

The current status of FTAG

Anne Abeillé, IUF, TALaNa & LaTTICe, Paris 7
 Marie-Hélène Candito, TALaNa & LaTTICe, Paris 7 and Lexiquist
 Alexandra Kinyon, TALaNa & LaTTICe, Paris 7 and IRCS, UPenn
 abeille, kinyon, candito@linguist.jussieu.fr

Introduction

We describe the current status and organization of a French Lexicalized Tree Adjoining Grammar (FTAG), developed over the last 10 years at TALaNa (Abeillé 91, Candito 99). The new version grammar is generated semi-automatically, independently of any corpus or application domain. It is intended to model speaker competence, and can be used both for parsing and generation. As far as parsing is concerned, we describe a general processing module which can rank the different parses produced based on linguistic information present in FTAG.

1. General linguistic choices

Most of our linguistic analyses follow those of Abeillé 91 (except that clitic arguments are substituted and not adjoined), complemented by Candito 99. We dispense with most empty categories, especially in the case of extraction.¹ Semantically void (or non autonomous) elements, such as complementizers, argument marking prepositions or idiom chunks, are coanchors in the elementary tree of their governing predicates.

1.1 A minimal tagset

We depart from traditional part of speech wherever the modern linguistic analyses have better to propose, especially in the generative tradition. We thus distinguish a special category for Clitics (weak pronouns) following Kayne 75, and for Complementizers. We collapse proper names, common nouns and pronouns into one category N, with features. We do not have a tag for subordinating conjunctions which are either Prepositions (followed by a complementizer: *pendant que* (during)) or (full) Complementizers (*si* (if), *comme* (as)...). Sentential structures are 'flat' (no internal VP). We thus have the following tagset.

Lexical categories: D (determiners), N (nouns, names, pronouns), V (verb), Cl (clitic pronoun), Prep (preposition), A (adjective), Adv (adverb), Conj (Coordinating conjunction), C (complementizer, subordinating conjunction),

Non lexical categories: SP (prepositional phrase), S (sentence). A and N are also used for nominal or adjectival phrases.

1.2 A rich set of grammatical functions

Tree sketches of the French TAG are compiled out of the French metagrammar (Candito 99), which expresses subcategorization in terms of grammatical functions. The functions used in the French MG for verbs are the following:

subject, object, dat-object, obl-object, gen-objet, locative, source-locative, manner, goal-infinitive, perception-infinitive, interrogative clause, "predicative complement"

All these functions can be both initial functions and final functions. An additional function "agt-object" is used as final function only, and is beared by a by-phrase in the case of passive.

We use several "complement" functions for complements of adjectives, prepositions, nouns, adverbs. And these categories may bear the function "modifier" with respect to the element they modify.

1.3. A parsimonious use of features

Most of the syntactic properties handled by feature structures in unification based linguistic theories (LFG or HPSG) are directly captured by the topology of the elementary trees in LTAG.

¹ We keep some empty categories for non realized arguments, such as PRO subjects (see Abeillé 91).

No use has to be made of valence or slash features to ensure subcategorization requirements or filler-gap relations. No feature passing principles, besides unification, are needed either.

We only rely on atomic valued features (which guarantees against any cyclic structure). We distinguish between:

- Morphological features, which are used in the morphological lexicon, in the syntactic lexicon when an argument is constrained for them (eg *trouver* has only indicative sentential complement) and for agreement in the elementary tree sketches,
- Syntactic features, used in the syntactic lexicon (for a verb to disallow passive for example) and in the tree sketches (to distinguish between trees in the same family or to further constrain tree combinations),
- Semantic features : these are gross classifications used for arguments (human, locative etc) which should be further refined.

We are currently using about 40 features as follows:

morphological features: <det>, <card>, <case>, <el>, <mode>, <num>, <ord>, <pers>, <P-num>, <P-pers>, <tense>.

syntactic features: <ant>, <ant-s>, <ant-v>, <aux>, <cq>, <det>, <extrap>, <gen>, <inv>, <modif>, <neg>, <nom>, <passive>, <part-num>, <part-gen>, <pred>, <princ>, <pro>, <quant>, <san1>, <san2>, <subj-gen>, <subj-pers>, <subj-num>, <sym>, <tense>, <wh>.

semantic features: <conc>, <degre>, <hum>, <loc>, <man>.

2. The Internal organization of FTAG

2.1. 3 sources of information for lexicalized elementary trees

Strict lexicalization at execution time does not prevent from representing the elementary trees in a less redundant way. Indeed it is required for any reasonably sized grammar, since for instance a verbal form may anchor dozens or hundreds of elementary trees. A first level of sharing between elementary trees was proposed within the XTAG system (XTAG group 1995) : elementary trees are compiled out of three sources of information:

- a set of tree sketches ("pre-lexicalized" structures, whose anchor is not instantiated)
- a syntactic lexicon, where each lexeme is associated with the relevant tree sketches
- a morphological lexicon, where inflected forms point to a lemma plus morphological features

Lexical selection of tree sketches is controlled by features from the syntactic and morphological lexicons, and uses the notion of tree families : sets of tree sketches that share the same initial argumental structure. The tree sketches of a family show all the possible surface realization of the arguments (pronominal clitic realization, extraction, inversion...) and all the possible transitivity alternations (impersonal, passive, middle..).

A lexeme selects one or several families (corresponding to one or several initial subcat frames) and with the help of features selects exactly the relevant tree sketches : The features may rule out some tree sketches of the selected family, either because of morphological clash (eg. the passive trees are only selected by past participles) or because of "idiosyncrasies" (eg. the French transitive verb *peser* -to weigh- disallows passive).

Figure 1 shows an elementary tree anchored by *parlait* (talked) and the corresponding tree sketch.

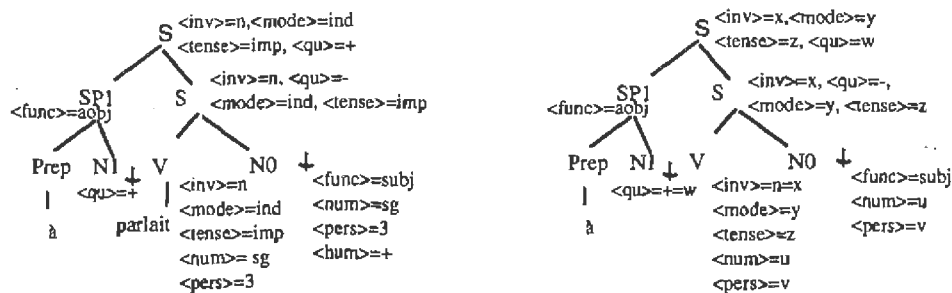


Figure 1. Lexicalized tree and tree sketch

The lexicalized tree is compiled out of the tree sketch and the following lexical entries with feature unification, (cf Barrier et al 00):

Morphological database:

parlait, V: [parler], {V.b:<mode>=ind, <tense>=imp, <num>=sg, <pers>=3}.

Syntactic database:

[parler], V: n0Vn1 {NO:<hum>=+}.

The inflected form *parlait* points to the lemma PARLER, and the lexeme /PARLER/, that comprise the single lemma PARLER, selects in turn the n0Vn1 family, where the preposition appears as a co-anchor (except in the case the argument 1 is cliticised).

2.2 The lexicons

Contrary to the English LTAG which reuses existing dictionaries (Collins 1979 for the morphological database, Oxford English Dictionary and COMLEX for the syntactic database), our French lexicons had to be done by us. They currently comprise the following items:

Morphological lexicons: over 50 000 (inflected) forms: 45800 for verbs, 3500 for nouns and pronouns, 950 for adjectives and 50 for determiners.

Syntactic lexicons: over 6000 (disambiguated) entries: 3700 for verbs, 500 for prepositions and adverbs, 800 for adjectives, 80 for determiners, 2000 for nouns, 350 for idioms

The lexical items chosen have been extracted as the most frequent ones from the frequency lists of Julliard 1970 and Catach 1984, except for idioms where one had to rely on personal intuitions. They have been disambiguated (and separated into different syntactic entries) with standard dictionaries as well as LADL lexicon-grammar tables (Gross 1975). The morphological lexicons have been automatically generated, using PC-Kimmo adapted to French. Both lexicons are organised in lexical databases, and the features normalized with templates.²

The morphological lexicon has nothing specific and associates lemmas, inflected forms and relevant morphological features. The syntactic lexicon associates lemmas with constructions (elementary trees or tree families with features) and performs some meaning disambiguation (based on different syntactic constructions, for example for the French verb *abattre* - knock down, shoot down) :

INDEX: abattre/1 (physical meaning)

ENTRY: abattre

POS: V

FAM: n0Vn1

FS:

INDEX: abattre/2 (psychological meaning, possible sentential subject)

ENTRY: abattre

POS: V

FAM: s0Vn1

FS: #N1_HUM+, #NO_HUM-

Future developments include integrating a more complete full form lexicon (over 400 000 forms independently developed for our tagger; cf. Abeillé et al 1998) into the morphological database, and developing the syntactic lexicon (with shallow parsed corpora and reuse of LADL valence tables for French verbs, cf. Namer and Hathout 1998).

2.3 The metagrammar

We use an additional layer of linguistic description, called the metagrammar (MG) (Candito 1996, 99) which imposes a general organization and formalizes the well-formedness conditions for elementary tree sketches. It provides a general overview of the grammar and makes it possible for a tool to automatically generate the desired tree sketches from the combination of smaller descriptions.

MG thus represents a TAG as a multiple inheritance network, whose classes specify syntactic structures as partial descriptions of trees (Vijay-Shanker & Schabes 92, Rogers & Vijay-

² For unknown words, a default tree assignment is used.

Shanker, 94). Partial descriptions of trees are sets of constraints that may leave underspecified the relation existing between two nodes. The relation between two nodes may be further specified by adding constraints in sub-classes of the inheritance network. Inheritance of partial descriptions is monotonic.³

In order to build pre-lexicalized structures respecting the Predicate Argument Cooccurrence Principle, and to group together structures belonging to the same tree family, MG makes use of syntactic functions to express either monolingual or cross-linguistic generalizations (as in LFG or Relational Grammar). Subcategorization of predicates is expressed as a list of syntactic functions, and their possible categories. The initial subcategorization is that of the unmarked case, and is modifiable by redistribution (or transitivity alternations).

Structures sharing the same initial subcategorization are grouped in a tree family. For verbal predicates, an elementary tree is partly represented with an ordered list of successive subcategorizations, from the initial one to the final one. Elementary trees sharing a final subcategorization, may differ in the surface realizations of the functions. MG represents this repartition of information by imposing a three-dimension inheritance network :

- Dimension 1: initial subcategorization
- Dimension 2: redistributions of functions
- Dimension 3: surface realizations of syntactic functions.

Dimension 1 describes a possible initial subcategorization (and possibly frozen elements). Dimension 2 describes a list of ordered redistributions (including the case of no-redistribution) which may impose a verbal morphology (eg. the auxiliary for passive). Dimension 3 represents the surface realization of a function (independently of the initial function).

The 3 dimension hierarchy is handwritten, the elementary trees are automatically generated with a two-step process. First the compiler automatically creates additional classes of the inheritance network : the "crossing classes". Then each crossing class is translated into one or several tree sketches (the minimal structures satisfying all inherited constraints). During the first step, crossing classes are automatically built as follows (with unification):

- a crossing class inherits one terminal class of dimension 1
- then, the crossing class inherits one terminal class of dimension 2
- then, the crossing class inherits classes of dimension 3, representing the realizations of every function of the final subcategorization.

The tree sketch of figure 1, for example, has been compiled, out of an initial subcategorization with nominal subject and dative object (dimension 1), an active canonical redistribution (dimension 2), a nominal inverted realization for the subject, and a fronted interrogative realization for the dative object (dimension 3).

3. Elementary trees in FTAG

3.1 Linguistic principles for elementary trees

Within FTAG, elementary trees respect the following linguistic well-formedness principles: (Kroch & Joshi 85, Abeillé 91, Franck 92, Candito 99, Candito & Kahane 98)

- Strict Lexicalization : all elementary trees are anchored by at least one lexical element, the empty string cannot anchor a tree by itself,
- Semantic Consistency : no elementary tree is semantically void (this ensures the compositionality of the syntactic analysis),
- Semantic Minimality : no elementary tree correspond to more than one semantic unit (modulo lexicalism : lexical anchors are not broken down into morphemes).
- Predicate Argument Cooccurrence Principle (PACP): an elementary tree is the minimal syntactic structure that includes a leaf node for each realized semantic argument of the anchor(s).

Initial trees are used for arguments, verbs with non sentential arguments, auxiliary trees are used for modifiers, determiners, modals, auxiliaries and verbs with sentential complement,

Some examples of elementary trees are the following:

³. In MG, nodes of partial descriptions are augmented with specific feature structures, called meta-features, constraining for instance, the possible parts of speech of a node or the index in the case of argumental nodes.

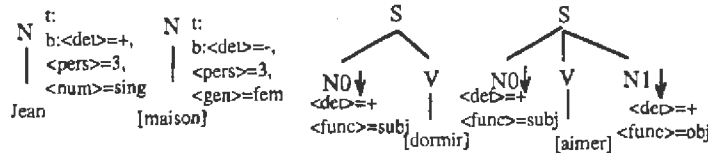


Figure 2. Initial elementary trees

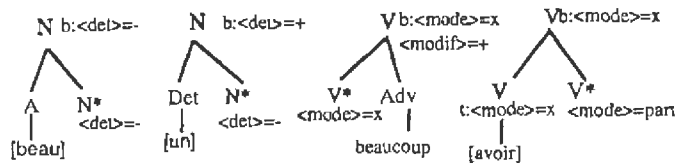


Figure 3. Auxiliary elementary trees

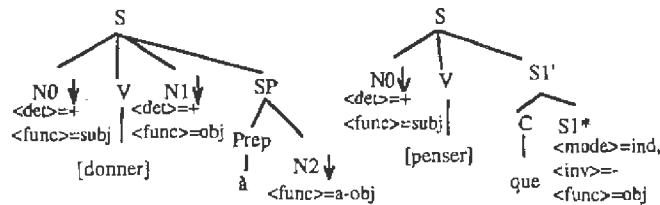


Figure 4. Elementary trees with functional co-anchors

In Figure 3, the relevant syntactic and semantic units are donner-à (give to) or penser-que (think that).

3.2. The metagrammar for FTAG

The set of tree sketches in FTAG is comprised over 5000 elementary tree sketches (not counting trees for causative constructions). Currently, all but 40 of them are compiled from the French metagrammar. The 40 remaining tree sketches are trees for determiners (plain and complex), nouns used as arguments, coordination conjunctions, clitics and "special" trees for deficient verbs such as raising verbs and auxiliaries.

The French MG comprises the description for the tree sketches anchored by full verbs, prepositions, adverbs, adjectives, and nouns (when used as modifiers).

Within dimension 1, it comprises 54 initial subcategorization frames for verbs (which means there are 54 tree families for verbs in FTAG), 4 initial subcategorizations for adjectives, 12 for adverbs, prepositions, subordinating conjunctions and nominal modifiers.

In dimension 2, primarily relevant for verbs, we have defined as redistributions the following phenomena:

- Passive (with or without agent) : additional V-headed elementary trees with auxiliary être substituted
- Causative constructions
- Reflexive
- Impersonal constructions (active and passive)
- Middle se (ces robes se lavent facilement)

In dimension 3, we define as realizations the following phenomena:

- Infinitival and sentential arguments (treated as S-complements)
- Relatives (qui, que, dont, Prep qui, Prep lequel), indicative, subjunctive
- Interrogatives (direct, indirect, est-ce-que)
- Cleft sentences (c'est que, c'est qui, c'est Prep N que)
- Clitic pronouns
- Subject inversion (nominal or subject clitic)
- Unbounded dependencies (with island constraints)

- Participials (past or present part, NP modifiers), Past participle agreement.
 - Null realization (empty subjects for infinitives, participials or imperatives)
 - Factorization (for subjects in coordinated phrases)
 - Word order variation (among complements)
- Work remain to be done of the syntax of quantifiers (often discontinuous in French), on negation (including negative concord), coordination and comparison (including superlatives).

4. Evaluating FTAG

Evaluating a wide coverage grammar is a difficult task, especially in the absence of reference tree banks for French. We performed a quantitative evaluation using the French test suite developed in the TSNLP project (Estival & Lehmann 96). Further evaluation will be done on newspaper corpora.

4.1. Evaluation using TSNLP

We have performed an external evaluation using the TSNLP multilingual data base, which aims at covering the major syntactic phenomena for each language, using a minimal vocabulary (a few hundred words). We have extracted all the French items of the TSNLP data base, classified by grammatical status (we only took 0 and 1), by length and by phenomenon (according to TSNLP original classification). For all grammatical items, the results with the 1998 version of our grammar are as follows:⁴

- over 80 % of the grammatical parsed, with an average of 2.9 parses per sentences
- over 82% of the agrammatical sentences have been correctly rejected.

There were no unknown words. The main failure cases are the following:

- missing lexical coding (transitive verb without object, transitive use of intransitives),
- missing elementary tree (causative trees, postverbal clitics with imperatives),
- feature unification clash (agreement with politeness forms: *vous êtes belle*, or with coordination : *deux bandes bleue et jaune*),
- missing phenomenon (tough construction, gapping...).

Cases of overanalysis either come from a disputable TSNLP coding (for example for sequence of times), or from the incompleteness of our representation (for example for coordination or negation, we overgenerate).

4.2. Comparison with other syntactic resources

The lexicon-grammar developed at LADL for more than 20 years is an unrivaled source of knowledge reusable in the sens that it is not designed for any program and not even dependent on any special formalism. However, it cannot be directly used to analyse (or generate) a text since it only lists some basic constructions (with their lexical head). It does not code the crossing of constructions nor the productive phenomena which are not clearly lexically sensitive (such as causative, quantifier floating or argument extraction for simple verbs). Thus, even though it is crucial to know that transitive *voler* (to steal) must be distinguished from intransitive *voler* (to fly), more general grammatical rules are needed to know that it is the transitive *voler* which is instantiated in examples (1)-(2) without a postverbal NP object, or that it is the intransitive *voler* which is instantiated in examples (3)-(4) (even though there is a postverbal NP):

- (1) Ils veulent tout voler
- (2) les bijoux qu'ils ont finalement avoué avoir volé...
- (3) Lufthansa fait voler ses avions 5 jours sur 7
- (4) A une altitude à laquelle ne vole normalement aucun avion ...

M. Salkoff (1973, 79) string grammar has listed numerous grammatical strings representative of French syntax but has never been associated with a sizable lexicon and cannot be reused independently of the parsing scheme it was made for. The HPSG like grammar developed for French by Namer and Schmidt 93 suffers from the same problems and is totally dependent upon the ALEP development platform.

⁴ A previous evaluation done in 1996 (Abeillé et al. 96), with a smaller coverage grammar (comprising about 830 elementary tree sketches), using the same lexicon, had the following results : 65% of the grammatical sentences (excluding coordination) parsed, with an ambiguity rate of 1,5 parses per parsed sentence.

The GB grammar developed for French at LATL (Wehrlé 97), is more modular and associated with a sizable dictionary. But it is not clearly separated from the program that uses it (extraction or passive phenomena are not handled as grammatical data but as types of action - attachment, trace creation... - that the program does at a certain stage in the parsing scheme) and thus cannot be reused as such for other applications.

5. Ranking parses

To be usable in practice, our grammar must associate one best analysis per grammatical sentence. The output of a TAG parser can be viewed as a derived tree (encoding phrase structure) or as a derivation tree (encoding dependencies). Since it is both more compact and more informative, we choose the derivation tree for parse ranking (contrary to Srinivas & al. 95).

5.1. General Disambiguation principles

Our parse ranker is based on empirical (i.e. corpus-based) and psycholinguistic-based preferences (Kinyon 99). It only uses lexical and syntactic sources of information (whereas a true disambiguator should also use semantic and discourse information). Since we work on the derivation tree which exhibits the lexicalized trees used for parsing, it is easy to mix lexical and syntactic preferences. Our parse ranker thus uses 3 types of preferences : lexical preferences (such as valence preference for verbs), grammatical preferences (construction types) and general principles which are structure-based, domain, language and application independent.

The lexical preferences code either a category preference or a valence principle. They have to be computed for each word, but we rely on the general tendency in French to favor grammatical categories over lexical categories for ambiguous forms (for example weak pronouns (clitics) to strong pronouns, or auxiliaries over full valence verbs).

The grammatical preferences code a construction preference, for example active over passive or personal over impersonal. In *Il est venu une nuit*, the personal interpretation (with *il* as personal subject and *une nuit* as adjunct) is to be favored over the impersonal one (with *une nuit* as deep subject).

The general principles assume the existence of a universal preference for economy (e.g. adjunction is more costly than substitution) and therefore favor analysis that needs to perform the fewer operations. Formulating structural preference principles in terms of derivation tree allows to capture widely accepted preferences, which turn out to be difficult to formalize in terms of constituent trees : idioms are preferred over literal interpretations, arguments are preferred over modifiers.

These general principles are the following :

- 1- Prefer the derivation tree with the fewer number of elementary trees (=fewer nodes)
- 2- Prefer to attach initial trees low
- 3- Prefer the derivation tree with the fewer number of auxiliary trees

Principle 1 favors the idiomatic interpretation of a sentence over its literal interpretation (a), since the different idiom chunks belong to the same elementary tree. It also favors the attachment prepositional phrases as arguments rather than modifiers (b). Principle 2 favors the low attachment of arguments, when several alternative attachments are possible : in (c) the PP *de la manifestation* is an argument of the N *organisateur* rather than of the V *soupçonne*. In (d), the PP *à Jean* is an argument of *dit* rather than of *parle*. Principle 3 favors the derivation tree involving the fewer number of adjunctions (i.e. modifiers) : in (e) *le matin* could be a modifier, but the attachment as an argument is preferred.

- (a) Jean brise la glace (Jean cuts the mustard > Jean breaks the ice)
- (b) Jean pense à la reunion (Jean thinks of the reunion > Jean thinks at the reunion)
- (c) Jean remercie l'organisateur de la manifestation (J. suspects the organizer of the demonstration > for the demonstration)
- (d) C'est à Jean que Marie dit que Paul parle (It's to Jean that Marie says that John thinks > It's of Jean that)
- (e) Jean attend le matin (Jean awaits the morning > J. waits in the morning)

In case of conflict, the priority is for lexical preferences, then grammatical preferences, then general principles.

5.2. Application to TSNLP

The parsed item from TSNLP had an average of 2.9 parses per item. No categorial ambiguity remained. Most feature ambiguities are handled via underspecification (eg "les enfants" feminine ou masculine). The remaining (structural) ambiguities are the following (not all of these are spurious) :

- modifier adjoined to S ou V after an intransitive verb (L'ingénieur viendra volontiers),
- prepositional phrase analysed as complement or modifier (L'ingénieur préfère le vin à l'eau; Il passe pour un spécialiste),
- passive with or without agent (the par-PP can be analysed as an agent phrase or as a modifier)
- several adjunction sites in case of multiple modifiers.

After applying the general preference principles, we are left with only 2.17 derivations / sentence (i.e. -24 %), while the number of sentences for which a "correct" parse is present only marginally decreased. After applying the language specific preferences, we are left with 1,5 derivation / sentence (i.e. - 47 % in total). It turns out that one of the main sources of spurious ambiguities lies in adverbial attachment. We are exploring how to add lexical preferences to deal with this case.

References

- A. Abeillé, 1991. *Une grammaire lexicalisée d'arbres adjoints pour le français*. PhD Thesis, University Paris 7. (to appear Éditions du CNRS)
- A. Abeillé, B. Daille, B. Robichaud, 1996, FTAG : un analyseur de phrases françaises, Proceedings ILN, Nantes.
- N. Barrier, S. Barrier, A. Kinyon, 2000, Lexik: a tool for FTAG, Proceedings TAG+5, Paris. this volume
- M-H. Candito, 1996. A principle-based hierarchical representation of LTAG, *Proceedings 15th COLING*, Copenhagen.
- M-H. Candito, 1999. Représentation modulaire et paramétrable de grammaires électroniques lexicalisées. Application au français et à l'italien. PhD dissertation. University Paris 7.
- M-H. Candito, S. Kahane, 1998. Can the TAG derivation tree represent a semantic graph ? an answer in the light of MTT, *Proceedings TAG+4 Workshop*, Philadelphia.
- N. Catach, 1984. *Les listes orthographiques de base du français*, Paris, Nathan.
- L. Danlos, 1998. GTAG : un formalisme lexicalisé pour la génération inspiré de TAG, *TAL* 39:2.
- C. Doran, D. Egedi, Hockey, B., Srinivas, B., Zaidel M. 1994. The XTAG system : a wide coverage grammar for English, *Proceedings 15th COLING*, Kyoto, pp. 922-928.
- C. Doran et al 2000, Evolution of the XTAG system, in Abeillé, Rambow (eds) *Tree Adjoining grammars*, CSLI.
- D. Estival, S. Lehmann, 1997. TSNLP: des jeux de phrases test pour le TALN, *TAL*, 38:1, pp. 155-172.
- R. Frank 1992. Syntactic locality and Tree Adjoining grammar: grammatical, acquisition and processing perspectives, PhD Dissertation, Univ. Pennsylvania, Philadelphia.
- M. Gross, 1975. *Méthodes en syntaxe*, Hermann, Paris.
- A. Julliard, D. Brodin, C. Davidovitch, 1970. *Frequency dictionary of French words: 5000 words*, Mouton, La Haye.
- R. Kayne, 1975. *French syntax : the transformational cycle*, Cambridge, MIT Press.
- A. Kinyon 1999. Hiérarchisation d'analyses basée sur des informations dépendancielles pour les LTAGs. Proceedings TALN'99. Cargèse.
- A. Kroch, A. Joshi, 1985. The linguistic relevance of Tree Adjoining grammars, Technical Report, University of Pennsylvania, Philadelphia.
- F. Namer, P. Schmidt, 1993. Une grammaire du français dans un formalisme à structures de traits typés, *Proceedings ILN*, Nantes.
- F. Namer, N. Hathout 1998. Automatic construction and validation of French large lexical resources: reuse of verb theoretical descriptions, *Proceedings 1st LREC Conference*, Granada.
- J. Rogers, K. Vijay-Shanker, 1994. Obtaining trees from their descriptions, an application to tree adjoining grammar, *Computational Intelligence*, 10:4.
- M. Salkoff, 1973. *Une grammaire en chaîne du français*, analyse distributionnelle, Paris: Dunod.
- M. Salkoff, 1979. *Analyse syntaxique du français, grammaire en chaîne* J. Benjamin, Amsterdam.
- M. Silberstein, 1993. *Dictionnaires électroniques pour l'analyse lexicale : le système INTEX*, Masson, Paris.
- B. Srinivas, Doran C., Kulick S. 1995 : Heuristics and Parse Ranking. 4th workshop on Parsing Technologies. Prag. Czech Republic.
- B. Srinivas, 1997. Complexity of lexical descriptions and its relevance for partial parsing, PhD thesis, University of Pennsylvania, Philadelphia.
- K. Vijay-Shanker, Y. Schabes, 1992. Structure sharing in LTAG, Proceedings 14th COLING, Nantes.
- E. Wehrli, 1997. *L'analyse syntaxique des langues naturelles : techniques et méthodes*, Paris: Masson.