Intelligent Information Systems 9999
ISBN 666-666-666, pages 1–10

Adjectives in TreeLex

Anna Kupść

Université de Bordeaux / ERSSàB / SIGNES / IPIPAN

Abstract

The paper presents an ongoing project on automatic extraction of subcategorization frames of French adjectives.[1] The extraction method is guided by syntactic corpus annotations and heuristic rules. In particular, we use general linguistic knowledge to distinguish constituents inherent to impersonal, comparative and certain adverbial constructions from subcategorized complements.

1 Introduction

TreeLex is a subcategorisation lexicon of French, automatically extracted from a syntactically annotated corpus (a treebank). Initially, we used the treebank to extract only verb valence, Kupść (2007). In this paper, we describe our ongoing work on automatic extraction of subcategorisation frames of French adjectives.

Subcategorisation lexicons, i.e., resources which store information about syntactic combinatory potential of a predicate, play a crucial role in various NLP applications, related both to parsing, e.g., Briscoe and Carroll (1993), Carroll and Fang (2004), Surdeanu *et al.* (2003), and generation, e.g., Danlos (1985), Han *et al.* (2000). For French, just like for all other languages, such resources have been mostly developed for verbs, applying diverse methods ranging from time-consuming but detail-oriented work of human experts, cf. Gross (1975); Guillet and Leclère (1992); Mel'cuk *et al.* (1984, 1988, 1992, 1999), to various recent automatic techniques: Bourigault and Frérot (2005), Chesley and Salmon-Alt (2005), Gardent *et al.* (2006), van den Eynde and Mertens (2003), Sagot *et al.* (2006). We can mention two European research and development initiatives, concerning French (among others), which resulted in creating valence lexicons: EAGLES (Genlex, Menon and Modiano (1993)) and LE-PAROLE (Ruimy *et al.* (1998)). The projects were focused on providing a general multilingual architecture and creating multilingual resources and were not specifically devoted to developing syntactic lexicons.

Valence lexicons for other types of French predicates are scarce: Gross (1986) contains information on subcategorisation frames of nouns, adjectives and adverbs but it has not been adapted for automatic text processing, whereas a syntactic lexicon of French prepositions, which can be used for NLP purposes, has been

[1] I am indebt to Anne Abeillé and Jesse Tseng for their help on this project. I also wish to thank two anonymous reviewers for their comments which guided the improvements made in the final version of the paper.

only recently created by Fort and Guillaume (2007). In this paper, we present our work on building a syntactic lexicon of French adjectives adjusted to NLP applications.

2 Method

The approach we have adopted explores syntactic annotations in a corpus. We used the treebank of Paris7, Abeillé *et al.* (2003), a journalistic corpus based on articles from *Le Monde* (1989–1993), a French daily newspaper. The corpus contains about one million words with morphological, syntactic and, for major constituents, also functional annotations. Grammatical functions are indicated only for verbal dependents. Hence, unlike for verbs, valence of adjectival predicates is not directly specified in the corpus. The procedure we applied to obtain adjective valence is presented below.

2.1 Adjectives in the Treebank

The treebank contains three levels of annotation: 1) morphosyntactic: the category, lemma and morphological features are associated with every (simple or compound) word, 2) syntactic: all major syntactic constituents, including adjective phrases (AP), are indicated in the corpus, and 3) functional: direct verb dependents have their grammatical functions assigned.

The treatment of adjectives in the corpus is not uniform. In NPs, simple prenominal attributive adjectives are not considered APs and have only morphosyntactic (i.e., word-level) tags, (1). Even if several simple adjectives precede a noun, they are not grouped as an AP, (2). On the other hand, if a prenominal adjective is modified by an adverb or has a complement, it does form an AP with its dependent, (3). All postnominal or predicative adjectives are annotated as APs, even if the adjective appears alone, (4). Additionally, since predicative adjectives are direct dependents of the verb (the copula), they are assigned a grammatical function, ATS (fr. *attribut sujet*), which indicates a predicate related to the subject (e.g., *valable* in (4)). Finally, we found certain constituents headed by adjectives labeled with a different tag; for example, the superlative construction in (5) is annotated as an NP rather than an AP, even though no noun is present. Note that the superlative form followed by a noun is treated as an AP in (3).

(1) [$_{NP}$ la [$_{A}$ moindre] [$_{N}$ réforme]] 'the slightest reform'
the slightest reform

(2) [$_{NP}$ les [$_{A}$ dix] [$_{A}$ derniers] [$_{N}$ mois]] 'the last ten months'
the ten last months

(3) [$_{NP}$ la [$_{AP}$ [$_{Adv}$ plus] [$_{A}$ grande]] [$_{N}$ discrétion]] 'the highest discretion'
the most big discretion

(4) [$_{NP}$ Cette [$_{N}$ comparison] [$_{AP}$ [$_{A}$ préliminaire]]] semble [$_{AP}$ [$_{A}$ valable]].
this comparison preliminary seems valid
'This preliminary comparison seems valid.

(5) [$_{NP}$ la [$_{Adv}$ plus] [$_{A}$ forte] [$_{PP}$ des [$_{NP}$ [$_{A}$ dix] [$_{A}$ derniers] [$_{N}$ mois]]]]
the most strong of ten last months
'the strongest of the last ten months'

In order to extract adjective valence from the corpus, we focused on AP constituents. This automatically excludes from our consideration simple (prenominal) attributive adjectives and superlative constructions as in (5). It is a desirable result as in neither case a subcategorised element is present: obviously, simple adjectives, as in (1)–(2), do not have dependents, whereas the PP in (5) is not part of the adjective valence but is an inherent part of the construction.

The category adjective, as specified by the annotation schema, comprises several subtypes, such as numerals (e.g., *trois* 'three', *deuxième* 'second'), quantifiers (*plusieurs* 'several') or interrogative adjectival pronouns (*quel* 'which'). These elements do not take complements so they have been excluded from the list of concerned APs. We kept only APs which contain a qualitative adjective: *chaotique* 'chaotic', *adorable* 'adorable', *possible* 'possible', etc.

We counted all qualitative adjectives in the corpus, whether appearing in an AP or not. This resulted in the initial list of 2198 different qualitative adjectives (types), or 17371 occurrences (tokens). Almost a quarter of adjective occurrences (3960 tokens, 23%) were found not in an AP, whereas only 140 adjectives (6% of all types) were never found in an AP in the corpus. This means that almost a quarter of occurrences (tokens) can be used as bare adjectives (prenominal attributes or in non-AP constituents), but very few adjectives (types) are solely used as such. After excluding the 140 adjectives which were never dominated by the category AP in the corpus, our preliminary list of argument taking adjectives contains 2058 candidates (13411 tokens).

2.2 Complements

In French, complements of adjectives can be realised by three syntactic categories: prepositional phrases (PP), subordinate clauses (Ssub) or infinitival verb phrases (VPinf), illustrated in (6). Thus, we limited our search of complement phrases to these three categories.[2] (We filtered out Ssub and VPinf introduced by purpose complementizers, such as *parce que* 'because', *pour que, afin de* 'in order to/so as', *comme* 'as', etc., which indicate adjunct clauses.)

(6) sûr [$_{PP}$ de sa réussite] / [$_{Ssub}$ qu'il réussira] / [$_{VPinf}$ de réussir]
sure of his success that he will succeeded to succeed
'sure [of his success] / [that he will succeed] / [to succeed]'

Not every constituent of this type is a complement of the adjective. As the treebank contains rich annotations but its size is not very big, our quest of real complements of adjectives was mostly guided by linguistic knowledge rather than by the corpus statistics.

Adjectives can appear in various syntactic constructions, e.g., comparative, superlative or impersonal. Our first objective was to identify these constructions

[2]Of course, not all adjectives have the three types of complements.

and separate adjective dependents which are triggered by these environments from real valence arguments.

In comparative constructions, the comparison is introduced by subordinate conjunction *que* 'then', (7). The comparison part is labelled Ssub in the corpus, even if it does not contain a subordinate clause, (8).

(7) La réunion était [$_{AP}$ plus intéressante [$_{Ssub}$ **que** [$_{S}$ je ne pensais]]].
the meeting was more interesting then I NEG thought
'The meeting was more interesting than I thought.

(8) La réunion était [$_{AP}$ plus intéressante [$_{Ssub}$ **que** [$_{NP}$ l'année dernière]]].
the meeting was more interesting then the year previous
'The meeting was more interesting than last year.

In order to distinguish comparative constructions from subcategorised subordinate clauses, we added modifying adverbs to the list of elements which are recognized within an AP: if an adjective is accompanied by a Ssub (i.e., a constituent annotated in the corpus as Ssub) and a comparative adverb (*plus* 'more', *moins* 'less', *aussi, si, tellement* 'as much as, so', *à ce point* 'to this point'), the Ssub is not considered part of the valence frame; both the adverb and the Ssub are removed from the list of dependents.

Keeping track of adverbial dependents also allows us to eliminate certain infinitival phrases introduced by *pour* 'to' (lit.: 'for'). In (9), the VPinf[*pour*] cannot be subcategorised by the adjective *fabuleuse* 'fabulous' since removing the adverb *trop* 'too' renders the sentence ungrammatical. Therefore, VPnf[*pour*] co-occurring with an intensifying adverb (*trop* 'too', *assez* 'enough', *suffisament* sufficiently, etc.) is excluded as a subcategorised element.

(9) Cette histoire est [$_{AP}$ *(trop) fabuleuse [$_{VPinf}$ pour être vraie]].
this story is too fabulous for be true
'This story is too fabulous to be true.'

Predicative adjectives have different properties in impersonal and personal constructions. In (10), Ssub is a true complement, whereas in impersonal constructions, as in (11), Ssub is not part of the adjective valence as it can become the subject; no such transformation is possible in (10). As indicated in (12)–(13), infinitive phrases behave similarly in impersonal constructions.

(10) Paul est heureux [$_{S}$ que Marie vienne].
Paul is happy that Mary comes
'Paul is happy that Mary comes.'

(11) C'est agréable [$_{S}$ que Marie vienne]. [$_{S}$ Que Marie vienne] est agréable.
it is pleasant that Mary comes
'It is pleasant that Mary comes.' 'That Mary comes is pleasant.'

(12) Paul est capable [$_{VPinf}$ de sortir tous les jours].
Paul is capable to get out every the day
'Paul is capable to get out every day.'

(13) C'est agréable [$_{VPinf}$ de sortir]. [$_{VPinf}$ (De) sortir] est agréable.
it is pleasant to get out
'It is pleasant to get out.' 'Getting out is pleasant.'

These two syntactic realisations are directly reflected in corpus annotations, as illustrated in (14)–(15).

(14) [$_{V}$ C'est] [$_{AP}$ agréable] [$_{Ssub}$ que Marie vienne].
it is pleasant that Mary comes
'It is pleasant that the Mary comes.'

(15) [$_{NP}$ Paul] [$_{V}$ est] [$_{AP}$ heureux [$_{Ssub}$ que Marie vienne]].
Paul is happy that Mary comes
'Paul is happy that Mary comes.'

In impersonal constructions, the subject is expressed by the invariable pronouns *ce* 'this' or *il* 'it'. They are specified as clitics attached to the verb rather than as independent NPs. The predicative adjective does not form a constituent with the following subordinate or infinitive clause; both of them are treated as direct dependents of the verb. Therefore, the Ssub (or VPinf) is correctly excluded from the adjective valence. On the other hand, if the subject of the copula is a regular NP, the subordinate (15) (or infinitive (12)) clause is a true dependent of the adjective. There are no functional annotations within AP so the complement vs. adjunct status of the dependent still needs to be established. Again, adverbs help us to specify this distinction.

The subordinate clause is treated as a complement, unless a comparative adverb is present. Note that the syntactic annotations for both (7) and (15) are identical. The Ssub in the former, unlike in the latter, is excluded as a complement due to the presence of the adverbial *plus* 'more'. Similarly, the VPinf clause is considered a complement if no intensifying adverb accompanies the adjective. According to corpus annotations, the two VPinf in (9) and (12) belong to the predicative AP. The former, unlike the latter, is analysed as an adjunct due to the presence of the intensifier adverb *trop* 'too'.

Extracted elements are not indicated in the corpus. Hence, we cannot distinguish between complements of the so-called *tough*-adjectives (the object of the complement VPinf is extracted, cf. (16)) and control adjectives (no element is extracted from the VPinf complement but the sentential subject is coreferential with that of the infinitive, (17)).

(16) Ces erreurs sont [$_{AP}$ faciles [$_{VPinf}$ à comprendre _]].
these mistakes are easy to understand
'These mistakes are easy to understand.'

(17) Jean est [$_{AP}$ long [$_{VPinf}$ à comprendre ces erreurs]].
Jean is long to understand these mistakes
'Jean is slow to understand these mistakes. (It takes Jean long to understand these mistakes).'

For PPs, we retained the lexical preposition which introduces the phrase. In superlative constructions, a PP can be used to specify the "range" of the comparison, (5) or (18). These PPs are not part of the subcategorisation frame of the adjective since they are present only in superlative constructions. This fact is reflected directly in corpus annotations: neither in (5) nor in (18) is PP marked as a dependent of the adjective.

(18) [$_{NP}$ la [$_{AP}$ plus sévère] [$_{N}$ récession] [$_{PP}$ parmi les Douze]]
the most severe recession among the Twelve
'the most severe recession among the Twelve'

For PPs which do appear within an AP, we used PrepLex, Fort and Guillaume (2007), a lexicon which specifies argumental and non-argumental prepositions, i.e., prepositions which may or may not introduce arguments in French. As observed by one of the reviewers, this lexicon has been created for PP arguments of verbs rather than adjectives. In the current experiment, we assume that the prepositions appropriate for the latter are a subclass of the former.

PrepLex contains 49 argumental prepositions, both simple (mono-word) and complex (multi-word). Actually, all of them have a double function and can be used in non-argumental PPs as well. Thus, if a PP is headed by a preposition listed as non-argumental, we can exclude this PP from the valence frame but argumental prepositions cannot reliably identify complements. For adjectives, we added one more non-argumental preposition, *comme* 'as', since it is used only in comparative constructions, i.e., with non-argumental PPs. Other "comparative" prepositions, e.g., *parmi* 'among' or *de* 'of', are retained as they can introduce real complements as well, for example: *nombreux parmi nous* 'many among us' or *satisfait des résultats* 'satisfied with the results'. Among PP dependents of adjectives, we found several complex prepositions (expressions tagged as prepositions in the corpus) which are not listed in PrepLex: *à la tête de* 'leading/at the head of', *à la limite de* 'at the borderline of', *à la suite de* 'as a consequence of', *au profit de* 'at the benefit of', *du fait de* 'from the fact of', *par l'intermédiaire de* 'by means of'. They were all considered to head non-argumental phrases.

3 Results

After adopting the modifications described in the previous section, the initial list of 2058 adjectives (13411 occurrences) was drastically reduced to 271 adjectives (811 occurrences), accompanied by 27 different frame types. It means that the vast majority of qualitative adjectives (almost 90% of all types, around 94% of tokens) do not have complements. More importantly, however, this result shows that only 10% of adjectives which potentially appear with a dependent (i.e., they are part of an AP in the corpus), have a subcategorised for element. 177 adjectives (types) in the final list always occurred within an AP, whereas the remaining 94 adjectives were either found as "bare" elements or appeared within an AP. Indeed, as discussed in sec. 2.1, APs may have no dependents at all, (4), or contain only a modifying adverb, (3). On the other hand, many of AP components do not belong to the adjective valence but are related to specific syntactic constructions, sec. 2.2.

frame	freq.	# adj.	frames with frequency 1
P-OBJ:PP[à]	269	74	P-OBJ:PP[devant]
P-OBJ:PP[de]	194	90	P-OBJ:PP[face à]
P-OBJ:VPinf[de]	90	26	P-OBJ:PP[sous]
P-OBJ:VPinf[à]	54	16	P-OBJ:PP[à]|P-OBJ:VPinf[de]
P-OBJ:PP[pour]	36	29	P-OBJ:PP[jusqu'à]
OBJ:Ssub[que]	31	7	OBJ:VPinf
P-OBJ:PP[en]	31	24	P-OBJ:PP[selon]
P-OBJ:VPinf[pour]	24	6	P-OBJ:PP[à]|P-OBJ:VPinf[sans]
P-OBJ:PP[dans]	23	15	OBJ:Ssub[que]|P-OBJ:PP[de]
P-OBJ:PP[par]	13	12	P-OBJ:PP[envers]
P-OBJ:PP[sur]	11	11	P-OBJ:PP[vis-à-vis de]
P-OBJ:PP[avec]	9	6	P-OBJ:PP[parmi]
P-OBJ:PP[entre]	5	3	
P-OBJ:PP[chez]	4	3	
P-OBJ:PP[depuis]	3	3	
P-OBJ:PP[après]	2	2	

FIGURE 1: The extracted frames and their frequency counts

The average number of frames per argument-taking adjective is about 1.25, which is a much lower ambiguity rate than the result we obtained for French verbs (around 2), cf. Kupść and Abeillé (2008). This is a natural consequence of the fact that the variety of adjective frames (27) is much more reduced in comparison to verb frame types (180). Also, the proportion of adjectives which appear with a single frame (82.3%, 223 adjectives) highly exceeds single-frame verbs (58.2% verbs). The frequency counts for the extracted adjectival frames are given in Fig. 1.

Fig. 1 indicates that the most frequent complements are PPs introduced by *à* or *de*. The same forms are also used with infinitival complements. Prepositions *pour* and *en* are also frequent, both in the text and in adjective frames. Only seven adjectives appear with the subordinate clause complement (*autre* 'other', *certain* 'sure', *conscient* 'conscious', *tel* 'such', *ciblé* 'targeted' and two irregular comparative forms *meilleur* 'better' and *pire* 'worse') but its use is quite frequent. Similarly, the infinitival clause with *pour* 'for' is more frequently used than it is found as a complement in an adjective frame. Among the six adjectives with VPinf[pour] frame, the first four (*indispensable* 'indispensable', *insuffisant*'insufficient', *suffisant* 'sufficient', *nécessaire* 'necessary') are correctly recognized, whereas for the last two (*énorme* 'enormous' and *étroit* 'narrow'), the infinitive complement has been confused with the *pour*-clause of the intensifier construction, (9).

As for frames which appeared only once, 3 entries were extracted with two arguments: *créateur* 'creator' (P-OBJ:PP[à]|P-OBJ:VPinf[sans]), *averti* 'informed' (OBJ:Ssub[que]|P-OBJ:PP[de]), and *heureux* 'happy' (P-OBJ:PP[à]|P-OBJ:VPinf [de]). All three frames are, however, incorrect. The first one should be completely ignored as *créateur* is a noun mistakenly tagged as an adjective. For the other

two, according to the corresponding source sentences, only one argument should be present: in *averti*, Ssub comes from a comparative construction and, thus, is not subcategorised, whereas in *heureux*, PP[à] is a locative adjunct (*heureux à Rome* 'happy in Rome'), and it shouldn't be present either. One adjective (*égoïste* 'egoistic') has a peculiar OBJ:VPinf frame, i.e., it takes an infinitival complement without any preposition. This frame results from an annotation problem where the infinitive is indicated as a complement of the adjective rather than of a verb.

As illustrated in (6), an adjective can appear with several different types of a complement. For the majority of adjectives, a single frame has been extracted (82.3%). The distribution of multi-frame adjectives is as follows: 1 adjective (*difficile* 'difficult') has 5 frames, 2 adjectives (*nécessaire* 'necessary' and *présent* 'present') appear with 4 frames, 13 — with 3 frames and 32 — with 2 frames. For most adjectives, several frames result from a different form of the preposition (or the complementizer) used with the same type of frame. For example, *présent* 'present' is always followed by a PP but 4 different prepositions are possible: *dans* 'in', *à* 'at', *en* 'in' and *sur* 'on'. Following traditional linguistic tests, if an element does not combine with a specific form of a preposition but various forms are possible, the phrase introduced by the preposition is likely an adjunct. This conclusion should be taken with caution as locative PPs can be complements (of certain verbs) despite the fact that the preposition is realised by multiple lexical forms. Therefore, if different prepositions are used, their semantic uniformity should be verified as well.

4 Conclusion and Further Questions

We presented initial results of our ongoing project on extracting subcategorisation frames for French adjectives. We obtained a list of 271 adjectives and 27 valence frames. The extracted frames as well as the list of argument taking adjectives should be considered preliminary. We focused mostly on separating productive constructions from valence realisations using corpus annotations and linguistic knowledge. This approach targets mainly subordinate and infinitive clauses since mostly these two frame types participate in adjective constructions. We used PrepLex, a lexicon of prepositions, in order to distinguish argumental and non-argumental PPs in APs. Argumental prepositions cannot reliably indicate complements as they also appear with non-argumental PPs. Hence, our extraction method should be completed by other techniques, e.g., linguistic tests or statistical approaches.

Although preliminary, the presented experiments allow us to ask questions which have not been frequently addressed so far: how to distinguish complements from adjuncts in APs (constructions vs. subcategorisation)? which traditional linguistic tests (applied most often to verbal dependents) can be helpful to determine adjective valence? do verbs and adjectives have the same argumental prepositions? how derivational morphology (for deverbal adjectives) affects valence?

Seeking answers to these questions will involve exploring different linguistic domains (syntax, morphology and possibly semantics) but also confronting techniques well-established in general linguistics.

The lexicon will be freely available from the site:
`http://erssab.u-bordeaux3.fr/article.php3?id_article=150`

References

Anne Abeillé, Lionel Clément, and François Toussenel (2003), Building a treebank for French, in *Treebanks*, Kluwer.

Didier Bourigault and Cécile Frérot (2005), Acquisition et évaluation sur corpus de propriétés de sous-catégorisation syntaxique, in *Actes des 12èmes journées sur le Traitement Automatique des Langues Naturelles*.

T. Briscoe and J. Carroll (1993), Generalised probabilistic LR parsing for unification-based grammars, *Computational linguistics*.

John Carroll and A. Fang (2004), The automatique acquisition of verb subcategorisations and their impact on the performance of an HPSG parser, in *Proceedings of the 1st International Conference on Natural Language Processing*, Sanya City, China.

Paula Chesley and Susanne Salmon-Alt (2005), Le filtrage probabiliste dans l'extraction automatique de cadres de sous-catǵorisation, in *Journé ATALA sur l'interface lexique-grammaire*, Paris.

Laurence Danlos (1985), *La génération automatique de textes*, Masson.

Karën Fort and Bruno Guillaume (2007), PrepLex: a lexicon of French prepositions for parsing, in *ACL SIGSEM07*.

Claire Gardent, Bruno Guillaume, Guy Perrier, and Ingrid Falk (2006), Extraction d'information de sous-catégorisation à partir du lexique-grammaire de Maurice Gross, in *TALN 2006*.

Maurice Gross (1975), *Méthodes en syntaxe*, Hermann.

Maurice Gross (1986), Lexicon–grammar: the representation of compound words, in *Proceedings of the 11th coference on Computational linguistics*, pp. 1–6, Association for Computational Linguistics, Morristown, NJ, USA.

Alain Guillet and Christian Leclère (1992), *La structure des phrases simples en français*, Droz, Genève.

Chung-hye Han, Juntae Yoon, Nari Kim, and Martha Palmer (2000), A Feature-Based Lexicalized Tree Adjoining Grammar for Korean, Technical report, IRCS.

Anna Kupść (2007), Extraction automatique de cadres de sous-catégorisation verbale pour le français à partir d'un corpus arboré, in *Proceedings of TALN'07*, pp. 153–162.

Anna Kupść and Anne Abeillé (2008), Growing TreeLex, in Alexander Gelbukh, editor, *Computational Linguistics and Intelligent Text Processing*, number 4919 in Lecture Notes in Computer Science, pp. 28–39, Springer.

Igor Mel'cuk, Nadia Arbatchewsky-Jumarie, and André Clas (1984, 1988, 1992, 1999), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques, vol. I, II, III, IV*, Les Presses de l'Université de Montréal.

Bruno Menon and Nicole Modiano (1993), EAGLES: Lexicon Architecture, Technical Report EAG-CLWG-LEXARCH/B, EAGLES.

Nilda Ruimy, Ornella Corazzari, Gola Elisabetta, Antonietta Spanu, Nicoletta Calzolari, and Antonio Zampolli (1998), The European LE-PAROLE Project and the Italian Lexical Instantiation, in *Proceedings of ALLC/ACH, 1998, Lajos Kossuth University, Debrecen, Hungary, July, 5-10 1998*, pp. 149–153.

Benoît Sagot, Lionel Clément, Éric Villemonte de La Clergerie, and Pierre Boullier (2006), The Lefff 2 syntactic lexicon for French: architecture, acquisition, use, in *Actes de LREC 06, Gènes, Italie*.

M. Surdeanu, S. Harabagiu, J. Williams, and P. Aarseth (2003), Using predicate-argument structures for information extraction, URL `http://citeseer.ist.psu.edu/surdeanu03using.html`.

Karel van den Eynde and Piet Mertens (2003), La valence: l'approche pronominale et son application au lexique verbal, *French Language Studies*, 13:63–104.