

Are the temporal structures of texts algorithms?

Nikolay Vazov and Guy Lapalme

Département d'Informatique et Recherche Opérationnelle
Université de Montréal, C.P. 6128, succursale Centre-ville
Montréal (Québec), Canada, H3C 3J7
{vazovn,lapalme}@iro.umontreal.ca

Abstract

This article proposes an algorithm for identification of temporal structures of French texts. The algorithm is based exclusively on the linguistic information found in the texts and uses two alternating operations which apply to the utterances in the text in their linear order. The first operation assigns an aspecto-temporal value to each next utterance. The second one controls the association of this value to the set of already identified values. The temporal structure of a text is seen as a dynamically built configuration of interacting semantic values which are organised according to the temporal relevance of each situation in the text unit.

Introduction

This article discusses an algorithm for automatic identification of temporal structures in French texts. This algorithm is an extension of our previous research on temporal structures whose goal was to propose a solution to two major problems in the analysis of temporality in French:

- how to identify the temporal structures of texts (paragraphs)?
- is the number of possible temporal structures finite?

Our first claim was that the aspecto-temporal information in a text is organised in a temporal structure. The constituents of the temporal structures are the aspecto-temporal values of all main clauses in the text. The type of temporal structures is thereby determined by the type and the relative positions of the values in the text.

Our second claim was that the number of temporal structures is finite. This article is an attempt to provide evidence to this hypothesis and to represent the temporal structures as dynamic semantic configurations determined by their own "grammar". Temporal structures can be of arbitrary length (as soon as texts can be of arbitrary length too) but the process of their construction is controlled by a finite set of rules. Both the speaker and the hearer apply this set to encode and,

respectively, to decode the temporal information in the text.

In all, our approach distinguishes 6 types of temporal structures: 3 primary and 3 coerced. The three primary temporal structures are *descriptions*, *narrations* and *comments*. The values in these structures are represented exclusively as "left-associative" operators (Steedman 1997). Hence, their order in the structure will correspond to the linear order of the utterances.

Very often, some values are semantically incompatible with the already built temporal structures and fail to apply the "association to the left". As a consequence, they operate on the structures resulting from the previous clauses and shift, or coerce, their type (Moens & Steedman 1988). Our approach distinguishes two *coerced* temporal structures: *coerced narrations* and the *coerced comments*. In turn, the coerced comments may contain inserted narrative structures in their bodies. These narrative structures are encapsulated substructures in the entire temporal structure of the text which have their internal organisation.

The identification of these structures is carried out by 91 context-scanning and 323 association control rules. The former assign aspecto-temporal values to each utterance in the text. The latter control the association of the next value in the text which has to be integrated in the temporal structure built so far. The dynamic modifications undergone by the temporal structures at each step of their construction are represented in a transition network containing 7 nodes and 13 arcs.

The algorithm has been tested manually over a corpus of 98 texts from different domains. The texts in the corpus were selected according to one criterion: the number of verbs in past tense had to be superior or equal to the number of all other tensed verb forms. This criterion applied exclusively to the verb forms in main clauses.

The details of the approach will be discussed in the following sections. Section 1 presents the notions of aspecto-temporal values as constituent elements of temporal structures and the strategy for their identification. Section 2 shows the rules of interaction of aspecto-temporal values in a text and the mechanism of identification of the temporal structures. Section 3 discusses

the properties of the identified temporal structures.

Aspecto-temporal values of situations and their identification

Definition and representation of aspecto-temporal values

Aspecto-temporal (AT) values of situations are semantic values assigned to predicative relations. They contain three types of information:

1. they encode the internal structure of situations (durative, punctual, etc.);
2. they fix the position of the situations with regard to the speaker's position (past, future, etc.);
3. they locate the time period where the speaker considers the situations as relevant (e.g. *Jane wrote a letter* vs. *Jane has written a letter*).

The values which denote the internal structure of situations are called aspectual values. According to their aspectual value, all situations fall into one of the three universal classes: *states*, *events* and *processes*, Desclés (1980; 1990; 1995), Guentcheva (1989; 1990). The three classes are represented by topological intervals. The topological representation reflects the intuitive segmentation of the linguistic temporal space which is based on the notions of continuity and transition, Desclés (1989; 1991)¹.

States denote situations with no explicitly marked beginning and end. They are represented by open intervals where the value of the predication x holds at every point of the interval, except for its left and its right bound. In the notation below, t_2 is last point of the situation preceding the state, t_1 is the first point of the situation following the state, t_0 is the time of utterance.

- (1) De Falla était compositeur espagnol.
De Falla was a Spanish composer



$$state(]t_2, t_1[) = \{x, t_2 < x < t_1\}$$

Processes express ongoing situations with an explicitly marked beginning (initial break), but without an explicitly specified end (final break). They are represented by semi-open intervals, closed to the left and open to the right. The predication x defined as a process holds over the whole interval except for its right bound. (t_2 = first point of the process, t_1 = first point of the situation following the process, t_0 = time of utterance).

¹The presented approach deals exclusively with the 3 aspecto-temporal parameters of the situations formulated in the beginning of this section. For the various aspectual phenomena resulting from the distribution of the predicates over their arguments, see (Verkuyll 1993).

- (2) De Falla écrivait une pantomime quand il a rencontré son impresario.
De Falla was writing a pantomime when he met his impresario



$$process([t_2, t_1[) = \{x, t_2 \leq x < t_1\}$$

Events describe situations perceived as an indivisible whole, containing initial and final break. They are represented by closed intervals, the predication x holding over the whole interval, bounds included. (t_2 = first point of the event, t_1 = last point of the event, t_0 = time of utterance).

- (3) De Falla est allé en Argentine en 1937.
De Falla went to Argentina in 1937.



$$event([t_2, t_1]) = \{x, t_2 \leq x \leq t_1\}$$

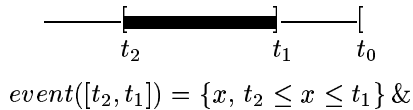
When the speaker produces an utterance, s/he locates the situation with regard to the time of utterance (t_0) and the aspectual value of the situation becomes aspecto-temporal. The above examples express a state, a process and an event which are anterior to the time of utterance.

The hashed intervals correspond to the *intervals of validation*. These intervals denote the segments of the temporal axis where the speaker considers the situation as relevant². In the examples above, the validation intervals overlap with the intervals of realisation of the situations. Yet, the interval of validation can be situated in a different segment of the temporal axis (example 4). In this case, the situation will be represented by a pair of related intervals where the hashed open interval ($]t_1 t_0[$) will stand for the interval of validation and the blank closed one ($[t_2 t_1]$) - for the interval of realisation of the situation:

²The intervals of validation can be seen as an extension of the Reichenbachian reference point R (Reichenbach 1947 second edition 1966). One advantage of representing the time of relevance as an interval is that it can receive different aspectual values (state vs. process vs. event) and, respectively, be assigned different types of intervals. These intervals may not necessarily overlap with the intervals of realisation of the situations (see example 4). For some other advantages of this notation see (Vazov 1998), ch.2.

- (4) La guitare espagnole est déjà devenue synonyme de la guitare classique.

The classical guitar has already become synonymous with the Spanish Guitar.



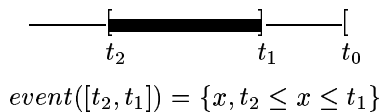
$$resulting\ state(]t_1, t_0[) = \{x, t_1 < x < t_0\}$$

The configuration of intervals in example 4 denotes that the classical guitar became synonymous with the Spanish one in a period of time anterior to the time of utterance (t_0). This event is marked by the closed interval ($[t_2 t_1]$) where $t_1 < t_0$. However, the speaker considers as relevant not the event itself but its result at the time of utterance. The result of the past event, i.e. the state of being synonymous, is represented by the open interval ($]t_1 t_0[$) whose right bound coincides with the utterance time (t_0). Hence, the situation in example 4 receives the value *resulting state at the time of utterance*³.

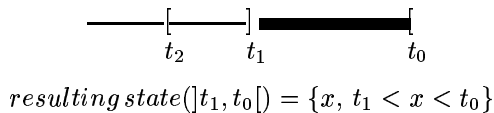
Sometimes the aspecto-temporal information in a clause is insufficient for the assignment of one particular AT value to the situation. In this case we will consider the situation as ambiguous and will assign it an indeterminate value:

- (5) La guitare classique est devenue synonyme de la guitare espagnole.

The classical guitar became synonymous with the Spanish Guitar.



The classical guitar has become synonymous with the Spanish Guitar.



This ambiguity is usually removed, if the AT value of the situation is considered in a larger context (text unit) where it interacts with the values of the neighbouring clauses.

³Our approach considers the utterance as a process rather than as a point. Hence, its corresponding semi-open interval $[t'_0 \text{---}]t_0$ where t'_0 stands for the beginning of the utterance and t_0 - for the first point of non-realised flow of time (Benveniste 1962), (Desclés 1973), (Steedman 1997). To simplify the notation in the paper, we shall represent the utterance by ($]t_0$). Thus, if a situation is defined as holding or relevant at the utterance time, the right bound of its interval of validation should overlap with ($]t_0$).

Identification of aspecto-temporal values

The identification of the AT values of situations is carried out by a context-scanning strategy (Berri, Maire-Reppert, & Oh 1991), (Maire-Reppert 1990), (Desclés *et al.* 1991), (Oh 1991), (Jouis 1994), (Vazov 1998). This strategy is based on the hypothesis that the semantic representations are not necessarily determined by world knowledge but can be considered as configurations integrating elements of local linguistic information. The interaction of these elements is controlled by a set of mechanisms which build the semantic representation by putting together all the pieces of local information. The elements of this representation can be derived from grammar or lexical semantics and from some syntactic phenomena. By the latter we shall understand:

- the position of aspecto-temporal markers within the clause
- the identity of grammatical subjects and/or objects in compound sentences
- the position of the clause within the text

The context-scanning strategy has a two-step organisation.

The first step establishes the list the markers to be detected by the scanning procedure and formulates the rules of their interaction. Every context-scanning strategy makes use of two types of markers: *triggering markers* and *contextual markers*. The triggering markers represent the core elements in the semantic configuration. Once identified, the triggering markers launch the searching procedure for contextual markers in the text. An example of a triggering aspecto-temporal marker is the morpheme of a finite (conjugated) verb form; a temporal adverbial detected in the same utterance is considered as a contextual marker.

The interaction of the information provided by both markers is encoded in a set of production rules. These rules are expressions in propositional logic ($A \rightarrow B$) in which A is the premise (triggering and contextual markers) and B is the conclusion (a semantic value obtained in the presence of the markers).

The second step is the application of the context-scanning procedure. Using a complete set of production rules, the strategy launches a search engine whose goal is to detect the markers in a given linguistic context. If the search engine finds relevant markers, the semantic values become automatically assigned to the linguistic forms.

An example of a real context-scanning rule for identification of the value resulting state at the time of utterance is given below. It specifies the triggering marker (form of passé composé) and the contextual marker (adverbial déjà 'already') of the searching procedure:

Rule 1: (applies to example 4)

IF a morpheme of the grammatical tense passé composé has been detected in the clause C

tense	value	comments
<i>passé simple</i>	past event	
<i>plus-que-parfait</i>	past event	its interval precedes the interval of the previous event value in the text
<i>imparfait</i>	past process	non-static verbs
<i>imparfait</i>	past state	static verbs
<i>présent</i>	present process	non-static verbs
<i>présent</i>	present state	static verbs
<i>futur</i>	future situation	unspecified at this stage of the study

Table 1: Default AT values assigned to the sentences

AND IF an element of the list *temporal adverbials indicating resulting state* has been detected in the clause C

THEN the situation in the utterance is assigned the value *resulting state at the time of utterance*

We have formulated 65 rules of this type. The goal of the rules is to identify the values of the clauses containing forms of *passé composé*. These forms are also the triggering markers in our context-scanning strategy. The choice of the triggering marker has been determined by the following 3 factors:

1. Texts referring to past events in French can be built out exclusively of forms of *passé composé*⁴.
2. *Passé composé* is the only grammatical tense whose forms can receive an event or a state value (see examples 3 & 4). The first one moves forward the narration in a text and the second one expresses a descriptive situation (Gagnon 1993).
3. In a text, *passé composé* can cooccur with all other grammatical tenses.

The above factors make the values of *passé composé* central for the temporal structures of texts expressing past events.

If the texts contain forms of other tenses, they will be assigned a default value as follows:

Our approach uses contextual markers of various types. We have collected 8 lists of adverbials:

1. temporal event adverbials (110), e.g. *il y a trois mois* (three months ago)
2. non-temporal event adverbials (14), e.g. *soudain* (suddenly)
3. adverbials of iteration (8), e.g. *souvent* (often)
4. future adverbials (6), e.g. *bientôt* (soon)
5. adverbials of experience (2), e.g. *jamais* (never)
6. resulting state adverbials (12), e.g. *enfin* (finally)
7. "reset" adverbials, or adverbials which start a new temporal structure (6), e.g. *d'autre part* (on the other hand)
8. link adverbials, or adverbials which link the values of two clauses in a sequence of events (22), e.g. *ensuite* (then)

⁴*Passé simple* and *présent* also allow for such a construction. The identification of their temporal structures is trivial as these forms always describe strictly successive events and the clauses where they occur usually contain explicit temporal markers. Examples of such texts are encyclopaedia entries (for *passé simple*) and direct sports broadcasting (for *présent*).

Another large group of contextual markers are some sets of verbs: static, declarative, verbs of perception, etc. For a complete list of contextual markers, see (Vazov 1998).

Our context-scanning strategy allows to fix a finite number of AT values. Thus, the main clauses in our corpus will always receive one of the 8 following values (Table 1):

1. past state
2. past process
3. past event
4. resulting state in the time of utterance
5. indeterminate
6. present state
7. present process
8. other (future events/states/processes, conditional, interrogative, imperative).

Identification of temporal structures

General principles

In the introduction we defined the temporal structure as a set of AT values of all main clauses in a text. However, this definition does not capture the relations among the elements of the set. In this section we shall represent the temporal structure as an applicative expression whose operators and arguments are the values of all main clauses in a text. This applicative expression encodes the order of application of these values as operators. For example, according to the relations among its values, the temporal structure of a text containing 4 main clauses, respectively 4 values $\{V_1, V_2, V_3, V_4\}$, could be represented as: $((V_1 V_2) V_3) V_4$, $((V_3 V_4) V_1) V_2$, etc. Hence, the identification of the temporal structure of a given text could be seen as construction of applicative expression encoding the relations among the AT values in this text.

The procedure of construction of applicative expressions is incremental and can be represented as a program consisting of two basic operations. The first operation identifies the value in each clause using context-scanning rules (see previous section). The second one associates the identified value to the previous value or structure. This procedure is carried out by a set of association rules.

values of clause 1	values of clause 2							
	past state	past process	past event	indeterminate	resulting state	present state	present process	other
past state	description $V_1 V_2$	description $V_1 V_2$	narration $V_1 V_2$	narration $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$
past process	description $V_1 V_2$	description $V_1 V_2$	narration $V_1 V_2$	narration $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$
past event	narration $V_1 V_2$	narration $V_1 V_2$	narration $V_1 V_2$	narration $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$
indeterminate	narration $V_2 V_1$	narration $V_2 V_1$	narration $V_2 V_1$	description $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	comment $V_1 V_2$
resulting state	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	comment $V_1 V_2$
present state	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	\emptyset	\emptyset	comment $V_1 V_2$
present process	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	description $V_1 V_2$	description $V_1 V_2$	\emptyset	\emptyset	comment $V_1 V_2$
other	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	comment $V_1 V_2$	\emptyset

Table 2: Structures resulting from the association of the values in clause 2 to the values of clause 1

Association rules for initial temporal structures (value $_1$ + value $_2$)

The association of values of the first 2 clauses sets up the initial temporal structure. The total number of rules controlling this structure will be $(8 \times 8) - 5 = 59$, or one⁵ for each pair of possible values in clause 1 and 2. Table 2 represents a chart of the possible associations of the value in clause 2 to one of the 8 values in clause 1. The value of the second clause in the initial structures is always a left-associative operator (see applicative expressions assigned to each temporal structure).

The text in 6 is an example of a *narration* obtained by the combination of the value ‘indeterminate’ of the first clause and the value ‘past event’ in the second clause (see Table 2).

- (6) [Il m’a écouté en fumant et en hochant la tête.
Puis il m’a demandé de la relire.]_{narration}
He listened to me smoking a cigarette and nodding his head.
Then, he asked me to read it again.

The association rules for the first two values control the construction of three initial temporal structures: *descriptions*, *narrations* and *comments*. These structures are further completed by the values of each next clause in the text.

Rules for association of values to the initial structures

Non-coerced structures. Table 3 shows the temporal structures which result from the association of the values in clause 3 to the three possible initial structures of clauses 1 & 2.

As we pointed out in the introduction, the linear order of values in the text may or may not correspond to

⁵We don’t consider texts whose first two clauses do not contain forms of past tense.

the order of their application as operators. Thus, some of the temporal structures in table 3 preserve their initial type. In these structures the value of the third clause is associated to the initial temporal structure by association to the left ($V_1 V_2 V_3$). This recursion (which can be repeated a number of times) is typical for narrative texts, though instances of such structures can be provided by some descriptive texts as well.

An example of an association to the left is the association of a ‘past event’ (value of clause 3) to the initial structure of ‘narration’ (example 7):

- (7) [[Il m’a écouté en fumant et en hochant la tête.
Puis il m’a demandé de la relire.]_{narration}
Il a été tout à fait content.]_{narration}
He listened to me smoking a cigarette and nodding his head.
Then, he asked me to read it again.
He felt very satisfied.
(A. Camus, *L’étranger*, Éditions Bordas, p.53)

Coerced structures. The initial temporal structure can be potentially modified by the third (or any next) value in the text. Each modification will then reflect in a rearrangement of operators in the applicative expressions. The rearrangement indicates that the new value is not left-associative but applies as an operator to the already existing initial structure and *coerces* it into a different structure. The coerced structures can be divided into two groups with respect to the presence or the absence within their bodies of inserted narrative substructures.

An example of a coerced structure without an inserted narrative substructure is the configuration which results from the association of a ‘past process’ value to the initial structure of *description* (see table 3, the first of the three possible resulting structures):

clauses 1 & 2	values of clause 3							
	past state	past process	past event	indeter- minate	resulting state	present state	present process	other
narration $V_1 V_2$	narration $V_1 V_2 V_3$	narration $V_1 V_2 V_3$	narration $V_1 V_2 V_3$	narration $V_1 V_2 V_3$	comment with inserted narration $V_3(V_1 V_2)$	comment with inserted narration $V_3(V_1 V_2)$	comment with inserted narration $V_3(V_1 V_2)$	comment with inserted narration $V_3(V_1 V_2)$
description $V_1 V_2$	coerced narration $V_3 V_1 V_2$ or coerced comment $V_3 V_1 V_2$ or comment with inserted narration $V_1(V_3 V_2)$	coerced narration $V_3 V_1 V_2$ or coerced comment $V_3 V_1 V_2$ or comment with inserted narration $V_1(V_3 V_2)$	coerced narration $V_3 V_1 V_2$ or coerced comment $V_3 V_1 V_2$ or comment with inserted narration $V_1(V_3 V_2)$	description $V_1 V_2 V_3$	description $V_1 V_2 V_3$	description $V_1 V_2 V_3$	description $V_1 V_2 V_3$	coerced comment $V_3 V_1 V_2$
comment $V_1 V_2$	comment $V_1 V_2 V_3$ or comment with inserted narration $V_1(V_3 V_2)$	comment $V_1 V_2 V_3$ or comment with inserted narration $V_1(V_3 V_2)$	comment $V_1 V_2 V_3$ or comment with inserted narration $V_1(V_3 V_2)$	comment $V_1 V_2 V_3$	comment $V_1 V_2 V_3$	comment $V_1 V_2 V_3$	comment $V_1 V_2 V_3$	comment $V_1 V_2 V_3$

Table 3: Association of the 8 possible values to the 3 initial temporal structures

- (8) [Si ces révélations n’ont pas fait grand bruit en France, elles ont quand même suscité quelques réactions.

Le Monde a rappelé avec passion ce qu’était l’affaire Rosenberg, la comparant au cas de Sacco et Vanzetti.]*description*
Though these "revelations" haven't stirred France, they have still provoked several reactions. Le Monde has passionately mentioned what the Rozenberg case was about, comparing it to Sacco and Vanzetti.

→→

[Si ces révélations n’ont pas fait grand bruit en France, elles ont quand même suscité quelques réactions. Le Monde a rappelé avec passion ce qu’était l’affaire Rosenberg, la comparant au cas de Sacco et Vanzetti.

Quant à l’hebdomadaire de droite Valeurs actuelles, son article du 19 août 1995 proclamait: "Les Rosenberg étaient bien coupables."]*narration*
Though these "revelations" didn't stir France, they did provoke several reactions.

Le Monde passionately mentioned what the Rozenberg case was about, comparing it to Sacco and Vanzetti.

As for the right weekly Valeurs actuelles, its article of Aug. 19 proclaimed bluntly: The Rozenberg were absolutely guilty.

(*Le Monde on CD*)

In turn, the association of a ‘present state’ value to the initial structure of *narration* results in a *comment with inserted narration*. The applicative expression in the parentheses ($V_1 V_2$) (see table 3) represents an inserted narrative structure in the body of a comment⁶:

- (9) [...Jadis, j’avais constamment des problèmes avec mon patron et mes collaborateurs. Mais maintenant, la situation a soudain changé.]*narration*

→→

[...Jadis, j’avais constamment des problèmes avec mon patron et mes collaborateurs. Mais maintenant, la situation a soudain changé.]*narration*

Mon patron et mes collègues tiennent à présent compte de mes idées.]*comment*

... In the past, I had problems with my boss and my colleagues all the time.

But now, the situation changed all of a sudden.

My boss and my colleagues do not ignore my ideas and proposals any more.

(*leaflet Hanussen, January 1996*)

⁶The number of propositions in the inserted narrative structure can be arbitrary. Any following value which does not allow the association to the left will mark the end of this structure and will resume the structure of the comment, respectively. The entire inserted structure plays the role of a single operand in the applicative expression of the entire structure of comment.

description		the value to be associated - V_n	new structure
V_{m-1}	last value - V_m		
indeterminate	indeterminate	past (state, process,event)	coerced narration $V_n V_1 \dots V_m$
\neg indeterminate	indeterminate	past (state, process,event)	comment with inserted narration $V_{m-1}(V_n V_m)$
\neg indeterminate	\neg indeterminate	past (state, process,event)	coerced comment $V_n V_1 \dots V_m$

Table 4: Possible associations to the structure of description

comment	the value to be associated - V_n	new structure
last value - V_m		
past (state, process,event) or indeterminate	past (state,process,event)	comment with inserted narration $V_{m-1}(V_n V_m)$
\neg past (state, process,event) or \neg indeterminate	past (state,process,event)	comment $V_{m-1} V_m V_n$

Table 5: Possible associations to the structure of comment

The scope of operation of ‘past’ values in the structures with inserted narration. The associated values in the structures with inserted narrative substructures can have different scope of operation. Thus, the value of ‘present state’ will apply to the entire initial structure of *narration* (example 9). Their corresponding applicative expression will be of the form $V_3(V_1 V_2)$. In turn, the value of ‘past event’ will operate only over the last value in the initial structure of *description* (example 10) and will be represented by the expression $V_1(V_3 V_2)$.

- (10) [... Le PDG est M.Andras Balazs, la quarantaine passée. Fils d’agriculteur, il a eu son premier emploi dans une ferme d’Etat.]*description*
 $\rightarrow \rightarrow$
 [... Le PDG est M.Andras Balazs, la quarantaine passée... [Fils d’agriculteur, il a eu son premier emploi dans une ferme d’Etat. Puis, il a travaillé pour son copmte...] *narration*] *comment*
The president of the company, M.Andras Balazs, is in his forties.
Son of a field worker, he had his first job in a state-owned farm.
Then, he started working for his own account.
(Le Monde Diplomatique on CD-ROM)

As seen in table 3, the association of the values ‘past state’, ‘past process’ and ‘past event’ to the preceding structures of *description* and *comment* produce different temporal structures according to the values which constitute the initial structures. The rules of association of these values are given in table 4 and table 5, respectively.

The associated past values modify the preceding

structure starting from the last value and moving to the beginning of the structure. If the last value (m) in the preceding structure is indeterminate, they will coerce it into a ‘past event’. Then they will check if the value ($m-1$) is also indeterminate and will coerce it to a ‘past event’ too.. If all the values in the preceding description are indeterminate (example 8), the text will be incrementally converted into a coerced narration whose corresponding applicative expression is shown in the column “new structures” of Table 4. In turn, if the last value of the preceding structure is not indeterminate, it the entire text will be automatically converted into a coerced comment (exemple 11).

- (11) [... Le Conseil d’Administration de l’Université Paris XI a décidé de limiter les crédits de fonctionnement de l’Université à 80% du budget 1995. Ils préparent ainsi le terrain aux restrictions budgétaires.]*description*
 $\rightarrow \rightarrow$
 [... Le Conseil d’Administration de l’Université Paris XI a décidé de limiter les crédits de fonctionnement de l’Université à 80% du budget 1995. Il prépare ainsi le terrain aux restrictions budgétaires. Il y a deux ans, beaucoup de postes ont été supprimés et les budgets réduits.] *comment*
The Administrative Council at the University Paris XI has decided to limit the University operating budget to 80% of the 1995 budget.
They are thereby preparing for further budget cuts.
Two years ago many teaching positions were closed and the budgets were reduced.
(Leaflet distributed at the University Paris XI)

coerced structure	values of clause 4							
	past state	past process	past event	indeterminate	resulting state	present state	present process	other
coerced narration $V_3 V_1 V_2$	coerced narration $V_3 V_1 V_2 V_4$	coerced narration $V_3 V_1 V_2 V_4$	coerced narration $V_3 V_1 V_2 V_4$	coerced narration $V_3 V_1 V_2 V_4$	comment with inserted narration $V_4(V_3 V_1 V_2)$	comment with inserted narration $V_4(V_3 V_1 V_2)$	comment with inserted narration $V_4(V_3 V_1 V_2)$	comment with inserted narration $V_4(V_3 V_1 V_2)$
comment with inserted narration $V_3(V_1 V_2)$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$	comment $V_3(V_1 V_2)V_4$
comment with inserted narration $V_1(V_3 V_2)$	comment with inserted narration $V_1(V_3 V_2 V_4)$	comment with inserted narration $V_1(V_3 V_2 V_4)$	comment with inserted narration $V_1(V_3 V_2 V_4)$	comment with inserted narration $V_1(V_3 V_2 V_4)$	comment $V_1(V_3 V_2)V_4$	comment $V_1(V_3 V_2)V_4$	comment $V_1(V_3 V_2)V_4$	comment $V_1(V_3 V_2)V_4$
coerced comment $V_3 V_1 V_2$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$	comment $V_3 V_1 V_2 V_4$

Table 6: Associations of values to coerced structures

The middle row of Table 4 shows a configuration where the last value (m) is indeterminate and the others ($m-1, \dots, n$) are not. In this case, the ‘past’ value will apply only to this value and will create an inserted narrative structure (example 10). In the entire temporal structure of the text this narrative structure will play the role of a single operand to the preceding descriptive structure. The result of the application of the temporal structure of description to the temporal structure of narration will be the structure of comment with inserted narration. The backward application of the ‘past’ value will repeat until it becomes blocked by the first non-indeterminate value in the preceding structure. Then all the indeterminate values processed so far will be integrated in the inserted narration substructure of the comment.

- (12) [... Onze personnes ont été brûlées vives vendredi 8 janvier á Bombay.
Les trois jours de violence ont déjà fait 66 morts et plus de cent blessés..]*comment*
→→
[... ... Onze personnes ont été brûlées vives vendredi 8 janvier á Bombay.
Les trois jours de violence ont déjà fait 66 morts et plus de cent blessés..
Des affrontements ont eu lieu vendredi dans quatre secteurs de la capitale économique de l’Inde.]
comment
Eleven people were burnt alive on January 8th, Friday, in Bombay.
66 people have died and more than a hundred have been injured during the three days of violence.
Clashes between the riot police and the protestors were reported Friday in the industrial capital of India.
(Le Monde on CD-ROM)

The association of ‘past’ values to the structure of

comment is carried out following the same principles (Table 5). The past values operate backwards on all past *and* indeterminate values in the preceding structure and integrate them into a narration inserted into the body of a comment. If the last value of the preceding structure is a non-past or a non-indeterminate value, the ‘past’ value will be attached to this structure by association to the left. The result of this association will be a new comment (example 12).

Association of values to coerced structures. All temporal structures in Table 3 which do not undergo coercion after the association of the third value preserve the characteristics of the initial structure of value₁ + value₂. Thus, the structure of narration will result in a new narration when it associates a value ‘past event’. The new narration will replace the old one in the chart and will apply recursively to the next values unless modified by one of the four values in the chart (‘resulting state’, ‘present state’, ‘present process’ or ‘other’). Descriptions and comment will behave similarly and will preserve the initial structure until the end of the analysed text.

Table 6 shows the chart of associations of the coerced values in Table 3.

According to the following value, the coerced narration can either preserve its characteristics or be entirely coerced into a comment with inserted narration (first row of Table 6).

The second row of the same table shows the possible associations of values to the comment with inserted narration. This type of comment can result either from the initial structure of narration (Table 3) or from the coerced narration given in the top row of Table 6. The inserted narration in this structure is an argument of the last value of the previous structure (for example, $V_3(V_1 V_2)$) and is not accessible to the next values in the text. Thus, the comment with inserted narration of this type necessarily transforms into a comment, if a

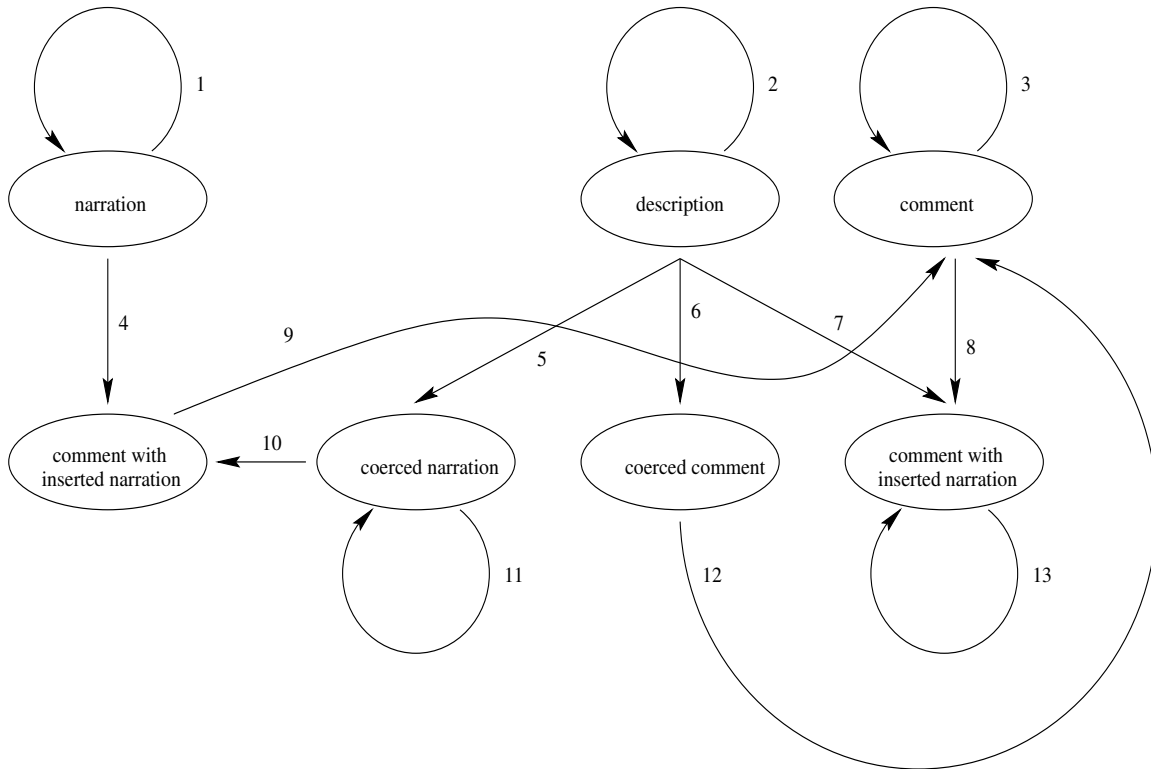


Figure 1: Network of possible associations in the temporal structures

new value is associated to it. Yet, it should be pointed out that the new temporal structure assigned to the text does not deny the existence of the narrative substructure within its body. The label ‘comment’ will only denote that the temporal structure will behave as a comment, should other values be further associated to it.

The comment with inserted narration given in third row of Table 6 is a result of application of a ‘past’ value to the initial structure of description or comment. The ‘past’ values coerce the last indeterminate values in the description or in the comment into ‘past events’ and create a narrative substructure. Unlike the inserted narrations which result from coercing narrations (see previous paragraph), these inserted structures remain open to the next values in the text. Thus, according to their type, the next values in the text can extend either the inserted narration substructure, or the structure of comment. The ‘past’ values will become incorporated in the narration whereas the values of ‘resulting state’, ‘present state’, ‘present process’ or ‘other’ will associate to the entire structure of comment. Their corresponding expressions will have the following form: $V_1(V_3V_2V_4V_n)$ for the extended inserted narration and $V_1(V_3V_2V_4)V_n$ for the extended comment.

Finally, the coerced comment preserves the structure of comment regardless of the value associated to its body.

All possible associations are represented by the fol-

lowing transition network (Figure 1). It shows that the temporal structure of a text is an algorithm containing 13 possible transitions (arcs) which link 7 possible temporal structures. All the nodes in the network allow either a recursion or a passage to another temporal structure thus providing the possibility to process texts containing infinite number of phrases without increasing the number of possible temporal structures.

Temporal structures

This section presents the internal organisation of the temporal structures identified by association control rules.

Description structure displays the following features (example 13 & Figure 2):

1. All validation intervals of description are situated in the same zone of the temporal axis of the speaker. A zone of temporal axis is any of the three time periods which are situated before, during or after the utterance time.
2. The structure of description is represented by the intersection of all intervals of validation in the text. This intersection is called common interval of validation.
3. The common interval is the biggest open interval within the intersection. Its right bound coincides with the right bounds of all validation intervals of the situations in the text.

- (13) [Meurtre d'une call-girl à Los Angeles.
 Deux flics mènent l'enquête dans une compagnie japonaise.
 Au bouillant Web Smith, noir et intrépide, a été adjoint John Connor, féru de culture japonaise.]^{description}
Murder of a call-girl in Los Angeles.
Two cops are investigating a Japanese company.
Web Smith, a ferulous, hotheaded black cop, has been teamed up with John Connor, a buff of Japanese culture.
 (Télérama, N.2455, 29.01.1997)

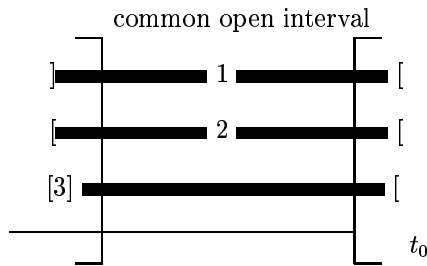


Figure 2: Description valid at the time of utterance

The temporal structure of *narration* has the following characteristics:

1. All validation intervals of narration should be situated in the same zone of the temporal axis of the speaker.
2. Narration should contain at least one of the two configurations:
 - two closed intervals of validation which represent the events as ordered, simultaneous or including each other⁷
 - one closed and one open/semi-open interval of validation

The text in example 7 has a temporal structure of narration. Its structure is represented graphically in Figure 3.

The temporal structure of *comment* imposes only one condition to the position of its constituents validation intervals: comment contains at least two validation intervals which are situated in different zones of the temporal axis of the speaker, for example in the zone of past *and* in the zone of present relevance. The situations within the comment structure lack internal temporal organisation. The aspecto-temporal values in this temporal structure do not interact with the values in the neighbouring utterances. The situations in comments

⁷In some earlier publications, Vazov (1997; 1999), the narration was defined as containing at least two *disjoint* closed intervals of validation. The analysis of larger text corpora showed that the weaker definition - two closed intervals of validation - allows to capture a wider range of texts with narration structure, namely those using other discourse techniques like flashback, explanation, elaboration and background, Asher and Lascarides (1991; 1993).

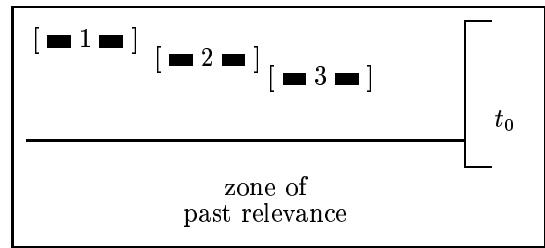


Figure 3: Narration

are linked exclusively by semantic factors, like shared topic.

Conclusion

The main goal of this paper was to provide evidence that the organisation of the temporal information in the texts has to obey a finite set of dynamic combinatory rules. These rules are combinatory as they determine the interaction of the already analysed temporal structure of the text and the semantic value of the upcoming clause. The rules are activated and selected every time a new value has to be integrated into the temporal structure, hence their dynamic nature. The third important feature of these rules is their finite number. It allows to define a “grammar” of temporal structures based on a fixed number of predictable temporal structures assigned to texts of arbitrary length.

The information about the temporal structure of a text has an important role in temporal reasoning as it allows to classify the texts with regard to the way they describe the scenes in the world of the speaker. For instance, the description temporal structures depict scenes of reality as a snapshot. Hence, they cannot express succession of events or causal relations between the described situations. In turn, narration structures are capable of expressing implicit causal relations due to the total/partial ordering of the events which constitute this structure (Battistelli & Vazov 1997b), (Battistelli 2000).

Identification of temporal structures will reduce the number of texts which require further analysis of the temporal order of situations. This analysis will not apply to descriptions where all situations are relevant over the same period of time, or to comments where situations are temporally independent.

References

- Asher, N., and Lascarides, A. 1991. Discourse relations and defeasible knowledge. In *29th Meeting of the Association for Computational Linguistics*, 55–62.
- Asher, N., and Lascarides, A. 1993. Temporal interpretation, discourse relations and common sense entailment. *Linguistics and Philosophy* 16(5):437–493.
- Battistelli, D., and Vazov, N. 1997b. Building of semantico-cognitive representation of different types of narratives in french. In *RANLP'97*, 373–378.

- Battistelli, D. 2000. *Passer du texte à une séquence d'images, analyse spatio-temporelle de textes, modélisation et réalisation informatique (système SPAT)*. Ph.D. Dissertation, Université Paris - Sorbonne.
- Benveniste, E. 1962. *Problèmes de linguistique générale*, volume 1. Éditions Gallimard.
- Berri, J.; Maire-Reppert, D.; and Oh, H. 1991. Traitement informatique de la catégorie aspecto-temporelle. *T.A.Informations* (1):77-90.
- Desclés, J.-P.; Jouis, C.; Oh, H.-G.; Reppert, D.; and Segond, F. 1991. Représentations et recherches des valeurs sémantiques des temps de l'indicatif du français pour une mise en oeuvre informatique. Technical Report 957, CAMS - Université Paris IV, Sorbonne.
- Desclés, J.-P. 1973. Formalisation des opérations énonciatives. In *Modèles logiques et niveaux d'analyse linguistique*. Metz, France: PUF, Klincksieck.
- Desclés, J.-P. 1980. *Construction formelle de la catégorie grammaticale de l'aspect (essai)*. Klincksieck. 198-237.
- Desclés, J.-P. 1989. State, event, process and topology. *General Linguistics* 29(3):159-200.
- Desclés, J.-P. 1990. *Langages applicatifs, langues naturelles et cognition*. Paris: Hermès.
- Desclés, J.-P. 1995. Transitivity sémantique versus transitivity syntaxique. In *Colloque "Transitivité"*.
- Gagnon, M. 1993. *Expression de la localisation temporelle dans un générateur de texte*. Ph.D. Dissertation, Université de Montréal.
- Guentchéva, Z. 1989. Implication aspecto-temporelles en français et en bulgare. *Linguistique comparative* XIV(5):26-37.
- Guentchéva, Z. 1990. *Temps et aspect: l'exemple du bulgare contemporain*. Éditions du CNRS.
- Jouis, C. 1994. Contextual approach: Seek, a linguistic tool and conceptual tool for use in knowledge acquisition. In *First European Conference "Cognitive Science and Industry"*, 259-274.
- Maire-Reppert, D. 1990. *L'imparfait de l'indicatif en vue d'un traitement informatique du français: étude théorique et méthode de description avec ses conséquences didactiques dans l'enseignement des langues*. Ph.D. Dissertation, Université Paris - Sorbonne.
- Moens, M., and Steedman, M. 1988. Temporal ontology and temporal reference. *Computational Linguistics* 14(2):15-28.
- Oh, H.-G. 1991. *Les temps de l'indicatif du français en vue d'un traitement informatique: passé composé*. Ph.D. Dissertation, Université Paris - Sorbonne.
- Reichenbach, H. 1947 (second edition 1966). *Elements of Symbolic Logic*. McMillan.
- Steedman, M. 1997. *Handbook of Logic and Language*. The MIT Press/ North Holland. chapter Temporality, 895-939.
- Vazov, N. 1997. Événement ou état, récit ou description - le rôle du contexte dans l'identification des valeurs aspectuo-temporelles. *Text Processing and Cognitive Technologies* 1:15-24.
- Vazov, N. 1998. *Identification des valeurs aspectuo-temporelles de situations en vue de traitement automatique. Le cas de passé composé*. Ph.D. Dissertation, Université Paris - Sorbonne.
- Vazov, N. 1999. Context-scanning strategy in temporal reasoning. In *Modeling and Using Context (Second International and Interdisciplinary Conference, CON-TEXT'99)*, 389-402. Springer-Verlag.
- Verkuyl, H. 1993. *A theory of aspectuality. The interaction between temporal and atemporal structure*. Cambridge Studies in Linguistics. Cambridge University Press.