

Temporalité linguistique et S-Langages

Delphine Battistelli, Jean-Luc Minel, Etienne Picard et Sylviane R. Schwer

Laboratoire LaLICC UMR 8139 (Université Paris IV, CNRS)
96, bd Raspail, 75006 Paris -France
Prénom.Nom@paris4.sorbonne.fr

Résumé – Abstract

Après un rappel de la problématique de l'ordonnement temporel dans un texte, nous décrivons les S-langages qui offrent une représentation unifiée des relations temporelles et une opération (la jointure) permettant de calculer les combinaisons entre celles-ci.

After a brief overview of the problem of text temporal structures calculus, we describe the formal language of S-languages. This framework offers an unified representation of temporal relations and an operation (the joint) which computes combinations between such relations.

Keywords – Mots Clés

Structure(s) temporelle(s) dans un texte narratif, S-langages.
Narratives temporal structure(s), S-langages.

1 Introduction

Notre travail se situe en aval de l'analyse aspecto-temporelle d'un texte effectuée par un linguiste, c'est-à-dire que nous avons à notre disposition l'expression d'un ensemble de relations temporelles entre procès dans une théorie linguistique donnée – dans le cadre des intervalles topologiques de Desclés (1989), de la DRT de Kamp (1981) ou encore du système d'intervalles de Gosselin (1996) – permettant la définition de l'extension temporelle des procès comme chaînes de points et/ou d'intervalles. Notant que pour la plupart de ces modèles, ces relations sont très souvent binaires et que les descriptions sont souvent incomplètes (au sens où plusieurs configurations temporelles peuvent par exemple être associées à une description donnée), nous nous assignons les deux objectifs suivants : (i) situer les procès *globalement* les uns par rapport aux autres ; (ii) expliciter toutes les relations potentielles entre procès. Nous rappelons dans la section 2 les principaux aspects de la problématique linguistique et nous décrivons dans la section 3 le formalisme des S-langages qui permet de prendre en charge certaines caractéristiques de la problématique linguistique.

2 Contraintes linguistiques et ordonnancement temporel

Les travaux linguistiques dans le domaine de la temporalité¹ se concentrent sur la mise en évidence des modes d'expression du temps, sur l'analyse de leurs modes d'interaction et sur la recherche des principes (ou contraintes) d'organisation temporelle des procès dans un texte. Ces travaux mettent alors en exergue le rôle des connecteurs ou adverbess temporels mais aussi celui des valeurs aspectuelles attachées aux propositions que l'on cherche à relier temporellement. Sans recourir à la terminologie propre à tel ou tel modèle linguistique, nous illustrons les principaux aspects de la problématique de l'ordonnancement temporel à travers les exemples (1) à (5).

- (1) *Zoé est partie* (P1). *Félix est arrivé* (P2).
- (2) *Félix est arrivé* (P1) *après que Zoé est partie* (P2).
- (3) *Félix est arrivé* (P1). *Zoé était partie* (P2).
- (4) *Félix est arrivé* (P1). *Zoé partait* (P2).
- (5) *Etant arrêté sur la file droite* (P1), *j'avais mis mon clignotant* (P2).

Dans l'exemple (1), l'ordre temporel entre les propositions P1 et P2 correspond à leur ordre linéaire : on a ainsi Sit1<Sit2, Sit1 et Sit2 étant les situations dénotées respectivement par P1 et P2 et '<' la relation temporelle de précédence. S'il arrive qu'il y ait comme ici isomorphisme exact entre les deux structures (la structure du référentiel narratif et la structure chronologique du référentiel externe), c'est cependant loin d'être toujours le cas, qu'il s'agisse par exemple de retours en arrière ou bien de relations de recouvrement partiel ou total entre situations. Ainsi, dans les exemples (2) et (3), c'est l'ordre temporel Sit2<Sit1 qui prévaut : dans (2), il est imposé par la marque explicite « *après que* » ; dans (3), il découle de l'utilisation du passé-composé pour P1 et du plus-que-parfait pour P2. Dans (4) et (5), on a Sit2 qui recouvre temporellement Sit1, ce qui est initié par le contraste aspectuel entre le passé composé et l'imparfait pour (4) et entre le participe passé et le plus que parfait pour (5).

A l'instar de Desclés (et al.1994), Gosselin (1996) ou encore (Vet 1994), nous nous situons dans une approche cherchant résolument à mettre à jour le plus possible de contraintes marquées linguistiquement pour l'interprétation des relations temporelles² et nous attirons l'attention sur deux points : (i) ce ne sont pas les temps verbaux qui instruisent directement l'ordonnancement temporel des situations dénotées par les propositions mais les valeurs aspecto-temporelles de ces propositions ; (ii) la position qui consiste à considérer que des séquences narratives comme (1), (3), (4) et (5) ne contiennent pas d'indices *explicites* d'ordonnancement temporel est discutable : on peut en effet considérer que les temps verbaux à eux seuls constituent des marqueurs d'ordonnancement temporel, relativement bien sûr au

¹ Vue l'étendue des travaux dans ce domaine, il est impossible de couvrir de manière exhaustive les références possibles. Nous nous contenterons ici de renvoyer principalement à Desclés (1989), Gosselin (1996) et Kamp (1981) pour des travaux sur le français.

² En tenant compte en particulier de la sémantique lexicale des verbes, de l'organisation thématique du texte (Vet 1994) ou encore des marques de ponctuation. Certains auteurs travaillant dans le cadre de la DRT (Kamp 81) ont eux préféré considérer que les questions de chronologie relative étaient presque uniquement régies par des principes discursifs s'appuyant sur des connaissances du monde, ces principes étant directement inspirés de la RST (Mann et Thompson 1988). Bras et Asher (1994) ont ainsi abandonné l'opposition aspectuelle établie par Kamp entre état (associé à l'imparfait) et événement (associé au passé simple), notant alors que la principale difficulté devenait celle de l'identification des relations de discours.

type de valeurs aspectuelles qu'ils induisent (cf. remarque (i)). Visualisée selon un certain aspect, une situation est assujettie à une certaine logique des relations temporelles qui détermine l'acceptabilité de certaines combinaisons de notions aspectuelles et l'inacceptabilité d'autres combinaisons. Parmi les principes les plus importants qui gouvernent la logique des relations temporelles pertinentes du point de vue aspectuel, certaines peuvent alors être considérées comme axiomatiques (Lyons 1990, p. 329).

Un autre aspect essentiel des travaux linguistiques sur la temporalité a trait au type de relations temporelles qui sont envisagées entre propositions. On remarque ainsi qu'elles sont incomplètement spécifiées : en effet, si l'on se réfère aux 13 relations théoriquement possibles entre deux intervalles quelconques³, il est en général impossible de préciser si l'on a affaire à l'une ou l'autre de ces relations ; on a plutôt recours à un sous-ensemble de ces relations pour expliciter les relations possibles entre deux propositions. La relation de couverture, notée θ , introduite par Kamp (1981) illustre bien cette problématique. Elle traduit la contemporanéité des deux procès et dit exactement ceci : peu importe les situations respectives de début et de fin des procès, la seule information pertinente est qu'ils aient une partie commune. Considérons à cet effet l'exemple (5). Il sera représenté par P1 θ P2 car rien ne dit si le clignotant a été mis avant, en même temps ou après le début de l'arrêt, et l'histoire n'étant pas achevée, nous n'avons aucune autre information pour affiner la situation entre les deux bornes de droite. Une autre illustration de cette problématique peut être donnée dans le cadre du modèle de Desclés (1989)⁴ où Sakagami (1997) a analysé le fonctionnement de plusieurs connecteurs temporels en français (« *quand* », « *chaque fois que* », « *après que* », ...). Elle a identifié des relations invariantes exprimées par les connecteurs et a exhibé une cinquantaine de relations possibles entre valeurs aspectuelles, certaines de ces relations se retrouvant dans des phrases contenant des connecteurs différents⁵. Elle a alors dégagé trois caractéristiques communes à ces relations, ce qui témoigne encore selon nous du caractère incomplètement spécifié des relations entre intervalles : (i) une valeur aspectuelle quelconque suivie d'une valeur aspectuelle quelconque dans la chaîne linéaire exprime toujours une certaine succession ; (ii) la fin d'une valeur en précédant une autre est difficilement repérable ; (iii) parmi les 13 relations possibles entre deux intervalles quelconques, seulement 12 sont attestées linguistiquement, celle de chevauchement n'ayant pas été identifiée. L'application sur le texte (6), issu d'un corpus de constats d'accidents, des règles d'ordonnement temporel proposées dans le cadre de ces travaux, conduit aux relations temporelles de la figure 1. On voit clairement que ces dernières sont incomplètement spécifiées (plusieurs lectures temporelles du texte sont possibles) et souvent issues de calculs envisagés au niveau local seulement, c'est-à-dire en ne considérant que les propositions deux à deux. Les opérateurs des S-langages permettent de résoudre les problèmes posés dans le cadre de ces deux remarques.

³ Et si l'on admet que l'on associe des intervalles aux situations dénotées par les propositions.

⁴ Le modèle de Desclés (1989) propose une tripartition des valeurs aspectuelles en état, événement et processus (lesquelles peuvent être ensuite spécifiées plus finement) interprétée dans le modèle topologique : à un état correspond un intervalle ouvert, à un événement un intervalle fermé, et à un processus un intervalle fermé à gauche et ouvert à droite.

⁵ Les régularités observées à ce niveau ont directement été exploitées pour exhiber des relations temporelles pouvant être établies par défaut (Battistelli 2000) *i.e.* en l'absence de connecteurs temporels. Elles indiquent que les valeurs aspectuelles codent à elles seules certaines instructions relatives à l'organisation temporelle dans un texte narratif.

(6) *Etant arrêté momentanément sur la file droite du Bd des Italiens (P1 : état), j'avais mis mon clignotant (P2 : état), j'étais à l'arrêt (P3 : état). Le véhicule B arrivant sur ma gauche (P4 : processus) m'a serré de trop près (P5 : évènement) et m'a abîmé tout le côté avant gauche (P6 : évènement).*

<p>0. Règle générale : si un état est dans une relation avec un évènement ou un processus, alors tous les états concomitants avec cet état héritent de la même relation</p> <p>00. Règle générale : tout état englobe l'ensemble des événements et des processus</p> <p>1. P1, P2, P3 CONCOMITANTS (*existence d'un sous-intervalle commun aux trois propositions*)</p> <p>2. P5 INCLUS_DANS P4 OU P5 SUCCEDE_A P4</p> <p>3. P6 COMMENCE_EN_MEME_TEMPS_OU_APRES P5</p>
--

Figure 1 : Règles générales du modèle et relations temporelles pour le texte (6)

Plus généralement, quelque soient les objets primitifs adoptés (points, intervalles, ...) dans un modèle linguistique donné, il s'agit en fait de situer des points isolés ou des bornes d'intervalles les uns par rapport aux autres. Or situer relativement deux points sur une droite, c'est les mettre soit en relation de précédence soit en relation de simultanéité. Le modèle mathématique qui synthétise ces deux relations est celui des préordres linéaires. L'ensemble de toutes les situations possibles entre n chaînes de points équivaut à rechercher l'ensemble de tous les pré-ordres linéaires constructibles en préservant ces n chaînes. Les préordres s'expriment simplement en termes de *S-langages* (Schwer 2002). La figure 2 exprime les relations temporelles de la figure 1 en termes de S-langages.

3 Temporalité et S-Langages

Ce qui distingue les S-langages des langages formels usuels est la nature des lettres utilisées : pour les S-langages, ce sont des ensembles de lettres – les S-lettres – et non des lettres simples. Une S-lettre possédant une unique lettre peut être confondue avec cette lettre ; ainsi, les mots et les langages deviennent des S-mots et des S-langages. Nous associons à chaque extension temporelle d'un procès A une identité a , pris comme une lettre de l'alphabet, et le mot a^p où p est le nombre des extrémités de ses intervalles et/ou de ses points isolés. Par exemple, si deux procès A et B , d'identité respective a et b , sont des intervalles, alors les (S-)mots aa et bb les représentent sur l'alphabet $A=\{a,b\}$. L'ensemble de toutes les relations entre aa et bb est représenté par le S-langage $aa \otimes bb = \{aabb, a\{a,b\}b, abab, \{a,b\}ab, ab\{a,b\}, abba, \{a,b\}\{a,b\}, baab, ba\{a,b\}, \{a,b\}ba, baba, b\{a,b\}a, bbaa\}$. Le S-mot $\{a,b\}ab$ traduit le fait que le procès A débute en même temps que le procès B mais se termine avant. Toute relation temporelle est un sous S-langage de $aa \otimes bb$. Par exemple, le S-langage $\{abab, \{a,b\}ab, ab\{a,b\}, abba, \{a,b\}\{a,b\}, baab, ba\{a,b\}, \{a,b\}ba, baba\}$ représente $A0B$. En dehors des opérateurs classiques des langages formels (produit, union, intersection), nous utilisons la S-projection et la jointure (Schwer 2001). Pour faciliter le calcul et en réduire la complexité (ce genre de calcul est NP-complet), nous introduisons la notion de S-langage élémentaire.

<p>0., 00. et 1. donnent : $[p1 \otimes p2 \otimes p3] [p4p4 \otimes p5p5 \otimes p6p6] [p1 \otimes p2 \otimes p3]$</p> <p>En raffinant avec 2. : $[p1 \otimes p2 \otimes p3] [\{p4p5p5p4, p4p4p5p5\} \otimes p6p6] [p1 \otimes p2 \otimes p3]$</p> <p>En raffinant avec 3. : $[p1 \otimes p2 \otimes p3] \{p4p5[p5p4 \otimes p6p6], p4p4p5[p5 \otimes p6p6], p4\{p5,p6\}[p5p4 \otimes p6], p4p4\{p5,p6\}[p5 \otimes p6]\} [p1 \otimes p2 \otimes p3]$</p>

Figure 2 : S-expressions des contraintes temporelles du texte (6)

Définitions :

- Soit X un alphabet dont les éléments sont appelés lettres. Le S -alphabet \hat{X} est l'ensemble des parties non vides de X . Par identification des S -lettres $\{x\}$ avec la lettre x , $X \subseteq \hat{X}$. Un S -mot est une suite finie de S -lettres et un S -langage un ensemble de S -mots. L'ensemble de tous les S -mots sur X est noté \hat{X}^* .
- La S -projection sur le sous-alphabet Y d'un S -mot f , notée $f_{|Y}$, est le S -mot obtenu en supprimant de f toutes les occurrences de lettres qui ne font pas partie de Y , puis toutes les S -lettres dont toutes les lettres ont été supprimées. Soit X un alphabet, notons pour $f \in \hat{X}^*$, $X_{|f}$ le sous-alphabet constitué uniquement des lettres composant f .
- Soit X et Y deux alphabets, $f \in \hat{X}^*$, $g \in \hat{Y}^*$, la *jointure* de f et g , notée $f \otimes g$, existe si sur l'alphabet commun $Z = X_{|f} \cap X_{|g}$ les projections des mots f et g coïncident ($f_{|Z} = g_{|Z}$). Dans ce cas, $f \otimes g$ est l'ensemble des S -mots écrits sur X et Y , dont la projection sur X vaut f et la projection sur Y vaut g . Formellement $f \otimes g = \{h \in \hat{T}^*, (T = X \cup Y), (h_{|X} = f) \text{ et } (h_{|Y} = g)\}$.⁶ La jointure de deux S -langages M et N est le S -langage union des jointures d'un S -mot de M et d'un S -mot de N . La jointure est une opération associative, commutative sur les S -langages d'alphabet X , avec pour élément neutre \hat{X}^* . Le traitement informatique du raisonnement temporel est fondé sur le fait qu'il existe une expression explicite (et régulière donc implantable) de $f \otimes g$ (Schwer 2001).
- Un S -mot est un S -langage élémentaire. Un S -langage élémentaire s'exprime uniquement à l'aide de mélanges et de produits de S -langages élémentaires. Il est trivial de montrer que tout S -langage peut s'exprimer comme une union de S -langages élémentaires (un S -langage pouvant toujours s'exprimer comme un ensemble de S -mots). Par exemple, le S -langage $A0B$ est élémentaire et s'exprime par $[a \otimes b][a \otimes b]$.

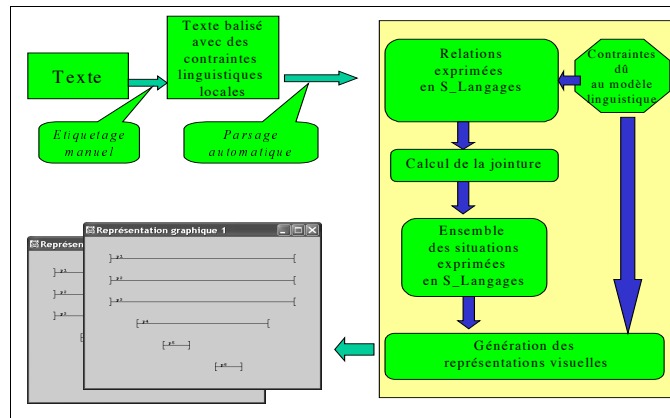


Figure 3 : Diagramme de flux de traitements

Un prototype, développé en Java, implémente ce calcul aspecto-temporel. A partir d'une analyse manuelle aspecto-temporelle réalisée dans le cadre du modèle de (Desclés 1989), le

⁶ Exemple (nous soulignons les lettres communes sur lesquelles la jointure se fait): $\{a,c\} \{a,b\} \{c,d\} \{a,b,c\} \otimes \{a,x,y\} \{a,b\} \{a,b,y\} = \{\{a,c,x,y\} \{a,b\} \{c,d\} \{a,b,c,y\}, \{a,c,x,y\} \{a,b\} \{c,d\} \{a,b,c,y\}, \{a,c,x,y\} \{a,b\} \{c,d\} \{a,b,c,y\}\} = \{a,c,x,y\} \{a,b\} \{c,d\} \otimes \{a,b,c,y\}$.

texte est découpé en propositions auxquelles sont attachées des valeurs aspectuelles. En sortie, le prototype affiche le résultat du calcul sous la forme d'une ou plusieurs représentations d'agencements temporels (cf. figure 3).

4 Conclusion

Dans cet article, nous avons présenté le formalisme des S-langages pour représenter les informations temporelles fournies par une analyse linguistique et réaliser le calcul global de l'ordonnement temporel des procès dans un texte. Cette approche oblige à expliciter les connaissances implicites qui sont mises en œuvre pour établir des relations temporelles entre propositions dans un texte narratif. Nous avons ainsi dégagé des S-langages élémentaires particuliers comme celui traduisant la relation de recouvrement de Kamp et qui sont des parties convexes du treillis des relations temporelles atomiques (Autebert et al. 2003), donc facilement exprimables et calculables. La démarche adoptée consiste en effet à calculer une expression pertinente de relations possibles à partir des marques linguistiques, plutôt qu'à calculer une seule relation possible ou l'ensemble de toutes les relations possibles entre propositions dans un texte.

Références

- AUTEBERT J.-M., SCHWER S.R. (2003), On generalized Delannoy Paths, *SIAM Journal on Discrete Mathematics*, 16 (2), pp. 208-223.
- BATTISTELLI D. (2000), *Passer du texte à une séquence d'images : analyse spatio-temporelle de textes, modélisation et réalisation informatique (système SPAT)*, Thèse de Doctorat, Univ. Paris-IV.
- BRAS M., ASHER N. (1994), Le raisonnement non monotone dans la construction de la structure temporelle de textes en français, *Actes RFIA '94*, Paris.
- DESCLÉS J.-P. (1989), State, Event, Process and Topology, *General Linguistics*, 29(3), Pennsylvania State University Press, University Park and London, pp. 159-200.
- DESCLES J.-P., MAIRE-REPERT D., OH H.-G., BERRI J. (1994), Traitement automatique du temps et de l'aspect pour l'analyse de textes, *T.A.L.*, vol. 35, n°1, pp. 83-106.
- GOSSELIN L. (1996), *Sémantique de la temporalité en français : un modèle calculatoire et cognitif du temps et de l'aspect*, Coll. Champs Linguistiques, Duculot.
- KAMP H. (1981), Evènements, représentations discursives et référence temporelle, *Langage*, n°64, Larousse, pp. 39-64.
- LYONS G. (1990), *Sémantique linguistique*, traduction de J. Durand et D. Boulonnais, Larousse.
- MANN W. C., THOMPSON S. A. (1988), Rhetorical Structure Theory : Toward a functional theory of text organization, *Text*, 8(3) p. 243-281.
- SAKAGAMI R. (1997), *Etude du fonctionnement de quelques connecteurs temporels en français en vue d'un traitement informatique*, Thèse de doctorat, Université Paris-IV.
- SCHWER R.S., (2002-2001), S-arrangements avec répétitions, *Compte. Rendu Acad. Sciences I 334*, pp. 261-266, Paris et Rapport Interne LIPN 2001-2.
- VET C. (1994), Relations temporelles et progression thématique, *Études Cognitives I, Sémantique des Catégories de l'aspect et du Temps*, Acad. des Sciences de Pologne, Warszawa, pp.131-149.