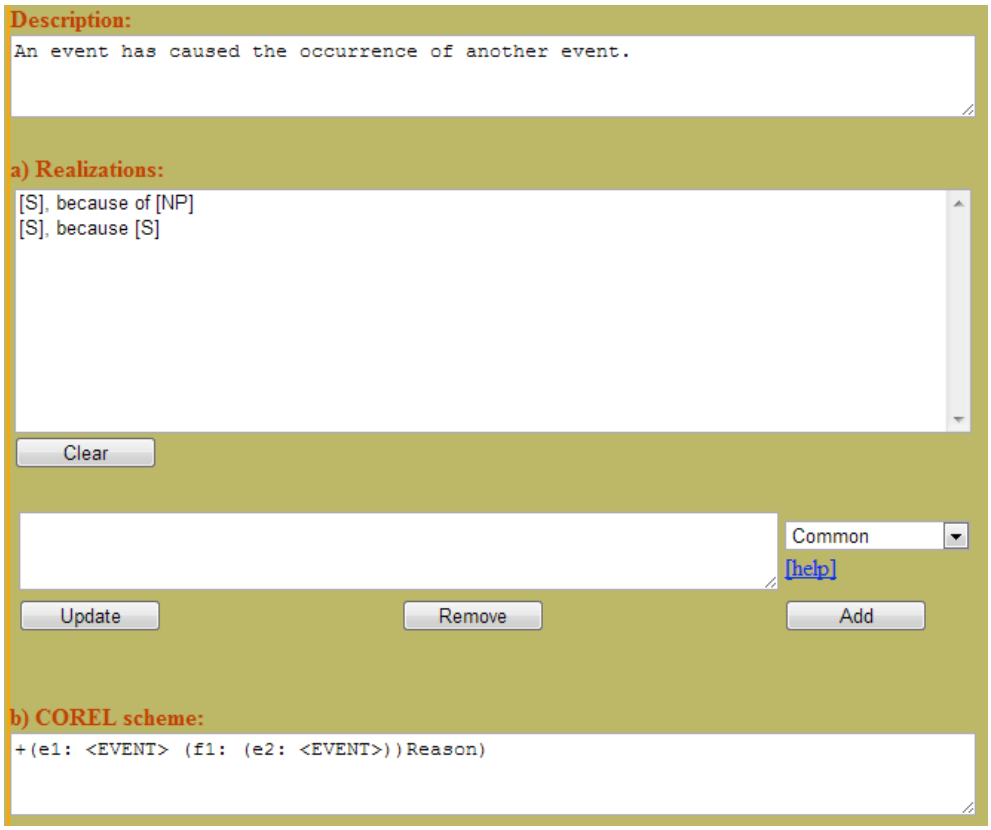

Figure 10. The interface of Cause constructions

The meaning of this construction is that an event has caused another event to occur. The X and the Y elements in the constructional realizations may be filled in by any sentences that bear a causal relation. The generic nature of this description is also reflected in the COREL representation:

(7) +(e1: <EVENT> (f1: (e2: <EVENT>))Reason)

In natural language, this COREL representation means that a given event has been caused by another event. An instantiation of this construction is to be found in the sentence *I went there because this restaurant was flagged in the*

slow food guide (cf. Chapter 7, section 7.2.4 for an analysis of this construction in terms of cognitive operations, i.e. saturation).

The fixed part of each construction ('Because' and 'Because of') are the activators for ARTEMIS to retrieve the information represented in this COREL scheme.

4.3.2. Condition

The interface in figure 11 shows the description, constructional realizations and COREL representation of condition constructions:

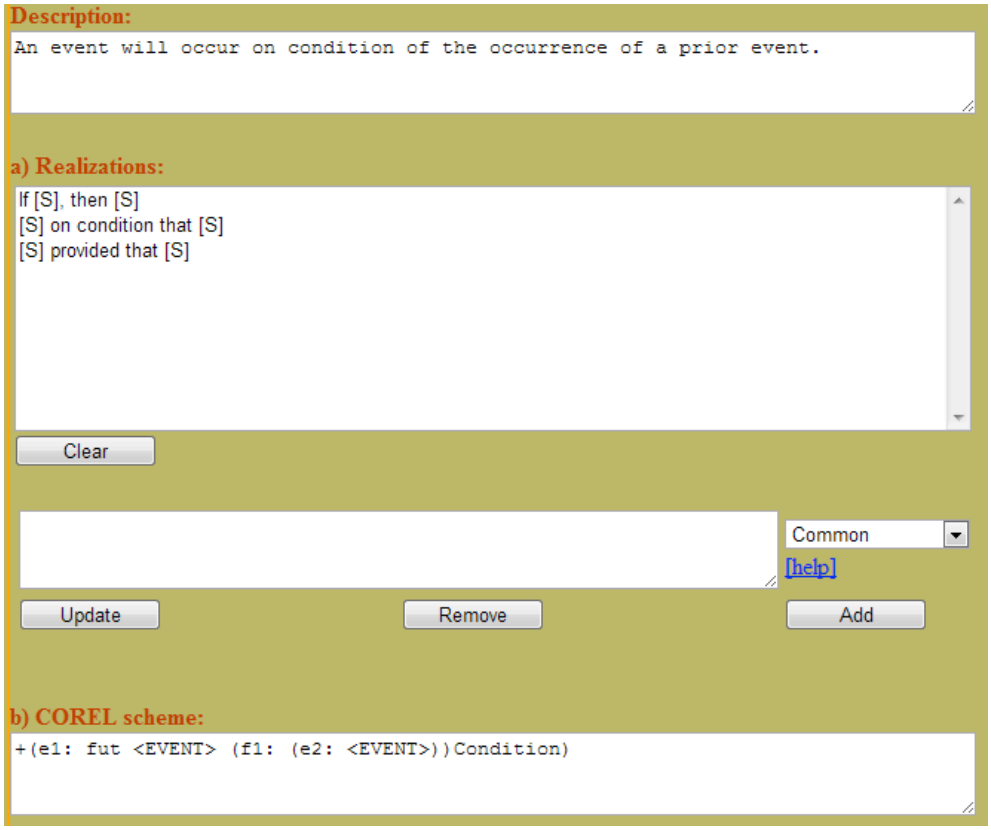

Figure 11. The interface of Condition constructions

As with cause constructions, the fixed elements of the constructional realizations determine their status as level-4 constructions for ARTEMIS. Again, the COREL representation is the same for all the constructions, because it captures the non-propositional meaning to be attributed to all of them:

(8) +(e1: fut <EVENT> (f1: (e2: <EVENT>))Condition)

This COREL scheme can be translated as follows: a future event will occur on condition of the occurrence of another event. An example of this construction, which encloses the semantic information encoded in the COREL representation, is the sentence *If you win, we'll let you walk with fifty more dollars than you got right now* (Chapter 7, section 7.2.5).

4.3.3. Consecution

The semantic relation of consequence dictates that the occurrence of an event results in the occurrence of a subsequent event (see Chapter 7, section 7.2.7). This description is captured in the interface presented in figure 12 below:

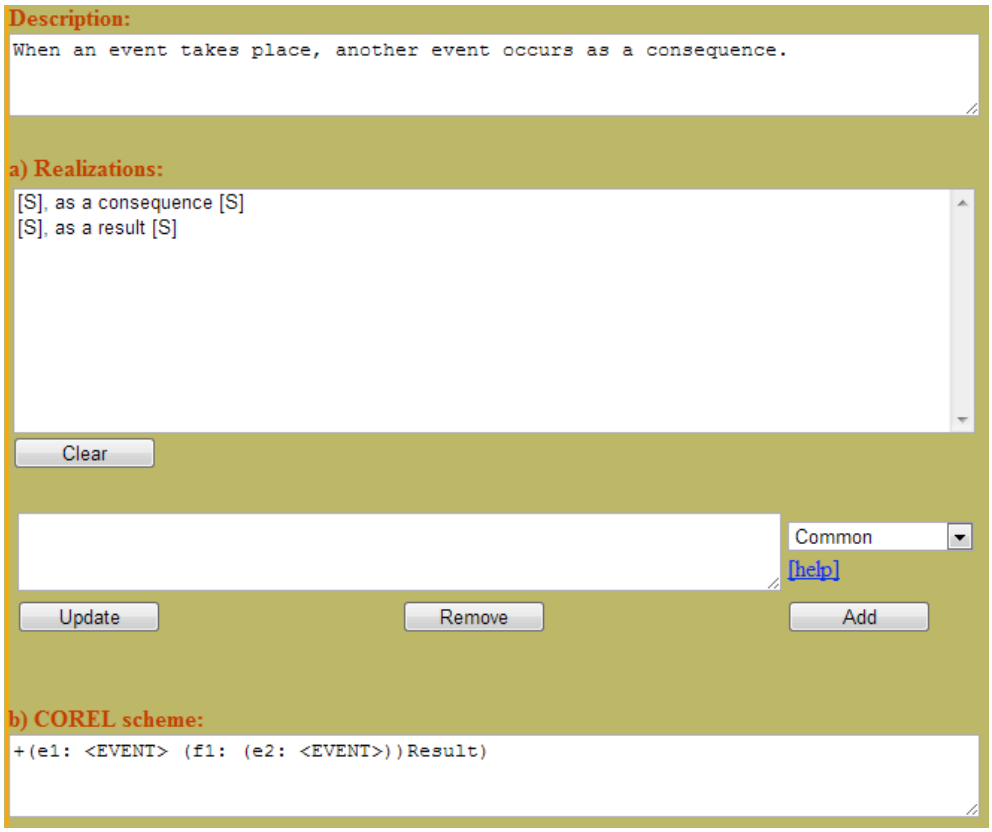

Figure 12. The interface of Consecution constructions

As we can see in figure 12 above, consecution constructions, some of which are presented in the realization box, share the following COREL representation:

(9) +(e1: <EVENT> (f1: (e2: <EVENT>))Result)

The COREL scheme presented above is to be interpreted as follows: the occurrence of an event has the occurrence of another event as a result.

As a concluding remark, we want to contend that, despite the limitations pointed out in our discussion, a broad-ranging fine-grained semantic representation of an input text containing not only lexico-conceptual but also constructional information, is possible in FunGramKB as an NLP system. We have shown throughout this chapter that the computer program ARTEMIS endows FunGramKB with a processing capacity that allows it to deal with non-propositional meaning, which is essential for the interpretation of natural language.