

## Matching Verbo-nominal Constructions in FrameNet with Lexical Functions in MTT

Myriam Bouveret

Université de Rouen & University of Berkeley\*

Charles J. Fillmore

University of Berkeley

*Matching verbo-nominal constructions in FrameNet with Lexical Functions in MTT*  
Multiword expressions are described in the FrameNet project as complex lexical units linked as wholes to their semantic frames, or as instances of special grammatical constructions. Amongst multi-word expressions, we explore in this paper verbo-nominal constructions in Framenet and more specifically support verb expressions. This category, described in FrameNet literature, has not yet received a systematic treatment in the project. We have looked at several problems regarding its actual classification and theoretical background. Instead of light verbs, or support verbs, we define the object of our study as Support Verb Constructions (SVC). We explore in this connection a well known model in lexicography, the Explanatory and Combinatorial Lexicology, part of the Meaning Text Theory to observe the feasibility of using Lexical Functions in FrameNet for the purpose of encoding SVC. The verbs have functions varying from that of light verbs, where they contribute only the ability to treat the noun's frame as a verbal-e.g., tense-bearing, entity-through a variety of aspect, perspective, and register values. These structures are treated in various ways in FrameNet and in the Explanatory and Combinatorial component of the Meaning-Text-Theory.

*Our goal in the present research is to understand the nature of these differences, and to consider whether the results obtained in one of them can be aligned with or incorporated into the other. We approach the parameters of comparison of the two systems by means of three themes characterizing verbo-nominal expressions:*

- *lexicalized collocations of the verb and the nominal head of its syntactic dependent,*
- *the manner in which the kind of situation-the semantic frame-evoked by the noun is given verbal expression,*
- *and the manner in which the syntactic arguments of the verb are interpreted as matching the semantic roles associated with the noun's frame.*

### 1. Introduction

This paper addresses a number of descriptive and theoretical problems surrounding the treatment of a class of verbo-nominal constructions, in French and English, connecting a governing verb and a dependent noun, in which it is the *semantic frame* encoded by the noun that contributes the main outlines of the meaning of the whole. The verb has functions varying from that of *light verb*, where it dispenses only the ability to treat the noun's frame as a verbal, e.g., tense-bearing, entity, through a variety of phasal, perspectival, and registral values. These structures are treated in various ways in FrameNet (Baker et al. 1998, Fillmore et al. 2003, Ruppenhofer et al. 2002), in Lexico-Grammar (Gross 1989 and 2004, Gross; De Pontonx 2005), and in the Explanatory and Combinatorial component of the Meaning-Text-Theory of Igor Mel'čuk (1998 and 2004) and Alonso Ramos (2007). Any theory of the lexicon and its

---

\* This paper was written while M. Bouveret was a visiting scholar at the Department of Linguistics, University of California in Berkeley

interaction with the grammar of a language needs to have something to say about the lexico-grammatical structure of sentences like (1a, 2a), and about its relation to the meaning of the sentences in (1b, 2b):

- (1a) She gave me wise advice.
- (1b) She advised me wisely.
- (2a) Les parents sont libres de donner leur consentement ou de refuser.
- (2b) Les parents sont libres de consentir ou de refuser.

Among the various systems that provide accounts of the phenomena represented in (1) and (2), the corpus-annotation-based FrameNet project (FN), and the lexical functions (LFs) part of the Explanatory and Combinatorial Lexicology (Alonso Ramos 2003, Kahane and Polguère 2001, L'Homme 2007, Mel'čuk 1974, 1996 and 2003, Polguère 2003, Wanner 1996, Zolkovskij and Mel'čuk 1967) have treatments of such phenomena that are built into quite different descriptive frameworks. Our goal in the present research is to understand the nature of these differences, and to consider to what extent the results of the one can be aligned with or incorporated into the other. We approach the parameters of comparison between the two systems by way of the three themes identified in (A, B, C). Both approaches are interested in characterizing, for verbo-nominal constructions,

- (A) lexicalized collocations between the verb and the nominal head of its syntactic dependent,
- (B) the manner in which the kind of situation (the semantic frame) evoked by the noun is given verbal expression, and
- (C) the manner in which the syntactic arguments of the verb are interpreted as matching the semantic roles associated with the noun's frame.

We note in sentences (1a) and (2a) that a collocational relation exists between the verb *give* and the noun *advice* (eng.), the verb *donner* and the noun *consentement* (fr.). We note too that the event recorded in the phrases *gave... advice* and *donner... consentement* is the event named by the nouns *advice* and *consentement*. And we note that the syntactic arguments of the verbs *give* in (1a) and *donner* in (2a) stand for the adviser and the advisee in the advising act<sup>1</sup>.

There are many cases for which, in both languages, the properties listed in (A, B, C) are all simultaneously present, as in sentences (1a) and (2a). There are, however, expressions that satisfy (B) and (C) but not (A), as in (3a) and (4a); there are expressions that satisfy criteria (A) and (C) but not (B), as in (3b) and (4b); and there are some that satisfy (A) and (B) but not (C), as in (3c) and (4c). *Breaking a promise* is a separate act from *making a promise* (i.e., from *promising*), though it remains clear that the person who breaks a promise is the one who did the promising; and the intransitive verbs *blow* (eng.) and *sonner* (fr.) when predicated of *wind* (eng.) and *la cloche* (fr.) do not contribute additional syntactic arguments that participate in the meaning of the whole.

- (3a) She carried out the work.
- (3b) You broke your promise.
- (3c) The wind was blowing.
- (4a) Il fait une ligne bien nette sur le papier.
- (4b) Ils ont trahi leur promesse.
- (4c) La cloche sonne.

A key element to understanding the difference between the two systems FN and LFs in regard to these structures is that the theory of lexical functions characterizes the properties of the verbs by establishing relations between deep and surface syntactic and semantic representations, whereas the FrameNet project first characterizes the semantic *frames* associated with the nouns

---

<sup>1</sup> Such information could be expressed entirely within a noun-headed phrase, as in *your advice to me*, allowing the predication of some evaluation, say, of that event: *your advice to me was wise*.

and then seeks to discover the phrases that can be construed as expressing the elements of those frames—the so-called *frame elements*—. In the cases that interest us, we find that some of the frame elements of a noun are realized as dependents on a verb that governs it.

## 2. The evolution of FrameNet’s treatment of verbo-nominals

The treatment of verbo-nominal constructions in FrameNet is related to the fact that the database is a corpus-based project in which certain kinds of information are derived automatically from manually created annotations, as summarizations of the properties of sentences annotated to exemplify the valencies of individually targeted lexical units. In FrameNet work, each target is identified as a member of a given semantic *frame* (Fillmore 1982, Fillmore and Atkins 1992), and a set of layers is dedicated to displaying the properties of that word; for example, a sentence like *The manager decided to cancel the meeting* could be annotated in separate layers with respect to any of its four frame-bearing words, *manager*, *decide*, *cancel*, and *meeting*. In an annotation of this sentence centred on *decided*, a FrameNet annotation would declare *the manager* to be the decider and the phrase *to cancel the meeting* to stand for the act that got decided. In the annotation layer for that same sentence centred on *cancel*, the annotation would regard *the meeting* as standing for the thing that got canceled, and would recognize that, by virtue of the grammatical requirements of the verb *decide*, *the manager* has to be construed as the subject (hence the “agent”) of *cancel*. The general practice is to associate a word with a particular *semantic frame* and then to look around, in the grammatically defined context of that word, to identify those phrases that represent participants in that frame.

One desideratum of FN annotation was to be able to collect minimal structured lexical sets, called kernel dependency graphs (syntactic trees in MTT), associated with particular predicating words: examples, presented linearly, are *{teacher, Demonstrate, proof, to-students}*, *{teacher, Examine, student, in-geometry}*, *{success, Depend, on-effort}*, *{manager, Cancel, meeting}*. The governors are capitalized, the other members of each set are the dependents. The dependent elements of such structures had to be given frame element names and grammatical function names, together with indications of how the dependent elements were “marked”, e.g., with specific prepositions. Finding the frame elements in arguments following the verbs does not raise any problems (i.e., where the relevant phrases are found “within the phrasal projection” of the predicating word—inside the VP for a verb, inside the NP for a noun—), but the problem was how to label the things that were found “to the left” of the target but which still could be seen as having semantic and structural relations to the frame-bearing lexical unit. Occurrence as the syntactic subject of a finite verb occurring immediately before the verb (*those teachers Taught geometry*), or the argument of a nominalization realized as a possessive modifier (*the president’s Decision to resign*) are clear cases. But in order to refer to phrases in other contexts that can be taken as satisfying the “subject”-like valent<sup>2</sup> of the predicate, a pseudo grammatical function EXT (“External argument”) was proposed. Thus, in *the president seems likely to resign*, the subject of *seems* is chosen as the EXT of *resign*. For simplicity, it was decided to use this grammatical-function name for the subject relation in general, whether locally or distantly realized: thus, *the president* would be tagged as the EXT of *seems*, and *likely*, as well, in the annotations centered on those words.

The verb *forget* is the controlling context that determines that *teacher* in (5) is the noun that belongs in the subject position for *demonstrate*, and the adjective *likely* is the context that allows *success* in (6) to be assigned that position relative to *depend*. Familiar theories of grammatical “control”—with certain cautions—justify such decisions.

---

<sup>2</sup> In this discussion we use valent to refer to either the syntactic constituent that is dependent on a frame-bearing lexical unit, or the semantic role that may or may not be realized as a syntactic constituent. Thus “semantic valent” has roughly the meaning of “actant” in the Tesnière tradition but includes more than the props and actors that occur in the *petit drame* of a sentence.

(5) The teacher forgot to *demonstrate* the proof to the students.

(6) Success is not always likely to *depend* simply on effort.

But verbs are not the only words that have valencies. When looking to “populate” kernel dependency graphs that would show the valency of nouns—like *decision*, *journey*, *complaint*, *operation*—it appeared that when looking for the EXT entities, there were often very specific verbs that “controlled” such identities, specific in the sense that they seemed to be individually selected by the nouns. For example, *make* is the verb most likely to introduce the EXT argument of the event noun *announcement*, and similar relations hold in phrases like *take a journey*, *have a relationship*, *give a demonstration*, *take a bath*. FrameNet made the decision to record such verbs in entries for the associated nouns, and called them, following long tradition, *support verbs* (SUPP) in these contexts. Hence, an entry for the noun *announcement* would include the information that its main support verb is *make*. Other kinds of facts required new decisions and sometimes changes of terminology, regarding support verbs, verbalization of event nouns that had no arguments (*the wind is blowing*), the functions of the copula and prepositions (see Atkins 2003), alternating perspectives (*the doctor performed an operation* vs. *the patient underwent an operation*), and other semantic features like non-fulfilment, inchoation, and duration (*break a promise*, *launch a project*, *continue a war*)<sup>3</sup>.

Recording the ways in which the valence requirements of a word were discovered in a sentence often depended on examining, not just the immediate syntactic dependents of an LU, but also those contexts that linguistically determined “argument sharing” relations between some contextual entity and the lexical unit being described. In the case of verbs, no special attention was paid to the sources of such argument sharing, since every grammatical system is provided with some theory of control by which, for example, the direct object of *persuade* is co-construed as the subject of its infinitival complement. In the case of frame-bearing nouns, however, the annotators noticed that some of the frame-requirements of certain nouns were satisfied through the syntactic arguments of verbs that occurred in construction with the noun. For example, there’s something wrong with a sentence like *\*I gave him advice to the student to enroll in the course*, since the object of the verb provides information about the recipient of the advice, blocking the need to include such information within the phrase headed by the noun. Unlike with the case of verbs, this time one couldn’t count on the grammar as a whole to explain such matchings, since the choice of the governing verb often depended on specific lexical properties of the noun. (*Persuade* doesn’t care what verb heads its infinitival complement.) It seemed important, therefore, to indicate the existence of such “support structures” as information about the LU itself. FrameNet workers became aware that such verb-noun pairings had parallels in the theory of lexical functions in the work of Igor’ Mel’čuk and his colleagues, but since all of the frame elements that participated in this relationship were included in the annotation, no information was lost—at least in the case of true light verbs.

As this work continued, a number of distinctions were taken note of. In some cases, the verb that determined a frame element of the noun was not a strict collocate of the noun. The primary semantic valent of *performance* is present as the subject of *resume* in *After the accident the musicians resumed the performance*, but this function of *resume* can’t be recorded as a lexicographically relevant property of the noun *performance*. The subject of *need* in *I need an operation* is construed as the secondary valent of the frame evoked by *operation*, but again, that verb is not linked to the noun as a true collocate. When it seemed necessary to include such facts in the annotation of individual sentences, the verb was marked as a *Controller* (CONTR). Further distinctions among verbs that were labeled *support* came to FN’s awareness, but their analysis was postponed. In the clearest cases, the verb was a true “light verb” and its co-occurrence with the noun added no information beyond the kinds of temporal, aspectual or modal information that goes with being a verb (*make an announcement*, *say a prayer*, *take a*

<sup>3</sup> At first these were referred to as support verbs by generalizing the concept; then were recognized as non-support *controlling* verbs because of the control or argument-sharing functions. Supports are also controllers in this sense, but the new category was for controllers that were not supports.

*shower*). But other cases provided more information: aspects of register and tone distinguish *make a complaint*, *lodge a complaint*, *file a complaint*, *register a complaint*, and the like, and FN had no way of recording such information. Sometimes temporal-schema information was included in the verb itself, as seen in the difference between *conduct an investigation* and *launch an investigation*. Still different situations appeared in the contrast between *make a promise* and *break a promise*; the collocational force is present in both cases, but talk about breaking a promise presupposes a promising event in the past—and an event in which the person who broke the promise is necessarily understood as the person who made that antecedent promise: it's not possible to break a promise made by somebody else.

Since FN is a large database founded on corpus evidence, it has seemed worthwhile in some FrameNet centres to attempt an alignment of the support structures found in FN data with descriptions of lexical functions that associate nouns with the words that can “verbalize” them.

### 3. Support verbs in the Meaning-Text Theory

The lexical functions in the Meaning-Text Theory (MTT) come at these issues from a different direction<sup>4</sup>—namely that of treating some words and constructions as elements of the value of “functions” that operate on particular words and yield larger phrases of various kinds (*fast friendship*, *strong tea*, *school of fish*)—. Collocations described by lexical functions within the phraseology of a language are described as *semi-phraseemes*. *Collocations*<sup>5</sup> are semantically compositional while *full phraseemes* (idioms) are not (for a fully detailed description see for e.g. Mel'čuk 1998 or 2003). The meaning of collocations encoded by means of LFs might easily be guessed but they could not easily be produced by a non-native speaker.

---

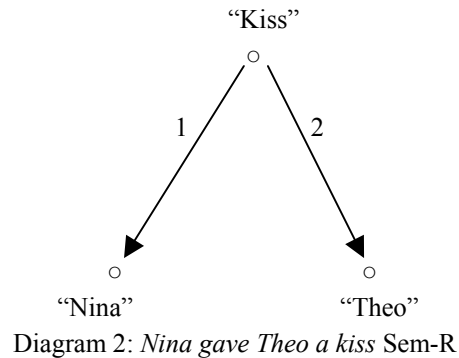
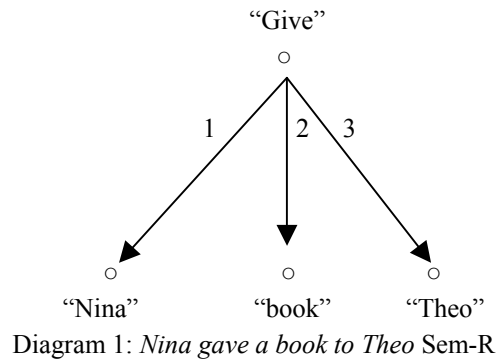
<sup>4</sup> The work on lexical functions began in 1962, in connection with Machine Translation: one had to formulate the rules to translate algorithmically such English words as *heavy* (as in *heavy drinker/smoker*) or *make* (as in *make trouble/love*), which feature hundreds of Russian translation equivalents. The first steps: Mel'čuk noticed a systematic dependency on the syntactic Governor or Dependent for the choice of translation; the LFs *Magn* and *Oper<sub>1</sub>* were discovered. The second step: Zholkovsky showed that there was an opposition between *Magn* and *Ver* and between *Oper<sub>1</sub>* and *Real<sub>1</sub>*. The third step: working together, Zolkovskij and Mel'čuk 1967 was produced; this was a fully developed presentation of LFs (translated and published in French in 1970, in *T.A. Informations*) (note from I. Mel'čuk).

<sup>5</sup> Multiword expressions are described in the FrameNet project (Ruppenhofer et al. 2002) as complex lexical units linked as wholes to their semantic frames, or as instances of special grammatical constructions (Atkins et al. 2003). To date FrameNet has not sought to include compound nouns, but mainly because the general annotation scheme is centred on recognizing a lexical unit as a target of analysis with its dependents labeled in terms of the semantic roles they satisfy in the frame linked to the target. In many cases, however, the modifier noun in a compound serves as a frame element of the head noun, as when we recognize that in *bear skin*, the noun *bear* satisfies a requirement of *skin*, just as the *peace* in *peace treaty* satisfies the topic indication in the frame to which *treaty* belongs. A class of “transparent nouns” recognized in FrameNet has been proposed for recognizing certain kinds of selection phenomena: in particular, for certain constructions of the type “N of N” in which the second noun is more relevant for answering to selectional preferences than the first noun. Thus what when eating a *type of fish*, one is eating a *fish*, and collocations like *wreak vengeance* are not violated by a phrase like *wreak this kind of vengeance*. The nouns that figure in such constructions identify types, portions, aggregates, and quantities. Some of these have important collocational force, such as *a pod of whales*, *a pride of lions*, *a flock of geese*, etc., but many others, such as *type*, *sort*, *kind*, are neutral. A new process for annotating grammatical constructions is currently being developed at FrameNet, whereby one can recognize simultaneously that *wheel chair* is a noun of a particular semantic type and that it has a substructure which is itself made up of two nouns. Verb-headed multiwords are treated as complex single lexical units, and among them are verb plus particle combinations (*give up*, *give out*), support constructions of various kinds (*give notice*, *give care*), and idiomatic expressions (*give a damn*, *give someone the slip*).

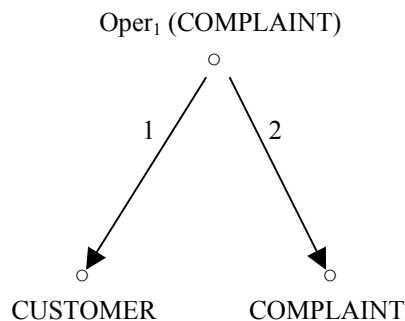
Lexical functions are known in lexicography as expressing a particular set of phrasemes in a language, the collocations, but they do not only encode collocations in fact. LFs more generally modelize semantic links on both the syntagmatic and the paradigmatic axes. “A lexical function is a function that associates with a specific lexical unit (L) which is the argument or keyword of  $f$ , a set of synonymous lexical expressions, the value of  $f$ ” (Mel’čuk 1998: 32). We focus our attention in this presentation on standard lexical functions of support verbs: [*Oper*, *Func*, *Labor*]. Each verb is established in relation to the syntactic role of the lexical unit in the verbo-nominal construction: subject, object, indirect object, as these are linked to their deep syntactic actants I, II, and III. The most frequent case is a support verb whose object is a nominal derivative. In the collocation *give a kiss*, *give* is the support, *kiss* is the object and *kiss* is the deverbal noun, which is represented:  $Oper_1$  (kiss) = *give*. This set of three verbal lexical functions allows one to modelize differences of perspective and differences of meaning beyond to the simple concept of *light verb* encoded with [*Oper<sub>1</sub>*] but are still named support verbs in MTT. The verb *receive* stands as [*Oper<sub>2</sub>*] for the noun *kiss* ( $Oper_2$  (kiss) = *receive*) since it changes the point of view from the first actant of the predicate *kiss* to the second actant, at the deep syntactic level the person who gives the kiss is I and the person who receives the kiss is II. [*Oper*, *Func*, *Labor*] are empty verbs but the set of corresponding realization functions [*Real*, *Fact*, *Labreal*] are not: they express a realization that is inherent in the noun. For example, the noun *promise* could collocate with its support verb *make* ( $Oper_1$  (promise) = *make*) or with the verb expressing the fulfillment meaning: *keep* ( $Real_1$  (promise) = *keep*). Support verbs (or the corresponding realization verbs) can also be compounded into configurations, in this case, the configuration of several lexical functions forming a configuration of LFs (see section 4) add semantic components to the support verb which is not empty anymore. Those semantic components are linked to the meaning of the noun in the verbo-nominal construction so that in an MTT dictionary, the definition is of primary importance. A complex lexical function is a set of Ls on the same key word, formed of simple lexical functions that are syntactically linked (i.e.  $IncepOper_1$  (love) = *fall in*). A *complex lexical function* has a global value which expresses the meaning of the whole. In a complex LF, all participating LFs have the same keyword and are syntactically related to each others; in a *configuration of LFs*, all participating LFs have the same keyword but are not syntactically related to each others. In the following configuration for example,  $AntiBon+Magn$  (sorrow) = *deep*, the LFs *AntiBon* and *Magn* are not syntactically related to each others so that this configuration is marked by the sign + differentiating configuration from complex LFs.

#### 4. Correspondences between semantic representation, Deep-Synt and Surface-Synt levels

In an MTT model of language, the semantic representation is linked to the Deep-Syntactic level and the Surface-Syntactic Level. A support verb lexical function appears at the Deep-Syntactic level. As a general rule, semantic actants match with syntactic actants, but paraphrasing rules are specific to each support verb construction. In such a collocation, the verb itself, being empty, has no semantic actant. In the *make a complaint* phrase for example, actants are semantically linked to the noun (the complaint on the part of X addressed to Y). The noun *complaint* as a predicate controls the semantic actants, the verb is not chosen for its semantic content since it is light. The verb is lexically constrained in a language at the S-Synt level. The situation is different from that of a plain verb in the following sentence for example: *Nina gives a book to Theo*. Without any ambiguity in this sentence, *give* is analyzed as the governor whilst *Nina* and *Theo* are semantically dependent as agent (addresser) and beneficiary (addressee), and the second actant *book* as an object while in *Nina gives her brother a kiss*, it is the noun predicate *kiss* that bears the actants, with *Nina* as an agent (addresser) and *her brother* as a patient (addressee) as shown in the following diagrams 1 and 2:



The support verb has a meaning as a lexeme, but may be deprived of it in the support verb construction (i.e. *give Theo a kiss* in diagram 2) so that the semantic structure does not represent it. The support verb may have a meaning in the collocation (i.e. *commit a crime*, *send a kiss*, *blow a kiss*), however it is not chosen for this meaning contrary to the non-support verb use (i.e. *give a book to Theo* in diagram 1) but for its role of supporting the syntactic realization of the noun in a given language. This characteristic would be enlightened in FrameNet by the fact that *give a book* in diagram 1 would refer to the *giving* frame while *give a kiss* in diagram 2 would refer to the frame encoded by the noun. In the support construction, syntactically the noun, selectionally the verb, is the dependent member of the pair which explains that the verb in the semantic representation does not appear since the noun is the predicate bearing the semantic actants. But a support verb does still behave like a verb, it can bear mood, modality, temporality for example in the support verb construction but still remains a support verb (i.e. *Nina did not send Theo a kiss*). In the syntactic representation then the verb and the noun have to be present, the verb is the governor and the noun is syntactically dependent at the deep syntactic level. We present the D-Synt trees of *make a complaint* in *The customer made a complaint [about the last delivery]*, and *face a complaint* in *Paul faced a complaint [about the last delivery]* in the following diagrams 3 and 4.



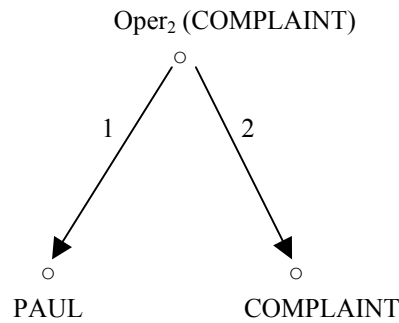


Diagram 4: *Paul faced a complaint [about the last delivery]* at the D-Synt level

A solution specific to support verb construction (SVC) is proposed in Ramos 2007 “since the support verb is empty and cannot have any semantic actant, its syntactic actants must be in correspondence with the semantic actants of the predicative nouns” and the “support verb adapts the diathesis of the noun to that of the verb” (Ramos 2007: 111). In MTT, DSyntSs are linked by rules of a universal paraphrasing system, which covers all possible DSynt-paraphrasing in all languages (for more, see e.g. Milicevic 2007).

In FrameNet terms, the noun evokes the appropriate frame in the support construction. The selection of the support verb is sometimes, but not always, sensitive to semantic properties of the frame: there is usually no predictable matching between semantic properties of the nouns in a frame and the choice of the appropriate support verb, but there are generalizations such as the preference for *make* with nouns involving acts of speaking: *declaration, announcement, statement*, etc.

### 5. Summarizing possibilities of integrating support verb lexical functions in FrameNet

In MTT the verbal collocate is described as an element of the value of the lexical function operating on the noun. In FrameNet, these verbs have in general been noted only during the lexicographic study of the nouns and have not been studied independently. In MTT, they are not studied independently either but are only described with respect to the noun, which is the keyword in the support construction. The following table illustrates the possible matching of support verb lexical functions with frames and frame elements (FEs):

1. Support verb lexical functions: [Oper, Func]	2. Related lexical functions	3. Lexical units inheriting LF column 1	FE: Core elements	Frame
<i>Oper<sub>1</sub></i> (complaint) = make	<i>Oper<sub>2</sub></i> (complaint) = face, get, receive  <i>V<sub>0</sub></i> (complaint) = complain	assertion, comment, complaint, concession, confession, declaration, exclamation, proclamation, remark, statement, allegation, announcement, speech, promise	Medium, Message, Speaker, Topic, Addressee	Statement Frame



$Oper_1$ (promise) = make	$Real_1$ (promise) = keep $AntiReal_1$ (promise) = break $V_0$ (promise) = promise	appointment, commitment, promise, threat, vow	Addressee, Medium, Message, Speaker, Topic	Commitment Frame
$Func_0$ (ceremony) = take place	$IncepFunc_0$ (ceremony) = begin $FinFunc_0$ (ceremony) = end	baptism, ceremony, communion, confession, wedding, mass, ordination, rite, sacrament, sacrifice	Desired State, Guardian, Instrument, Leader, Member, New Status, Object, Organization, Type	Rite Frame
$Oper_2$ (exam) = take $Oper_1$ (exam) = give	$Real_2$ (exam) = pass $AntiReal_2$ (exam) = fail $V_0$ (exam) = examine	exam, test	Examination, Examinee, Examiner, Knowledge, Qualification	Examination Frame

Table 1: Support Verb Lexical Functions matching Frames

The first column illustrates the set of support verb lexical functions taking the noun as an object [ $Oper$ ], a subject [ $Func$ ]; examples of [ $Labor$ ] encoding lexical unit as the second object are less frequent ( $Labor_{12}$  (interrogation) = submit to). Examples of support verbs matching strictly light verbs are represented with [ $Oper_1$ ]. Other examples show either a difference of perspective [ $Oper_2$ ], or a difference of syntactic function concerning the noun in the verbo-nominal construction [ $Func_0$ ]. In MTT though, they are still called “light” verbs since they do not have their own plain meaning in the support construction. Beyond true support verb constructions, the table shows other functions in column two, in order to cover and systematize relations actually shown as Support and Controller annotations in FN. Realization verbal lexical functions [ $Real, Fact, Labreal$ ] and additional component [ $Magn$ ] expressing intensity are also needed. Phasal meaning [ $Incep, Fin, Cont$ ] expressing the different phases of an event, its beginning, continuation and end ( $ContLabreal_{12}$  (mind) = keep in,  $Magn+ContFunc_0$  (war) = rage on) and causative meaning [ $Caus$ ] would also help to take care of various situations sometimes “abusively” described in FN as Support (e.g., *keep* and *break* + nouns of commitment *vow, promise, oath, appointment*). Semantic nominal derivatives [ $S_1, S_2, S_3$ ], syntactic derivatives [ $S_0, V_0$ ], other paradigmatic relationship encoding synonyms and contrary [ $Syn, Anti$ ], and conversives [ $Conv$ ] could add transparency to relations between lexical units in a frame. Several other lexical functions expressing typical circumstances: instrument, medium, mode, result and localization [ $S_{instr}, S_{med}, S_{mod}, S_{res}, S_{loc}$ ] request to raise the question of adjuncts and circumstances within frames. Those simple or complex lexical functions are to be accompanied by actants numbers that would help matching valency in FN.

## 6. Conclusion

In this paper we have studied verbo-nominal constructions in FrameNet and have considered what it would mean to import lexical functions, part of the Meaning-Text Theory, into the FrameNet project in order to give a systematic and more lexicographically transparent way of organizing those constructions. We have shown that in principle information about lexical functions could be imported into the FrameNet database, but there are technical and methodological reasons why that would be fairly difficult. Although FrameNet and MTT have in common analyses of the kinds of verbo-nominal structures that have concerned us in this paper, the backgrounds, terminology, and manner of approach are importantly different. The system of lexicographic annotation used for FrameNet was designed for capturing certain

distributional facts associated with frame-bearing lexical units, but not all. Summaries of recorded facts, based entirely on the annotations, lexical unit by lexical unit, are reported in the FrameNet database as lexical entries that show how the elements of the LU's "frame",—both "core" and "peripheral" frame elements—were syntactically realized in the sentences that actually got annotated. The reports are generated automatically, from the annotations, and there is no provision for generalizations.

Introducing LFs into a FN database would require a non-trivial modification of the FN procedures. One possibility would be to allow a large number of V+N pairs to be extracted from the noun entries and entered as multiword LUs in their own right, explaining in the text of the lexical entries, information about the nature of the argument-sharing as well as information about the semantic or pragmatic contribution of the verb. This would be extremely labor-intensive: current lexical entries in FrameNet are created automatically, where the annotator needs only to insert a brief definition or mnemonic for the LU. Another possibility would require creating lexical entries (and hence appropriate annotations) for each of the verbs that can participate in support constructions, but that would require a departure from the frame-based nature of FrameNet research, since these verbs could not be relevantly studied frame-by-frame.

Regarding support verb annotation with lexical functions or corpus retrieval, the project is at its beginning. As for a general matching of both models, questions regarding argument sharing, deep syntactic structure of the verbo-nominal construction, actants and arguments of the predicate, adjuncts and circumstants, definitions in MTT compared to core and peripheral components in a frame, still need to be analyzed carefully since the two models have many theoretical differences despite their proximity. Taking care of those questions is also part of the ongoing project.

### **Acknowledgements**

We thank Igor Mel'čuk for his collaboration, his patience and suggestions. Members of the FrameNet team and particularly Collin F. Baker, Michael Ellsworth, Russell Lee-Goldman, Birte Lönnecker-Rodman, Miriam R. L. Petruck, Carlos Subirats, are warmly thanked for their comments and interaction on previous drafts of the work. François Gaudin and Laurent Gosselin are also thanked for their reading and advice.

## References

- Alonso Ramos, M. (2007). "Toward the synthesis of support verb constructions: distribution of syntactic actants between the verb and the noun". In Wanner, L. (ed.). *Selected Lexical and Grammatical issues in the Meaning-Text Theory*. Amsterdam: John Benjamins. 93-134.
- Alonso Ramos, M. (2003). "Eléments du frame vs. Actants de l'unité lexicale". In *Proceedings of the First Meaning-Text Theory, MTT 2003, Paris, 16-18 juin 2003*. 77-88.
- Atkins, S.; Fillmore, C. J.; Johnson, C. R. (2003). "Lexicographic Relevance: Selecting Information From Corpus Evidence". *International Journal of Lexicography* 16 (3). 251-280.
- Baker, C. F.; Fillmore, C. J.; Lowe, J. B. (1998). "The Berkeley FrameNet Project". In *COLING-ACL 98: Proceedings of the Conference, held at the University of Montreal, Montreal*. 86-90
- Fillmore, C. J. (1982). "Frame semantics". In *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Company. 111-137.
- Fillmore, C. J.; Atkins, S. (1992). "Towards a frame-based organization of the lexicon: The semantics of RISK and its neighbors". In Lehrer, A.; Kittay, E. (eds.) *Frames, Fields, and Contrast: New Essays in Semantics and Lexical Organization*. Hillsdale: Lawrence Erlbaum Associates. 75-102.
- Fillmore, C. J.; Johnson, C. R.; Petruck, M. R. L. (2003). "Background to Framenet". *International Journal of Lexicography* 16 (3). 235-250.
- Fontenelle, T. (1992). "Using a bilingual computerized dictionary to retrieve support verbs and combinatorial informations". *Acta Linguistica Hungarica* 41 (1-4). 109-121.
- Gross, G. (1989). *Les constructions converses du français*. Genève, Paris: Librairie Droz.
- Gross, G. (dir.) (2004). "Introduction" *Linguisticae Investigationes* 27 (2). Amsterdam, New York: John Benjamins Publishing Company. 167-169.
- Gross, G. ; De Pontonx, S. (dir.) (2005): "Verbes supports: nouvel état des lieux". *Linguisticae Investigationes* 24 (2). Amsterdam: John Benjamins Publishing Company.
- Kahane, S.; Polguère, A. (2001). "Formal Foundation of Lexical Functions". In *Workshop proceedings: "COLLOCATION: Computational Extraction, Analysis and Exploitation, 39th Annual Meeting and 10th Conference of the European Chapter of the Association for Computational Linguistics"*. 8-15.
- L'Homme, M. C. (2007). "Using Explanatory and Combinatorial Lexicology to Describe Terms". In Wanner, L. (ed.) *Selected Lexical and Grammatical Topics in the Meaning-Text Theory*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 163-198.
- Mel'čuk, I. (1974). *Opyt teorii lingvističeskix modelej "Smysl-Tekst". Semantika, Sintaksis.— Esquisse d'une théorie des modèles linguistiques du type "Sens-Texte". La sémantique, la syntaxe*. Moscow: Nauka. 314.
- Mel'čuk, I. (1996). "Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon" In Wanner, L. (ed.) *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam, Philadelphia: John Benjamins Publishing Company. 37-102.
- Mel'čuk, I. (1998). "Collocations and Lexical Functions". In Cowie, A. (ed.) *Phraseology. Theory, Analysis and Applications*. Oxford : Clarendon Press. 23-53.
- Mel'čuk, I. (2003). "Collocations dans le dictionnaire". In Szende, T. (réd.). *Les écarts culturels dans les dictionnaires bilingues*. Paris: Honoré Champion. 19-64
- Mel'čuk, I. (2004): " Verbes supports sans peine". *Linguisticae Investigationes* 27 (2). 203-217.
- Milicevic, J. (2007). *La paraphrase*. Bern: Peter Lang.
- Polguère, A. (2003). "Étiquetage sémantique des lexies dans la base de données DiCo". *Traitement Automatique des Langues (TAL)* 44 (2). 39-68.

- Ruppenhofer, J.; Baker, C. F.; Fillmore, C. J. (2002). "Collocational Information in the FrameNet Database". In Braasch, A.; Povlsen, C. (eds.) *Proceedings of the Tenth Euralex International Congress. Copenhagen, Denmark*. Vol. I. 359-369.
- Wanner, L. (1996). "Introduction". In Wanner, L. (ed.) *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam: John Benjamins. 1-36.
- Žolkovskij, A. K.; Mel'čuk, I. (1967). "A. O semantičeskom sinteze—Sur la synthèse sémantique". In *Problemy kibernetiki* 19. 177-238. [Three translations of the paper exist: **a.** in Hungarian: *Dokumentáció és nyelvészet*, Budapest, 1969, 61-176; **b.** In English: *Systems Theory Research*, N.Y., v. 19; **c.** In French: *T.A.Informations*, 1970, n° 2, 1-85.]