The Scope and the Sources of Variation in Verbal Predicates in English and French

Tanja Samardžić, Lonneke van der Plas, Goljihan Kashaeva, Paola Merlo
University of Geneva
Linguistics Department, Rue de Candolle 2, 1204 Geneva
E-mail: {Tanja.Samardzic, Lonneke.vanderPlas,
Goljihan.Kashaeva, Paola.Merlo}@unige.ch

Abstract

In this article, we examine the variation of the predicate-argument structure between English and French in an experimental and data-driven approach. We annotate a corpus of 1000 French sentences with predicate-argument structure using a framework originally developed for English. We identify a number of non-matching predicates and examine them in a qualitative analysis. We show that these two languages do not differ substantially in the inventory and the nature of verbal predicates, even though certain general grammatical properties may result in some variation at this level of representation. We argue that the proposed comparative study can provide a basis for identifying the level of specificity needed for developing a framework for multilingual annotation of predicate-argument structure.

1 Introduction

The analysis of the predicate-argument structure of a sentence results in a linguistic representation that defines relations between the constituents in the sentence that cannot be defined by the rules of syntax only.

Cross-linguistically, the predicate-argument structure of a sentence is considered to be more stable (less varying) than its syntactic form. The English sentence in (1a) can be considered as equivalent to the French sentence in (1b), despite the fact that the positions of their syntactic subjects are occupied by different kinds of lexical elements and that the complements of the verbs differ both syntactically and semantically. The predicate-argument structure of (1a) and (1b) is equivalent since verbs (*liked*, *a plu*), which are the predicates of the sentences, take the same kind of arguments (EXPERIENCER, CONTENT) in both languages.

(1) a. [EXPERIENCER Mary] liked [CONTENT the idea]. (English)

b. [CONTENT L'idée] a plu [EXPERIENCER à Marie]. (French)

199

Not all predicates behave in the same way in a cross-linguistic context. Some of them are known to give rise to more parallel or cross-linguistically stable syntactic structures, while others result in more divergent syntactic realizations. For instance, the predicates expressed by verbs meaning "creation", such as English *create, build, construct* are likely to be realized as transitive verbs with direct objects in many different languages, while predicates denoting some state of mind, such as the verbs in (1), can be involved in different syntactic structures in different languages [13].

Linguistic investigations into predicate-argument structure have been mostly concerned with the kinds of labels (semantic or thematic roles) that can be assigned to the arguments. In addressing cross-linguistic variation, these approaches assume that the same predicates are associated with the same argument structure in different languages. Their focus is on the differences in syntactic realisation of the arguments and the semantic features that give rise to the variation. Little attention has been given to the cross-linguistic variation in the argument structure itself. However, a corpus-based analysis of the variation between English and Chinese [5] suggests that cross-linguistic parallelism in the predicate argument structure cannot be assumed. This study shows that 17% of arguments of English predicates do not map to the arguments of the corresponding Chinese predicates.

The aim of our study is to examine the variation of the predicate-argument structure between English and French predicates in an experimental and data-driven approach. We perform a qualitative analysis of divergent predicates to identify general principles underlying the variation. We argue that the proposed comparative study can provide a basis for establishing the level of specificity needed to develop a framework for multilingual annotation of predicate-argument structure.

We identify the varying predicates by annotating the predicate-argument structure in a corpus of one-thousand naturally occurring French sentences using an annotation framework originally developed for English. Cases in which the English annotation framework fails to be applicable to French predicate-argument structure are selected as strong examples of cross-lingual variation. Such an approach rests on two important methodological decisions.

First, we compare two closely related and well-documented languages. This choice is based on the assumption that English and French constitute a minimal pair suitable for micro-comparisons [9]. Since the variation in the predicate-argument structure is related to lexical items rather than to broader syntactic structures, micro-comparison seems to be the adequate approach, providing a good setting for identifying the elements of grammar that potentially underlie the variation.

Second, we take a corpus-based approach analysing a big sample of naturally occurring instances of predicates using an annotation scheme as a framework. This decision is motivated by the fact that studying linguistic variation requires analysing large amounts of data. Furthermore, the results of such an analysis can be directly incorporated in developing or improving tools for automatic natural language processing, especially automatic analysis of the predicate-argument structure.

2 Choosing the annotation framework

There are three current frameworks proposed for annotating corpora with predicate-argument structure: FrameNet [2], VerbNet [11], and PropBank [16]. All of them have been developed on the basis of English data and have been used to annotate English corpora. However, these frameworks are implementations of linguistic theories of the predicate argument structure that have been developed to account for universal phenomena, which is why they can be expected to apply to other languages as well.

All three resources describe the same linguistic phenomena covering approximately the same range of lexical items. Nevertheless, they are conceived with different purposes and with different theoretical backgrounds. As a result, the choice of the framework can have important consequences for the outcome of the annotation, especially in a cross-linguistic setting.

Predicate labels used in FrameNet (frames) and VerbNet (verb classes), are intended to capture the level of lexical semantics which is common to a group of lexical items. Different words can bear the same label. For instance, the verbs *accomplish, achieve, bring about* all bear the label *Accomplishment* in FrameNet. This is not the case in the PropBank framework, where predicate labels capture the specific meaning of verb senses. This is why the style of annotation used in FrameNet and VerbNet can be considered as more abstract and, thus, more portable across languages than the annotation in PropBank. Indeed, FrameNet has often been used as a basis for developing similar resources for other languages [4].[1] Furthermore, argument labels in FrameNet are assigned without taking into account the syntactic function of the constituents that bear them, while the PropBank argument labels can depend on the syntactic function. This feature makes the FrameNet annotation less tied to the syntactic representation and thus to a particular language than it is the case with the PropBank annotation.

In our study, however, we use the PropBank framework to annotate the predicate-argument structure of French sentences. We take this decision for two reasons.

First, the lexicon in this resource is built by extracting and describing all the predicates that occur in a predefined sample of naturally occurring sentences. Since our aim is to annotate exhaustively a corpus of naturally occurring sentences, we expect that such a lexicon provides a better coverage than the lexicon in FrameNet, which is not developed with a corpus-driven methodology.

Second, the labels used in PropBank both for predicates and arguments involve fewer theoretical assumptions than the labels in FrameNet. While FrameNet labels capture mostly linguistic intuition about the targeted level of lexical semantics and the relations between the lexical items, the PropBank labels rely strongly on the observable behaviour of words. The distinctions between verb senses, for instance, are made taking into account the differences in subcategorisation frames of each

---

[1]VerbNet is a relatively new resource and it is rarely used for automatic labelling of semantic roles.

sense. This approach can be expected to provide more tangible criteria for annotators in deciding how to annotate each instance of the predicate-argument structure found in the corpus, ensuring a more reliable and more consistent annotation.

3 Materials and methods

In identifying the points of variation in the predicate-argument structure between English and French, we rely on corpus data. We manually annotate a corpus of one-thousand French sentences using the PropBank annotation framework originally developed for English. Those French predicates for which no appropriate PropBank label (English verb sense) can be found are considered as varying. These sentences are automatically extracted from the corpus and manually analysed. We normalised the instances of predicates extracted to their citation form and looked up a comprehensive French-English dictionary to identify the English translation of the French predicate that is closest to the literal translation. If needed, the translations are verified with a native speaker. For example, we find the English expression *shed light* as the translation for the French expression *faire la lumière*. Finally, we classify and analyse the types of varying predicates, discussed in Section 4.

We envisage to detect cross-linguistic variation by looking at those cases where the manual annotation using a resource developed for a source language (English) fails to be applicable to the target language (French). It must therefore be ensured that failure is due to a genuine inability to provide a common analysis for the two languages and not to other causes. To ensure good quality and reliability of the annotation, we provided the annotators with detailed guidelines, that consist of an extensive description of the PropBank annotation framework (adapted from the original PropBank guidelines [1] to the French language), parallel English and French examples of predicate-argument analysis, and a manual on how to use the annotation tool[2]. Furtermore, we offered the annotators intensive training. The main annotation is preceded by a two-stage training phase, and a calibration phase, following [14]. More details on the annotation procedure can be found in [17]. For our current investigations into divergences in predicate argument structure, it is important to stress that the task of using an English resource to annotate French predicates is well-defined. This is reflected in high inter-annotator agreement (an average F-score of 95%) after discussion and individual re-annotation. In the calibration phase, when annotators were not able to discuss the annotations, an average F-score of 81% is reached. This score is calculated after reducing the very fine-grained PropBank verb sense labels to more general verb class labels with the predicate-argument structure preserved. We can assume that failure to annotate French predicates using the English resource is due to a genuine inability to provide a common analysis for the two languages and not to other causes.

The French corpus that we annotated consists of 1040 sentences drawn from

---

[2]The annotation tool we use is an adaptation of the user-friendly, freely available Tree Editor (TrEd, [15]).

the French portion of the Europarl corpus [12]. This corpus contains translations of the proceedings of the European Parliament in 11 languages parallelised at the sentence level.[3] We use 40 sentences for the training, so that the final French corpus annotated with predicates and arguments contains 1000 sentences.

4   Non-matching predicates

We find 90 instances of predicates that could not be annotated for a total of 1985 predicates in the 1000 sentences of the annotated corpus. This ratio indicates that a great majority of French predicates found in our corpus directly corresponds to an English verb sense with the same predicate-argument structure. The PropBank lexicon provides a better coverage for our corpus than the FrameNet lexicon for annotating a German corpus of a similar size [4], where 30% of instances of German predicates required adding a label to the existing FrameNet set.

The mismatching predicates are mostly conventionalised expressions with different degrees of semantic compositionality, including light verb constructions, collocations, and fully uncompositional idioms.[4]

In this section, we group and analyse the non-matching predicates in an attempt to identify the categories that are potential sources of variation. We classify the predicates according to the degree to which they differ from the closest English counterparts.

The degree of parallelism between the corresponding expressions in the two languages is defined in terms of the cross-linguistic stability of the lexical categories and syntactic properties of the lexical items involved in the expressions. On the basis of the assumption that the lexical category is not an inherent feature of a lexical item ([7]), we consider items such as the English verb *need* and the French noun *besoin* as matching, even though they are associated with different lexical categories in the two languages. By syntactic properties we understand the subcategorisation frame of a lexical item, specifying the number and form of its complements.

The French structures that are most parallel with English are those where the corresponding lexical item preserves its category across languages, and the arguments of the predicates preserve their syntactic form (Group 1). Expressions that include predicates that require different syntactic forms of their arguments in the two languages are considered less parallel (Group 2). Group 3 includes the expressions where both the category of corresponding lexical items differs and syntactic

---

[3]Even though we use the French side as a monolingual corpus in our study (the annotators do not have access to the actual translations), we take the sample from a parallel corpus because this gives us a possibility to look up the actual translations of the expressions we are interested in while analysing the variation. In addition, manual annotation of one side of a parallel corpus can be a useful resource for experiments in automatic annotation transfer.

[4]Our analysis does not address idioms (3 out of the 90 instances). We also exclude one mismatch which is clearly not due to linguistic reasons (the PropBank lexicon did not contain the English verb *dramatise* for the French verb *dramatiser*). This leaves us with 86 instances that are discussed.

divergences take place. Lastly, for a number of French predicates and their English counterparts it was not possible to find any lexical correspondence. These predicates are classified as Group 4.

4.1 Characteristics of the non-matching predicates

The first property which can be noticed about the non-parallel transitive constructions is that they tend to be headed by very frequent verbs with rather general meaning. The most characteristic verb for these constructions is the verb *faire* which occurs in all four types of mismatches. It can have two kinds of meaning. One of its meanings can be described as vaguely causative, denoting that its argument which is syntactic subject (ARG0 in the PropBank annotation) brings about the entity denoted by its object (ARG1), e.g. *faire pression* : *put pressure*. The other meaning is copulative, denoting that the argument which is its syntactic object (ARG1) denotes some property of the other argument (ARG0), e.g. *faire l'objet de* : *be the objective*. The meaning of the other verbs used in these constructions can be described in these terms as well.

The arguments considered as ARG1 in these constructions are headed by abstract nouns, including deverbal nouns, such as *pression* and *objectif* in the examples above. The meaning of these nouns is generalised with no specific reference, which is often reflected in the fact that they are not preceded by an article.

There are two dominant ways in which French transitive constructions are transformed into the corresponding English expressions. The transitive verbs of the groups 1 and 2, where the lexical categories are preserved, correspond either to an English verb with the same properties (general meaning of the same type), or to the copulative *be*. Transitive expressions of the group 3 are typically transformed into a single English verb, if they involve a verb with vaguely causative meaning. Otherwise, they are transformed into a copulative construction. Expressions of group 4 are mainly transformed into a single English verb.

The mismatching expressions with intransitive verbs as predicates are less numerous than those with transitive verbs. The examples found in our corpus provide a basis for identifying two factors that can make intransitive verbs in French hard to match with English verbs. We note that most of the examples are pronominal verbs (*s' exprimer, se mettre d'accord, se réjouir, se féliciter, se prononcer*), where the reflexive particle does not bear a semantic role. Another group of predicates are impersonal verbs (*il suffit, il convient*), which do not assign a semantic role to their syntactic subject and whose use is limited to a single morphological form (3rd person singular). Some predicates are characterised by both features (*il s'agit de, il se peut que*). We note that these predicates are most often transformed into a copulative construction in English.

There are two expressions with intransitive verbs that are neither pronominal nor impersonal (*tenir de, tomber bien*). Their transformation to English forms follows the patterns identified with other predicates. The expression *tenir de* corresponds to English copulative constructions, while the English form that corre-

sponds to *tomber bien* is another verb (*come*), but with the same kind of general meaning. Both French *tomber* and English *come* are unaccusative verbs of motion.

4.2 Overcoming the cross-linguistic divergence

The structures associated with French predicates identified as not matching with English predicates during the annotation still retain a certain level of parallelism with the corresponding English structures. In most of the cases, it is precisely the number and the meaning of the predicates' arguments that remain unchanged. Variation can be identified at other levels of the structure. It is limited to two domains: the choice of the lexemes which encode the impoverished meaning of the predicates and the choice of lexical category of the predicating lexical unit.

For cross-linguistic mapping of the predicate-argument structures which involve variation in one of the two domains, the representation of the structure needs to be slightly more abstract than the one that is currently used in PropBank. In this section, we propose the representations of the structures which are valid cross-linguistically introducing only minimal generalisations needed to address the observed variation.

4.2.1 Parallel structures: Group 1

In the case where the heading verbs are not translations of each other, while the rest of the structure is preserved across languages, we propose assigning a special label to the verbs. Since the meaning of these verbs is impoverished in these usages, no specific verb sense label can account for it. The label on these verbs needs to express their general meaning and it needs to be applicable to multiple verbs. Such a representation is given in (2).

(2) a. [ARG0 Nous] [REL-CAUS **tirons**] [ARG1 les leçons du passé]. (French)

   b. [ARG0 We] [REL-CAUS **learn**] [ARG1 the lessons of the past]. (English)

The following predicates belong to this group:

avoir lieu : take place; faire alliance : form an alliance;
avoir rien à voir : have nothing to do; faire pression : put pressure;
attirer l'attention : draw attention; faire appel : make an appeal;
céder la place : make room; tirer la leçon: learn the lesson
faire la lumière : shed light;

The semantically impoverished verb in French is assigned a more abstract label CAUSE. This label marks two characteristics of the predicate at the same time: the fact that it is not an ordinary predicate and its general meaning. All the impoverished verbs in our sample have the same general meaning, vague cause, but other labels could be used for other general meanings that can potentially occur.

This approach would require a set of abstract labels to be defined within the framework of PropBank in addition to the verb sense labels.[5] These labels would resemble the FrameNet labels, since they would express an abstract layer of the meaning of the verbs and since they would be applied to multiple lexical units. The difference is that these labels would encode a more abstract meaning than it is the case with the FrameNet labels. Also, they would be used only for a very limited set of lexical items in special contexts where their meaning is impoverished.

4.2.2 Almost parallel structures: Group 2

A case of almost parallel structures is represented in (3), where the French pronominal verb corresponds to the English transitive verb. A parallel representation of this sentence requires treating the pronominal clitic in French (*m* in (3a)) as one of the verb's arguments and assigning it the same label as to the corresponding argument in English (ARG1 in (3b)).

(3)   a. [ARG0 Je] [ARG1 **m'**] [REL-EXPRESS.01 **exprime**] [ARG2 sur le texte de M. L]. (French)

        b. [ARG0 I] [REL-EXPRESS.01 **express**] [ARG1 **my opinion**] [ARG2 on the Mr L.' s document]. (English)

The meaning of the ARG1 in the English sentence can be interpreted as non-specific or general, but this has no consequences for the annotation framework, since the argument labels are already general.

Apart from the intransitive verb given in (3), this group includes the following transitive verbs:

faire l'objet de : be the objective;
laisser qqn. sceptiqe : be sceptical;
tenir compte de : take into account;

avoir qqch comme objectif : be the objective;
donner à qqn. à penser : make sbd. think;

4.2.3 Category changing structures: Group 3

The non-parallel structures that involve a change in the lexical category of the predicating word require a stronger deviation from the standard PropBank representation. Most of these cases involve a predicate that is expressed synthetically in one language (as a single verb), while it is expressed analytically in the other language (as a combination of a verb and a predicating complement). The corresponding lexical units in these cases are the synthetic verb in one language and the verb's complement in the analytical expression in the other language. The following predicates belong to this group:

---

[5]Efforts in this directions have already been made [8]

Transitive:
avoir besoin de : need;
avoir trait de : have to do;
avoir retard : be late;
donner suite à : follow up;
y faire face : face;
faire état : state;
faire défaut : lack;
mettre qqch en exergue : highlight;
porter remède : remedy;

prendre conscience de : be/become aware;
rendre compte : be accountable

Intransitive:
se mettre d'accord : agree;
se réjouir : be happy;
il suffit : be enough

A possible way of representing the cross-linguistic parallelism in the predicate-argument structure of these expressions is to label the corresponding lexical items with a predicate label, leaving the heading verb of the analytical expression without a label, as illustrated in (4).

(4) a. [ARG0 Le marché] **a** [REL-NEED.01 **besoin**] [ARG1 de règles]. (French)
     b. [ARG0 The market] [REL-NEED.01 **needs**] [ARG1 rules]. (English)

The verbal head of the French analytical predicate (*a*) is not assigned a predicate label. Instead, the predicate is the noun heading its complement (*besoin*), which is, at the same time, lexical counterpart of the verb in English (*need*). The unlabelled verb in this case would be treated as a functional word which has no semantic arguments — a lexicalised light verb ([7], [6], [10]).

Such representation is not possible in the current PropBank setting due to two limitations. First, all the verbs except the copulative *be* are considered as predicates. Leaving a non-copulative verb without a label would go against this principle. Moreover, strictly functional treatment of light verbs is subject of debate in the literature, with some authors arguing in favour of such approach for certain verbs ([6], [10]) and others providing evidence of semantic content of light verbs [3]. Second, only verbs are considered as predicates in PropBank which is why the frames are not specified for other lexical categories. This means that an existing verbal frame in the English resource would need to be adapted to be used with a French nominal predicate. The limitation is even more important in predicates which are expressed in an analytical form in English, such as *be enough* in (5b), corresponding to the French *suffit* in (5a). The argument structure of the English adverb enough is not specified at all in PropBank's frame files.

(5) a. Il [REL-ENOUGH **suffit**] [ARG0 de lire le programme]. (French)
     b. It **is** [REL-ENOUGH **enough**] [ARG0 to read the manifesto]. (English)

4.2.4 Expressions with no direct translations: Group 4

Finally, the cases identified as the strongest mismatches in our classification can be resolved applying one of the described approaches too, as illustrated in (6).

The only difference is that the English label will not be a direct translation of the French word in question. This, however, should not pose a problem for the result of annotation. The label does not have to be a direct translation. It only needs to be consistent (and consistently found in a corpus) with the word.

(6) a. [ARG0 Les dirigeants politiques] doivent **faire** [REL-SHOW.01 **preuve**] [ARG1 d' un réel courage]. (French)

b. [ARG0 Political leaders] must [REL-SHOW.01 **show**] [ARG1 real courage]. (English)

This group includes the following predicates:

Transitive:
faire preuve de : show;
se faire un plaisir : be happy;
mettre qqch en oeuvre : implement;
remettre qqch en état : repair;
porter atteinte : undermine;
prendre la parole : speak;
prendre la peine : bother;
assigner qqn. à résidence : put sbd. un- der house arrest;

metre qqch. en cause : call stg. into question

Intransitive:
il s'agit de : be about;
il convient : should;
se féliciter : be pleased;
il se peut que : it is possible that;
se prononcer : give an opinion;
tenir de : be;
tomber bien : come in a good moment

The analysis of the mismatching expressions with lexical categories preserved across languages (groups 1 and 2) suggests that using a special label for verbs with impoverished meaning would ensure a more adequate and cross-linguistically valid annotation framework.

The analysis of the category changing predicates suggests that the lexical category of the predicating lexical items can change across languages, while the argument structure is still preserved. Therefore, specifying predicate-argument structures for all predicating lexical items, including verbs, nouns, and adjectives is important for making an annotation framework directly portable across languages.

5 Conclusions

A translation is an implicit representation of the meaning of a sentence. Applying the same annotation scheme for different languages is a way of making explicit the components of meaning that are expressed in a predicate.

The general linguistic conclusion of our investigation is that most mismatches concern the lexical realisations of the predicates and not the argument structure itself. Even for the cases identified as mismatches in our study, the number and the kind of arguments are unchanged across languages. The mismatches are due to the

different choices the two languages make with respect to lexical items that realise the meaning of semantically impoverished verbs, and to the lexical category of the predicating word. Our analysis also suggests that the structural correspondence fails more when a clause contains a verb with less specific meaning. This correlation between specificity of meaning and structural realisation is a confirmation of the cross-linguistic validity of those approaches, such as [13], that hypothesize a direct correspondence between the components of meaning of a predicate and the syntactic realisation of the arguments.

Applying a specific annotation framework such as PropBank in a cross-linguistic study proved plausible. A great majority of French predicates found in our corpus (95 %) corresponded directly to an English verb sense specified in PropBank. Our findings about the sources of the mismatches in the remaining predicates indicate that using a small set of special labels for verbs with impoverished meaning and providing annotation for all categories of predicating words (verbs, nouns, and adjectives) would improve the empirical adequacy and cross-linguistic validity of this framework.

Acknowledgements

References

[1] Olga Babko-Malaya. Propbank annotation guidelines. 2005.

[2] Collin F. Baker, Charles J. Fillmore, and John B. Lowe. The Berkeley FrameNet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, pages 86–90, Montreal, Canada, 1998. ACL / Morgan Kaufmann Publishers.

[3] Claudia Brugman. Light verbs and polysemy. *Language Science*, 23:551–578, 2001.

[4] Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal. FrameNet for the semantic analysis of German: Annotation, representation and automation. In Hans Boas, editor, *Multilingual FrameNets in Computational Lexicography: Methods and applications*, pages 209–244. Mouton de Gruyter, 2009.

[5] Pascale Fung, Zhaojun Wu, Yongsheng Yang, and Dekai Wu. Learning bilingual semantic frames: Shallow semantic parsing vs. semantic role projection.

In *11th Conference on Theoretical and Methodological Issues in Machine Translation (TMI 2007)*, pages 75–84, Skovde, Sweden, 2007.

[6] Jane Grimshaw and Armin Mester. Light verbs and theta-marking. *Linguistic Inquiry*, 19:205–232, 1988.

[7] Kenneth Hale and Samuel Jay Keyser. On argument structure and the lexical representation of syntactic relations. In Kenneth Hale and Samuel Jay Keyser, editors, *The View from Building 20*, pages 53–110. MIT Press, 1993.

[8] Jena Hwang, Archna Bhatia, Claire Bonial, Aous Mansouri, Ashwini Vaidya, Nianwen Xue, and Martha Palmer. Propbank annotation of multilingual light verb constructions. In *In Proceedings of the 4th Linguistic Annotation Workshop (The LAW IV)*, pages 82–90, Uppsala, Sweden, 2010.

[9] Richard Kayne. *The Oxford Handbook of Comparative Syntax*, chapter Some notes on comparative syntax, with special reference to english and french. Oxford University Press, 2005.

[10] Kate Kearns. Light verbs in English. Manuscript, 2002.

[11] Karin Kipper Schuler. *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, University of Pennsylvania, 2005.

[12] Philipp Koehn. Europarl: A parallel corpus for statistical machine translation. In *Proceedings of MT Summit 2005*, Phuket, Thailand, 2005.

[13] Beth Levin and Malka Rappaport Hovav. *Argument realization*. Cambridge University Press, Cambridge, 2005.

[14] Sebastian Padó. *Cross-Lingual Annotation Projection Models for Role-Semantic Information*. PhD thesis, Saarland University, 2007.

[15] Petr Pajas and Jan Štěpánek. Recent advances in a feature-rich framework for treebank annotation. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 673–680, Manchester, UK, 2008.

[16] Martha Palmer, Daniel Gildea, and Paul Kingsbury. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–105, 2005.

[17] Lonneke van der Plas, Tanja Samardžić, and Paola Merlo. Cross-lingual validity of propbank in the manual annotation of french. In *In Proceedings of the 4th Linguistic Annotation Workshop (The LAW IV)*, pages 113–117, Uppsala, Sweden, 2010.