

# Une formalisation plus précise pour une annotation moins confuse: la relation d'Élaboration d'entité

LAURENT PRÉVOT<sup>1,2</sup>, LAURE VIEU<sup>3,4</sup>  
ET NICHOLAS ASHER<sup>3</sup>

<sup>1</sup>CLLE-ERSS, CNRS et Université de Toulouse

<sup>2</sup>LPL, CNRS et Université de Provence

<sup>3</sup>IRIT, CNRS et Université de Toulouse

<sup>4</sup>LOA-ISTC, CNR, Trento

(Received September 2008; revised December 2008)

## ABSTRACT

Les théories cherchant à capturer la cohérence discursive (Mann & Thompson, 1987; Asher & Lascarides, 2003) offrent des cadres théoriques stimulants pour analyser de nombreux phénomènes discursifs. Ils peinent parfois cependant à fournir des définitions de relations combinant précision et robustesse pour des travaux d'annotation à grande échelle permettant, par exemple, de constituer des corpus annotés en relations discursives. Le cas des relations d'Élaboration et d'Arrière-plan constitue une parfaite illustration de ce problème. La plupart des théories discursives les incluent, parfois sous des noms différents, et l'on pourrait penser que leur définition et la manière de les reconnaître est maintenant établie. Cependant, en dépit du fait que ce sont les relations les plus courantes dans les corpus, c'est à leur sujet que les annotateurs ont le plus de problèmes. Elles sont souvent prises l'une pour l'autre. Dans cet article nous examinons la source de ces problèmes et nous proposons une solution basée sur la distinction d'une relation d'Élaboration d'entité, que nous formalisons en SDRT.

## I INTRODUCTION

La cohérence du discours est fondée sur les liens de cohérence qui existent entre ses différentes parties. Les relations de discours constituent par conséquent un élément essentiel pour étudier cette cohérence. Le cas de la relation d'Élaboration est particulièrement intéressant. Premièrement, c'est une relation présente dans la plupart des théories du discours car elle est intuitivement primordiale dans ce contexte. Deuxièmement, les théories qui l'incluent lui font jouer un rôle organisateur central. Par ailleurs, elle reste au centre de nombreuses discussions puisque malgré sa popularité et la reconnaissance de son importance, sa caractérisation peine à atteindre le consensus atteint par certaines autres relations (comme les relations causales). Ce dernier point est corroboré par les difficultés rencontrées pour faire annoter cette relation de manière consistante par des personnes différentes. Dans cet article nous affinons l'étude de certains phénomènes



discursifs liés à cette relation dans le cadre de la sémantique du discours de la SDRT (*Segmented Discourse Representation Theory*, Asher & Lascarides, 2003) afin de contribuer à l'établissement de règles d'identification plus précises qui devraient aider les annotateurs à être plus consistants au sujet de cette relation. Une proposition essentielle pour atteindre cet objectif est l'introduction d'une nouvelle relation de discours, l'*Élaboration d'entité*.

L'article commence (Section 2) par présenter les travaux pertinents réalisés au sujet de la relation d'*Élaboration* tant à un niveau théorique que dans le cadre de campagnes d'annotation de discours. Cette section se termine par l'énumération d'un certain nombre de problèmes concernant cette relation, notamment sa confusion fréquente avec la relation d'*Arrière-plan*, et suggère l'introduction d'une relation de discours supplémentaire pour résoudre ces problèmes. La section 3 présente et définit cette nouvelle relation, l'*Élaboration d'entité*, que nous appliquons à plusieurs exemples problématiques. La section suivante examine des passages et des discours purement descriptifs, un type de discours non traité jusqu'à présent en SDRT, et pour lesquels il a été avancé que la relation d'*Élaboration d'entité* était impliquée. Enfin, la dernière section discute des prolongements à apporter à cette étude.

## 2 UNE CERTAINE CONFUSION À PROPOS DE LA RELATION D'ÉLABORATION

### 2.1 Un peu de contexte sur l'*Élaboration*

La relation d'*Élaboration* est une relation à part dans les théories du discours. La plupart des modèles l'incluent alors que le consensus autour d'une relation et d'un nom de relation est assez rare. Par ailleurs cette relation est l'une des plus fréquentes si l'on en croit les projets d'annotation de relation de discours (Carlson et al., 2003; Wolf & Gibson, 2005; Reese et al., 2007; Prasad et al., 2007). Elle est cependant régulièrement confondue avec la relation d'*Arrière-plan*. Avant de définir précisément les relations de discours qui nous concernent dans les sections suivantes, nous donnons ici une première caractérisation intuitive de ces relations.

La relation d'*Élaboration* est utilisée pour attacher une contribution apportant des précisions sur un élément déjà introduit (ceci correspond d'assez près à la définition introduite par Halliday (1994) – voir aussi la contribution de Georges Kleiber et Hélène Vassiliadou dans ce numéro spécial). La relation d'*Arrière-plan* concerne l'attachement de contributions secondaires ou périphériques: le contexte ou décor de ce qui serait tenu comme l'élément essentiel du discours, comme par exemple la trame narrative.

Les analyses divergentes de l'exemple (1) illustrent la confusion au sujet de ces deux relations.

- (1) [Pourquoi a-t-on abattu Paul Mariani, {cinquante-cinq ans,}<sub>4</sub> {attaché au cabinet de M. François Doubin,}<sub>5</sub> {ministre délégué,}<sub>6</sub> {chargé du commerce



et de l'artisanat}<sub>7</sub> ?<sub>8</sub> [Cet ancien directeur de l'Office HLM de Chelles, {en région parisienne,}<sub>9</sub> était maire {(divers gauche)}<sub>10</sub> du petit village de Soveria,]<sub>11</sub> [situé à 10 kilomètres au nord de Corte,]<sub>12</sub> [sur la route nationale Ajaccio-Bastia.]<sub>13</sub><sup>1</sup> (*Le Monde*)

Les informations additionnelles, les segments 4, 5, 6, 7, 9, 10, 12, 13 de l'exemple (1), peuvent être interprétées soit comme des éléments secondaires se positionnant dans l'arrière-plan discursif ou dans l'arrière-plan de la trame événementielle principale; soit comme des informations additionnelles décrivant des propriétés, et par conséquent élaborant, des objets du discours. Les analyses proposées durant les campagnes et exercices d'annotation reflètent bien cette ambivalence et se partagent largement entre les différentes possibilités: *Élaboration*, *Arrière-plan*, ou les deux. Le projet d'annotation sur l'anglais Discor (Reese *et al.*, 2007) (National Science Foundation Grant IIS-0535154) confirme cette tendance en pointant *Élaboration* et *Arrière-plan* comme les relations qui sont le plus souvent prises l'une pour l'autre. Wolf et Gibson (2005) vont dans le même sens en notant qu'*Élaboration* est la relation pour laquelle l'accord inter-annotateur est le plus faible.

Cette confusion est entretenue par l'absence de marqueurs lexicaux fiables pour les relations d'*Élaboration* et d'*Arrière-plan* (Carlson *et al.*, 2003) alors que de tels indices sont généralement présents et sont les plus faciles à utiliser pour les autres relations. Carlson et ses collègues utilisent les chaînes lexicales et d'autres indices cohésifs pour repérer ces deux relations mais ces indices sont bien moins fiables et plus ambigus que les marqueurs lexicaux. L'absence de marqueur lexical est par ailleurs une des raisons de la mise au ban de cette relation dans le travail de Knott *et al.* (2001).

Une solution pourrait être de ne pas considérer les «petits» segments comme ceux de l'exemple (1) comme des segments de discours et de ne considérer que les segments dans lesquels ils semblent s'insérer, un seul segment par phrase dans cet exemple. Ce serait de fait une approche plus classique de l'interface entre syntaxe et discours. Cependant, selon nous, c'est précisément une telle approche qui génère la confusion dans les tentatives d'annotation à grande échelle comme Carlson *et al.* (2003), Wolf et Gibson (2005), Prasad *et al.* (2007) et Reese *et al.* (2007). En résolvant le problème de l'attachement de ces éléments par leur suppression, on complexifie la question de l'attachement des segments englobants dont la fonction discursive est souvent brouillée. Par conséquent, notre hypothèse est d'une part que l'attachement de ces petits segments devrait être relativement aisé et d'autre part que cette segmentation plus fine rend les segments englobants moins ambigus

<sup>1</sup> Dans les exemples extraits de corpus annotés, les segments discursifs sont marqués par des [crochets] ou des {accolades} dans le cas de segments en incise, et sont indicés par un numéro correspondant à leur position dans le corpus dont ils sont extraits. Ainsi, par exemple, si un segment est indicé 4 cela indique qu'il s'agit du quatrième segment dans le texte du corpus (même s'il s'agit du premier segment reproduit dans cet article). Par ailleurs ces exemples proviennent de corpus variés et en particulier ceux constitués dans le cadre des projets RCFA et ANNODIS.



quant à leur contribution discursive. L'exemple (3), un peu plus bas, illustre ce point. Par ailleurs, une fois établi le besoin de traiter comme segments certains de ces éléments, il est délicat de tracer une ligne claire entre des éléments méritant un traitement discursif et d'autres éléments similaires en surface qui ne le mériteraient pas. Notre stratégie consiste donc à considérer l'attachement discursif de tous ces petits segments (voir la section 3.2 pour leur analyse et leur caractérisation précise).

D'un point de vue plus général, nous considérons que la confortable frontière entre phrase et discours est bien souvent mise à mal. Ce fait a déjà été observé à de multiples reprises, voir par exemple Combettes (1998) pour une description des détachées dans laquelle le rôle discursif de ces dernières est affirmé. De plus, au niveau syntaxique, l'insertion de ces éléments dans l'arbre syntaxique de la principale est délicat. Des propositions, notamment en macro-syntaxe, soulignent les limites de la syntaxe traditionnelle pour décrire ces phénomènes (Deulofeu, 2003). Enfin, la prise en compte de segments non-propositionnels pour analyser le discours a été mise en œuvre à de multiples reprises, notamment dans le cadre de travail de la SDRT. Les introducteurs de cadres, qui sont des modificateurs de phrase (Inflectional Phrase-Adjuncts) et donc se positionnent au niveau discursif (Vieu *et al.*, 2005), en sont une illustration. Le traitement des présuppositions et l'intégration des topiques de discours implicites dans la structure du discours ouvrent également de larges brèches dans une éventuelle correspondance directe entre les phrases ou propositions d'une part et constituants de discours sur lesquels portent les relations de discours d'autre part.

Pour en revenir à la relation d'Élaboration dans la littérature, la RST (*Rhetorical Structure Theory*) (Mann & Thompson, 1987) propose une définition disjonctive qui inclue six conditions: de membre à ensemble (*member/set*), d'instance à abstraction (*instance/abstract*), de partie à tout (*part/whole*), d'étape à processus (*step/process*), d'attribut à objet (*attribute/object*) et de spécialisation à généralisation (*specialisation/generalisation*). Knott et ses collègues (2001) remettent en question le statut de relation de discours pour certaines de ces élaborations et plus précisément pour la condition objet/attribut qui semble être utilisable pour plusieurs segments de l'exemple (1). Cette condition correspond aux cas d'«*Individual-Elaboration*» que Fabricius-Hansen & Behrens (2001) distinguent de l'«*Eventuality-Elaboration*».

La difficulté de cerner précisément la notion d'Élaboration est visible au sein du manuel d'annotation du PDTB (Penn Discourse TreeBank) (Prasad *et al.*, 2007). La classe de relation «*expansion*» correspondrait intuitivement à la relation d'Élaboration mais un examen attentif de ses sous-classes pointe vers une relation beaucoup plus équivoque encore: *Conjunction*, *Instantiation*, *Exception*, *Restatement*, *Alternative*, *List*. La plupart des ces sous-relations ne correspondent pas à l'idée d'Élaboration. Une autre possibilité semble être offerte pour annoter une partie des élaborations: l'annotation de relation basée sur une entité (*Entity-Based coherence*) qui est utilisée quand aucune relation de discours définie dans le manuel ne semble appropriée, mais que les deux segments traitent de la même entité. Cette annotation correspond



parfois aux cas que nous considérons problématiques comme l'exemple (2) entre les segments (a) et (b).<sup>2</sup>

- (2) (a) Pierre Vinken, 61 years old, will join the board as a nonexecutive director Nov. 29. (b) Mr. Vinken is chairman of Elsevier N.V., the Dutch publishing group, ...] (Exemple (4) du manuel *PDTB* (Prasad et al, 2007)).

## 2.2 Élaboration et sa confusion en SDRT

Le cadre théorique qui nous choisissons, qui est décrit avec plus de détails dans la section 2 de l'article de Myriam Bras et Anne Le Draoulec de ce volume, présuppose que les relations de discours soient caractérisées de façon précise. Leur caractérisation est donnée par leur contribution sémantique (proposition logique) et structurelle (nature coordonnante ou subordonnante) au contenu discursif; celle-ci s'accompagne en outre d'indices linguistiques de surface, rassemblés sous forme de règles de déclenchement, qui permettent de reconnaître chaque relation. La SDRT devrait par conséquent être à l'abri des problèmes de confusion évoqués plus haut au sujet des relations d'Élaboration et d'Arrière-plan. Malheureusement au vu des campagnes d'annotation en cours, il n'en est rien. Ce couple de relations continue de faire chuter les taux d'accord entre les annotateurs.

### 2.2.1 La théorie

*Sémantique des relations.* La sémantique des relations de discours est définie de manière plus stricte en SDRT que dans les autres théories discursives. Lorsque la relation d'Élaboration est vérifiée (on a  $\text{Élaboration}(\alpha, \beta)$ ) les *éventualités principales*<sup>3</sup> (respectivement  $e_\alpha$  et  $e_\beta$ ) décrites par les segments attachés (respectivement  $\alpha$  et  $\beta$ ) entretiennent une relation de partie à tout ( $\text{PartieDe}(e_\beta, e_\alpha)$ ). La relation d'Élaboration considérée par la SDRT est donc l'*Eventuality-Elaboration* de Fabricius-Hansen et Behrens (2001). Une des conséquences de cette sémantique est que  $e_\beta$  est incluse spatio-temporellement dans  $e_\alpha$ . Cette contrainte est bien vérifiée entre les segments 11 et 15 de l'exemple (3). De plus la relation de partie à tout impose l'homogénéité entre les entités qu'elle concerne, les éventualités  $e_\alpha$  et  $e_\beta$  doivent donc être soit deux événements, soit deux états.

- (3) [Le juge Gilbert Thiel, {qui sera installé le 7 janvier substitut général à la cour d'appel de Metz},<sub>9</sub> {après treize années passées à l'instruction à Nancy,}<sub>10</sub> a refusé d'accéder à la demande pressante des Japonais.]<sub>11</sub> [Pour son dernier jour de magistrat instructeur,]<sub>12</sub> [vendredi 28 décembre,]<sub>13</sub> [M. Thiel {— qui fut chargé du dossier de Mme Simone Weber et de celui des fausses factures de

<sup>2</sup> Cet exemple emprunté reprend la segmentation faite par les auteurs. Ainsi que nous venons de le discuter, nous distinguerions nous-mêmes deux segments supplémentaires, pour «61years old» et pour «the Dutch publishing group», segments qui sont tous les deux justement des cas d'élaboration qui nous intéressent ici.

<sup>3</sup> Les éventualités, parfois appelées procès, subsument dans l'approche que nous suivons les états et les événements.



Nancy —<sub>14</sub> a rendu une ordonnance de refus de restitution des pâtes de verre]<sub>15</sub> [— trois Gallé et un Daum.]<sub>16</sub> (*Le Monde*)

La relation d'*Arrière-plan* possède une sémantique moins contraignante qu'*Élaboration* puisqu'elle impose seulement que les éventualités principales introduites par les segments qu'elle relie se recouvrent temporellement. A ceci se rajoute cependant que l'éventualité en avant-plan est un événement, alors que celle en arrière-plan est un état. Cette contrainte est bien vérifiée entre d'une part le segment 24 et d'autre part le segment 23 (*Arrière-plan* vers l'avant) ou le segment 25 (*Arrière-plan* vers l'arrière) de l'exemple (4). Nous renvoyons à Vieu et Prévot (2004) et à Asher *et al.* (2007) dans lesquels nous avons proposé un traitement homogène des différents types d'arrière-plans.

- (4) Paul Mariani se trouvait à son domicile le 31 décembre, vers 18 h 30,]<sub>23</sub> [lorsque des gravillons ont été jetés contre la fenêtre.]<sub>24</sub> [Le village était alors plongé dans l'obscurité]<sub>25</sub> [en raison, {croyait-on,}<sub>26</sub> d'une coupure de courant]<sub>27</sub> (*Le Monde*)

La sémantique de ces relations est donc claire: tandis qu'un segment élaborant un autre décrit une partie de l'éventualité de ce dernier, un segment en arrière-plan décrit un état concomitant à un événement décrit dans le segment en avant-plan.

*Règles de déclenchement.* *Élaboration* est inférée (par défaut) sur la base de la relation entre les prédicats utilisés dans les segments. Plus précisément si le prédicat introduisant l'éventualité principale du second segment est potentiellement un sous-type de celui introduit par l'éventualité principale du premier (une structure hiérarchique des prédicats provenant du lexique et des connaissances du monde est ici assumée), alors une relation d'*Élaboration* peut être inférée par défaut (voir pour l'exemple (3), les prédicats «refuser» et «rendre une ordonnance de refus»).

La relation d'*Arrière-plan* est inférée sur des critères complètement différents ayant trait à l'aspect. Une relation d'*Arrière-plan* est inférée (par défaut) si les segments présentent une nature aspectuelle différente. Un état succédant à un événement, ou un événement succédant à un état, sont tenus pour être de bons indices d'*Arrière-plan*. L'état décrit l'arrière-plan de l'événement comme dans l'exemple (4).

Du point de vue des règles de déclenchement, ces relations sont par conséquent assez différentes: *Élaboration ne met pas en jeu de changement aspectuel tandis qu'Arrière-plan le fait.*

### 2.2.2 La pratique

Bien que les relations reposent sur une définition solide qui a autorisé leur utilisation dans un cadre d'analyse théorique, leur utilisation systématique par des annotateurs non-spécialistes se révèle difficile. Elles sont fréquemment prises l'une pour l'autre par ces derniers.

Au niveau théorique, bien que les effets sémantiques de ces relations soient différents, leurs effets temporels ne sont pas mutuellement exclusifs. De ce point de vue temporel, toutes les *Élaborations* pourraient *a priori* être des *Arrière-plans*



puisque la relation de partie implique l'inclusion temporelle, qui à son tour implique le recouvrement temporel. Une différence sémantique claire entre *Élaboration* et *Arrière-plan* repose toutefois sur la nature des éventualités. *Elaboration* requiert des éventualités de même nature tandis que nous considérons qu'*Arrière-plan* impose une alternance entre états et événements.

D'autres difficultés viennent s'ajouter pour distinguer systématiquement ces deux relations lors de l'examen des indices de surface. Tout d'abord, les critères aspectuels sont parfois difficiles à appliquer. La relation entre temps verbal et catégorisation événement / état est parfois indirecte comme par exemple dans le cas de l'imparfait de narration. Par ailleurs de nombreux segments comme 16 dans l'exemple (3) n'ont pas de verbes, donc n'ont pas de valeur aspectuelle claire qui pourrait être utilisée pour déterminer facilement la nature de l'éventualité concernée.

### 2.3 Vers une solution

L'idée force de cet article est qu'étant donné qu'il y a une confusion fréquente entre *Arrière-plan* et *Élaboration* mais que ces deux relations sont relativement bien définies, un certain nombre de cas qui ne sont ni des *Élaborations* ni des *Arrière-plans* canoniques au sens de la SDRT pourraient correspondre à une troisième relation qui reste à introduire dans ce cadre théorique. Cette troisième relation correspond au cas *Object-attribute* d'élaboration identifié dans la RST et examiné dans Knott *et al.* (2001), et à ce que Fabricius-Hansen et Behrens (2001) appellent *Individual-elaboration* ainsi qu'à certains cas d'*Entity-based coherence* dans le PDTB (Penn Discourse TreeBank) (Prasad *et al.*, 2007).

## 3 LA RELATION D'ÉLABORATION D'IDENTITÉ

Dans l'exemple (1), les informations des segments 4 et 5 décrivent des propriétés de l'entité *Paul Mariani* et dénotent des états qui ne peuvent faire partie de l'événement de son assassinat, ni ne caractérisent particulièrement la situation dans laquelle celui-ci s'est produit. Ceci est encore plus clair pour les segments 6 et 7, qui eux décrivent des propriétés de *François Doubin*, totalement exclu de l'assassinat. De même, ni le segment 9, qui décrit la localisation géographique de *Chelles*, ni les segments 12 et 13, qui décrivent celle du village de *Soveria*, ne peuvent dénoter des états qui feraient partie de l'état d'être maire. Par ailleurs, la sémantique d'*Arrière-plan*, ainsi que nous l'avons vu plus haut, interdit d'avoir un arrière-plan sur un état.

Il n'y a donc pour tous ces cas, au sens de la SDRT, ni *Élaboration*, ni *Arrière-plan*. Le rôle de chacun de ces segments n'est pas de décrire une partie de l'éventualité principale du segment antérieur auquel il s'attache, ni la situation dans laquelle cette éventualité a lieu, mais est d'ajouter des informations sur une entité introduite dans ce segment précédent, afin de l'identifier ou de mieux la caractériser. Comme nous venons de le voir, la caractérisation plus précise d'un des participants à une éventualité ne peut être assimilée à l'élaboration de cette éventualité, i.e. à la



description d'une de ses parties, puisqu'une partie d'éventualité est également une éventualité. On retrouve le même phénomène, relativement répandu dans les corpus d'articles de journaux, dans le lien entre les segments 15 et 16 dans (3), ainsi que dans les liens entre (a) et (b) et entre (a) et son apposée dans (2).

La relation rhétorique, que nous appellerons «*Élaboration d'entité*» (en anglais *Entity-Elaboration* ou *E-Elab*), impliquée dans tous ces exemples est nouvelle pour la SDRT. Comme nous l'avons anticipé, cette relation n'est cependant pas une nouveauté dans les études sur le discours. Elle correspond étroitement à la sous-relation de la relation d'Élaboration de la RST nommée «*Object-attribute Elaboration*» et présentée plus haut, que Knott *et al.* (2001) examine principalement. Dans l'étude de Hovy et Maier, (1991), 10 travaux sont recensés comme mentionnant un usage de la relation *Object-attribute*, là aussi considérée comme cas particulier d'*Elaboration*. Enfin, nous avons vu que dans Prasad *et al.* (2007), la relation *Entity-Based coherence* est utilisée en référence à Knott *et al.* (2001).

### 3.1 Sémantique et nature de la relation d'élaboration d'entité

En accord avec l'approche néo-davidsonienne standard (Higginbotham, 2000, Kratzer, 2003, Parsons, 1990), nous considérons que tous les segments, y compris ceux en apposition, comportent la prédication d'une propriété sur au moins une entité. Ces prédications, suivant l'approche néo-davidsonienne, introduisent des éventualités parmi les arguments des propriétés, éventualités qui dans le cas de la SDRT et de la DRT, correspondent soit à un événement (télèque), soit à un état (atélique). Les propositions décrivant des propriétés d'une entité, comme 4, 5, 6, 7, 9, 12 et 13 de (1), 16 de l'exemple (3) ou (b) de l'exemple (2), impliquent en l'occurrence des états.

Ces «états descriptifs» impliqués dans l'*Élaboration d'entité* sont, intuitivement, d'un type particulier: ils décrivent des propriétés caractéristiques d'une entité plutôt que des propriétés transitoires, qui, elles, donnent lieu plutôt à des relations d'*Arrière-plan* (entre autres). Cette distinction intuitive entre propriétés accidentelles ou temporaires (e.g. «être debout, malade, en train de lire») et propriétés plus permanentes (e.g. «être intelligent, avoir les yeux verts»), voire essentielles (e.g. «être un homme») est présente dans la littérature linguistique (cf. Fernald, 2000 pour une synthèse et analyse critique). Elle a été étudiée pour ses effets grammaticaux dans plusieurs langues, et a donné lieu à deux catégories disjointes de prédicats connues sous le nom de prédicats épisodiques (*stage-level predicates* ou SLP) comme 'être debout' et prédicats d'individus (*individual-level predicates* ou ILP) comme 'être intelligent' selon la terminologie de Carlson (1977).

Le corrélat sémantique des manifestations grammaticales serait que les SLP sont les seuls à dénoter des propriétés davidsoniennes, c'est-à-dire ayant pour argument supplémentaire une éventualité (en l'occurrence un état, e.g. *debout(e,jean)*), tandis que les ILP dénotent des propriétés sans argument d'éventualité supplémentaire (e.g. *intelligent(jean)*) (Kratzer, 1995). De ce point de vue, les segments décrivant des



propriétés d'entités impliqués dans la relation d'Élaboration d'entité ne feraient donc pas intervenir d'état du tout, il n'y aurait donc pas d'état descriptif.

Nous n'adopterons pas cette approche pour caractériser sémantiquement les propositions descriptives, car, comme il a souvent été dit, aucune des distinctions ontologiques plausibles pour tenter d'isoler les propriétés dénotées par les prédicats ILP (par exemple leur caractère essentiel, permanent, inhérent, ou encore non-spatial) ne correspond aux distinctions linguistiques observées. Par exemple, la distinction temporaire/permanent n'est pas une distinction ontologique, puisque la plupart des propriétés considérées comme permanentes (e.g. «être grand, être agriculteur»...) le sont relativement à un contexte temporel donné. De fait, des explications alternatives de nature pragmatique ont été proposées (Maienborn, 2004). Nous prendrons nous aussi en compte cette distinction au niveau pragmatique, puisqu'elle donne lieu à des relations rhétoriques différentes (typiquement, *Élaboration d'entité* et *Arrière-plan*).

Au niveau sémantique, nous utiliserons simplement des états distingués comme *états descriptifs d'une entité*. Il s'agit des états caractérisés par un aspect prédicatif statif (verbe statif, copule plus adjectif ou substantif) et dont l'argument correspondant au sujet est l'entité en question.<sup>4</sup> La contrainte sur l'aspect prédicatif élimine les états dus à un aspect grammatical imperfectif mais caractérisés par un prédicat non statif (e.g. *Léa observait Luc* ou *Léa est en train d'écrire un article*). La contrainte sur le sujet permet de considérer que *Léa est la sœur de Luc* et *Léa porte un chapeau* sont deux états descriptifs de Léa et pas de Luc ou du chapeau.<sup>5</sup>

La relation d'élaboration d'entité relie deux segments  $\alpha$  et  $\beta$ , tels que:

- (i)  $\alpha$  dénote une proposition introduisant une éventualité principale  $e_\alpha$  (événement ou état) et un référent de discours accessible  $x$  ( $\exists P \exists \bar{y} (P(e_\alpha, \bar{y}) \wedge x \in \bar{y})$ )<sup>6</sup>;  $e_\alpha$  peut être un état, mais ce n'est pas un état descriptif de  $x$  ( $\neg \text{Etat-Descriptif}(e_\alpha, x)$ ).

<sup>4</sup> Le terme entité est ici utilisé dans un sens non restrictif. Or, tous les exemples d'Élaboration d'entité présentés dans cet article concernent des entités qui ne sont pas des éventualités. On pourrait en fait croire qu'une Élaboration d'entité où l'entité est une éventualité se ramène à une simple Élaboration (au sens de la SDRT), mais les choses sont compliquées. D'une part, les éventualités introduites par un nominal se prêtent à des phénomènes discursifs bien distincts des éventualités principales des constituants introduites par des syntagmes verbaux. D'autre part, les «états descriptifs d'éventualité» (qui restent à examiner de près vu qu'il n'est pas clair du tout que certaines propriétés et relations sur des éventualités, e.g. temporelles ou causales, donnent lieu à des états, même dans une approche néo-davidsonienne) ne peuvent être considérés comme des parties de ces éventualités, spécialement s'il s'agit d'événements. Nous laissons pour le futur les recherches pour établir les frontières entre relations d'Élaboration et d'Élaboration d'entité dans le cas où l'entité est elle-même une éventualité.

<sup>5</sup> Il est probable que la notion de sujet utilisée ici doive être examinée plus en détails pour prendre en compte correctement la structure informationnelle de la phrase et les effets de topique/focus et de figure/fond. Nous laissons ceci pour des approfondissements futurs.

<sup>6</sup> Les formules ne font qu'exprimer dans un langage logique ce qui est expliqué dans les gloses. Leur lecture n'est donc pas requise pour la compréhension de l'article. Par ailleurs



- (ii)  $\beta$  dénote une proposition dont l'éventualité principale est un état  $e_\beta$  descriptif de  $x$  (Etat-Descriptif( $e_\beta, x$ ))
- (iii)  $e_\beta$  et  $e_\alpha$  se recouvrent temporellement ( $O_t(e_\alpha, e_\beta)$ ).<sup>7</sup>

où  $e$  est un état descriptif de  $x$  (Etat-Descriptif( $e, x$ )) si et seulement si  $e$  est un état caractérisé par un prédicat statif dont  $x$  est l'argument sujet.

Les effets sémantiques de cette relation sont donc, formellement :

$$E - \text{Elab}(\alpha, \beta) \rightarrow \exists x \exists P \exists \vec{y} (P(e_\alpha, \vec{y}) \wedge x \in \vec{y} \wedge \neg \text{Etat} - \text{Descriptif}(e_\alpha, x) \wedge \text{Etat} - \text{Descriptif}(e_\beta, x) \wedge O_t(e_\alpha, e_\beta))$$

avec

$$\text{Etat-Descriptif}(e, x) \leftrightarrow \text{Etat}(e) \wedge \exists P \exists \vec{y} (P(e, x, \vec{y}) \wedge \text{Statif}(P))$$

en supposant que l'argument sujet est le premier des arguments de  $P$  après celui de l'éventualité.

Nous remarquons que cette sémantique et celles d'*Élaboration* et d'*Arrière-plan* ne sont pas incompatibles de façon flagrante. Supposons que nous ayons  $E - \text{Elab}(\alpha, \beta)$ . Dans le cas où  $e_\alpha$  est un état, il ne peut y avoir  $\text{Arrière-plan}(\alpha, \beta)$  puisqu'on n'a pas d'arrière-plan entre deux états, mais la simple nature des éventualités ne s'oppose pas à une  $\text{Elaboration}(\alpha, \beta)$ . On peut toutefois faire l'hypothèse qu'un état décrivant une entité ne peut être une partie que d'un autre état décrivant cette même entité, ce qui est alors exclu explicitement par notre définition. Dans le cas où  $e_\alpha$  est un événement, on ne peut avoir  $\text{Elaboration}(\alpha, \beta)$  puisqu'un état ne peut être une partie d'un événement, mais on pourrait *a priori* avoir  $\text{Arrière-plan}(\alpha, \beta)$ . Ici, il convient d'observer qu'une contrainte prise plus ou moins en compte dans l'ancien traitement de la relation d'*Arrière-plan* (Asher et Lascarides, 2003) a disparu du traitement actuel (Vieu et Prévot, 2004; Asher *et al.*, 2007). Il s'agit de la contrainte de lien thématique entre les deux éventualités. L'événement et l'état d'un *Arrière-plan* doivent tous les deux avoir un «topique» en commun. En quelque sorte, l'état doit être fonctionnel à l'histoire, ce qui est effectivement le cas de 23 et 25 dans l'exemple (4), mais pas de 4, 5, 6, 7, 9, 12 et 13 dans (1); en fait, le simple partage d'une entité est insuffisant. En supposant que cette contrainte sémantique soit effectivement prise en compte pour *Arrière-plan*, les cas de recouvrement sémantique restants sont en fait limités.

le lecteur peut se reporter à la section 2 de l'article de Myriam Bras et Anne Le Draoulec de ce volume pour plus de détails sur le cadre d'analyse.

<sup>7</sup> Pour ce dernier point, il ne s'agit en l'état actuel que d'une hypothèse théorique, validée seulement sur un corpus limité. Il semble naturel de supposer que  $x$  présente la propriété caractérisant  $e_\beta$  pendant au moins une partie de l'éventualité  $e_\alpha$ . L'annotation de l'*Élaboration d'entité* sur un large corpus nous permettra de vérifier cette hypothèse. Il conviendra alors d'examiner particulièrement les cas où  $e_\alpha$  est un événement, pour lesquels on peut s'attendre à ce que la relation de recouvrement temporel se précise en inclusion temporelle ( $\subseteq_t$ ).



Un cas résiste cependant à cette contrainte supplémentaire dans (1) même, où l'on peut envisager soit que 11 est un *Arrière-plan* de 8, si on considère que le texte suggère que le fait que *PM* soit *maire de Soveria* joue un rôle dans son assassinat, soit que 11 continue l'*Élaboration d'entité* de 8 en 4 et 5, dans le cas contraire. Ces deux attachements sont possibles, car en effet, du point de vue de sa nature structurelle, la relation d'*Élaboration d'entité* est, tout comme *Élaboration* et *Arrière-plan*, subordonnante. Une *Élaboration d'entité* peut être continuée (5 continue 4 et 7 continue 6), et laisse le premier argument libre pour de futurs attachements, par exemple 8 est libre pour l'attachement de 11 (Asher & Vieu, 2005). Dans la suite, nous considérerons toutefois que l'attachement de 11 à 8 dans (1) continue bien l'*Élaboration d'entité* faite en 4 et 5.

### 3.2 Règles de déclenchement

Contrairement à un *arrière-plan*, une *élaboration d'entité* n'est pas déclenchée par un changement aspectuel. Un tel changement peut se produire mais n'est pas nécessaire à l'inférence de cette relation. Elle n'est clairement pas non plus déclenchée par un lien sémantique entre les prédicats des deux segments, contrairement à une *élaboration*. Elle est en fait principalement déclenchée par un changement de *perspective*:  $\beta$  est une description de l'entité  $x$ , et  $\alpha$  l'est pas, même si  $x$  y est mis en jeu.<sup>8</sup> Ce point est pris en compte par l'adoption de la règle de déclenchement suivante:

$$\begin{aligned} & (?(\alpha, \beta) \wedge P(e_\alpha, \vec{y}) \wedge x \in \vec{y} \wedge \neg \text{Indice} - \text{état} - \text{descriptif}(e_\alpha, x) \\ & \wedge \text{Indice} - \text{état} - \text{descriptif}(e_\beta, x)) > E - \text{Elab}(\alpha, \beta) \end{aligned}$$

Cette règle spécifie que si l'on cherche à attacher  $\beta$  à  $\alpha$  (le '?' dans la formule précise que l'on ne sait pas encore par quelle relation cet attachement va se faire), que  $x$  est un des arguments du prédicat correspondant à l'événement principal de  $\alpha$ , qu'il n'y a pas d'indice qu' $\alpha$  introduit un état descriptif de  $x$  mais que par contre il y a des indices que  $\beta$  en introduit un, alors par défaut (l'inférence par défaut est notée '>') la relation entre  $\alpha$  et  $\beta$  est une *élaboration d'entité*.

Il reste bien évidemment à établir dans quelles circonstances nous avons des *indices d'état descriptif d'une entité*. Nous venons de voir que l'utilisation de la distinction ILP/SLP pour la caractérisation sémantique des états descriptifs est difficile. En pratique, elle est aussi assez délicate à exploiter comme indice lexical de description d'entité pour déclencher une *Élaboration d'entité*. En effet, la distinction ILP/SLP est controversée même au niveau grammatical, au vu de la facilité de construire des exceptions à toute catégorisation lexicale. Ceci amène Fernald (2000) à considérer

<sup>8</sup> Le segment  $\alpha$  peut aussi être une description d'entité, mais d'une autre entité que  $x$ . Ce cas se produit dans l'attachement de 6 à 5 dans (1), puisque 6 décrit *François Doubin*, alors que 5 décrit *Paul Mariani*. Voir la section 3.3 sur ce point. Les cas où  $\alpha$  est une description de  $x$  seront traités dans la section 4.



que la catégorisation grammaticale ne doit être faite au niveau lexical que par défaut, et qu'elle n'est vraiment opératoire qu'au niveau de la proposition, ce qui implique une analyse complexe.

Toutefois, nombre de cas d'*Élaboration d'entité* sont repérables de façon plus simple grâce à la structure syntaxique. En effet, en particulier dans le style concis des dépêches d'agence, nombre de segments élaborant une entité apparaissent sous la forme d'appositions ou de relatives attachées au NP décrivant l'entité élaborée.

Les appositions, comme dans 4, 5, 6, 7, 9, 10, 12 et 13 de (1), dans (a) de (2) et dans 16 de (3), sont des éléments *incidents* (Marandin, 1998) ou *parenthétiques* (Bonami & Godard, 2007), c'est-à-dire des structures syntaxiques indépendantes insérées dans une autre structure syntaxique, l'hôte. Les parenthétiques ont fait l'objet de travaux assez précis à la croisée entre syntaxe, sémantique et pragmatique. Dès Asher (2000), un traitement des parenthétiques a été proposé dans un cadre discursif. Néanmoins, ce travail ne couvre pas les types d'incidentes qui nous intéressent ici. Récemment, Potts (2005) a analysé les suppléments (appositions, «*as*-clauses», relatives non-restrictives et certains adverbes) comme des implicatures conversationnelles, alors que Beaver (2006) y oppose une analyse en termes présuppositionnels (voir la section 5 pour quelques éléments supplémentaires à ce sujet). Pour Beaver, l'application des tests présuppositionnels à ces éléments (comme le test de projection) montre un comportement similaire aux présuppositions classiques. Ces travaux se focalisent cependant sur l'interface sémantique/pragmatique et proposent plus un cadre global pour l'analyse de tous ces phénomènes qu'un traitement ou une description détaillée qui nous serait plus utile ici pour caractériser les différents types d'incidentes que nous considérons. Les constructions détachées de Combettes (1998) correspondent plus étroitement à notre sujet mais cette étude qui met en avant la fonction discursive de ces constructions mériterait elle aussi d'être systématisée.

Les appositions, contrairement aux incises (cf. 26 dans l'exemple (4)), ne forment pas une proposition complète. Ce sont des groupes nominaux, des adjectifs ou des participes. D'après Marandin (1998), les appositions se distinguent en incidents prédicatifs à la phrase, les *appositions qualitatives*, et en incident prédicatifs au NP, les *appositions ordinaires* comme illustré dans les exemples (5) à (7) ci-dessous. Cette distinction est importante pour nous car les appositions qui donnent lieu à une *Élaboration d'entité* ont toutes pour hôte un NP, alors que celles qui donnent lieu à un *Arrière-plan*, voire une *Explication*, un *Commentaire*, ou une autre relation rhétorique encore, ont pour hôte la phrase. Marandin (1998), suivant Milner (1978), explore nombre de phénomènes syntaxiques pour établir la différence entre hôtes; nous reprenons ici le test du déplacement. L'apposition incidente au NP ne peut se déplacer hors du NP ailleurs dans la phrase, ce qui est possible pour les appositions incidentes à la phrase. Ceci distingue bien, notamment les cas suivants:

- (5) Léa, malade, est restée à la maison aujourd'hui. / Malade, Léa est restée à la maison aujourd'hui. / Léa est restée à la maison aujourd'hui, malade. (Exemple forgé)



- (6) Marie, toute agitée de soupçons, demanda à rencontrer le commissaire. / Toute agitée de soupçons, Marie demanda à rencontrer le commissaire. / Marie demanda à rencontrer le commissaire, toute agitée de soupçons. (Marandin, 1998)
- (7) Léa, maire de Teilhet, est restée à la maison aujourd'hui. / \*Maire de Teilhet, Léa est restée à la maison aujourd'hui. / \*Léa est restée à la maison aujourd'hui, maire de Teilhet. (Exemple forgé)

Dans (5) et (6), le déplacement est possible et marque une apposition qualitative qui est soit une *Explication* (interprétation préférée pour les deux premières variantes), soit un *Arrière-plan* (préférée pour la troisième). Dans (7), nous avons par contre une *Élaboration d'entité* dans le cadre d'une apposition ordinaire, ce qui est démontré par l'impossibilité de la troisième variante.

On ne peut cependant pas considérer que toutes les appositions ordinaires déclenchent des élaborations d'entités, car par exemple, nombre de prédicats au participe passé caractérisent en fait un événement plutôt qu'un état descriptif, comme pour 3 et 4 dans (8), ce qui se démontre par l'usage des adverbes de localisation temporelle:

- (8) [Les pâtes de verre historiques du Musée de l'Ecole de Nancy, {volées en 1985}<sub>3</sub> {et récupérées en juin 1990 au Japon par la police française}<sub>4</sub>, vont-elles repartir dans ce pays ?]<sub>5</sub> (*Le Monde*)

Dans ces cas, on est en présence d'une relation de *flashback*, en quelque sorte une *narration* inverse. Ces cas peuvent être filtrés simplement en testant la nature de l'éventualité du segment en apposition, gardant les états et éliminant les événements.

Nous obtenons donc une règle sélectionnant les cas d'apposition au NP comme des indices privilégiés d'états descriptifs:

(Apposition – NP( $\alpha$ ,  $\beta$ )  $\wedge$  Etat( $e_\beta$ )  $\wedge$  P( $e_\beta$ , x,  $\vec{y}$ ))  $\rightarrow$  Indice – état – descriptif( $e_\beta$ , x)

En guise de conclusion sur le cas des appositions déclenchant ou non une *Élaboration d'entité*, nous observons que les appositions construites avec des participes présents (non adjectivés) ne permettent pas non plus d'introduire des élaborations d'entités, car l'apposition est alors systématiquement une apposition incidente à la phrase, comme on peut le voir sur l'exemple suivant.

- (9) Marie, étant la cousine de Léa, a été invitée à la fête. / Marie a été invitée à la fête, étant la cousine de Léa. (Exemple forgé)
- (10) Marie, la cousine de Léa, a été invitée à la fête. (Exemple forgé)

Avec (9), mais pas (10), la Marie dont on parle est déjà clairement identifiée et il est très possible que le fait qu'elle soit la cousine de Léa fasse partie des connaissances partagées du locuteur et de son interlocuteur. Le rôle de l'apposition dans (9) n'est donc pas d'identifier ou de caractériser Marie, mais de présenter le fait que Marie soit la cousine de Léa comme la raison pour laquelle elle a été invitée à la fête.

On voit ici confirmation que la distinction ILP/SLP ne peut être utilisée en tant qu'indice puisqu'un même prédicat, clairement ILP comme «être la cousine



de Léa», peut donner lieu soit à une *Élaboration d'entité*, soit à une *Explication* (ou une autre relation), selon qu'il est utilisé en prédication simple ou bien en participe présent. L'emballage informationnel, ici à l'interface entre syntaxe et structure discursive, prend donc le pas sur la sémantique lexicale.

L'usage d'une apposition n'est pas le seul moyen d'exprimer une relation d'Élaboration d'entité. Nous avons vu que les propositions indépendantes peuvent être également utilisées, comme avec (b) dans l'exemple (2) voire 11 dans (1) ainsi que discuté plus haut. Une autre famille assez typique d'Élaboration d'entité est constituée par les relatives, comme 14 dans (11):

- (11) [Jean Baudesson, {maire échevin de Saint-Dizier},<sub>12</sub> {autorisé par lettres patentes d'Henri IV},<sub>13</sub> installa à Marnaval {— qui signifiait val ou vallée de la Marne ou bien en aval de la Marne—}<sub>14</sub>, une forge]<sub>15</sub> [qui connut son apogée au XIXe siècle.]<sub>16</sub> (*L'Est Républicain*)

Cependant toutes les relatives ne donnent pas lieu à *Élaboration d'entité*. Il convient tout d'abord de noter que les relatives restrictives (ou déterminatives) ne nous concernent pas,<sup>9</sup> puisqu'elles ne donnent en fait pas lieu à un segment séparé, comme dans *Les Américains qui mangent des hamburgers sont gros*. Seules les relatives dites explicatives (ou appositives) comme dans *Les Américains, qui mangent des hamburgers, sont gros* déterminent des segments.

Mais toutes les relatives explicatives ne sont pas des cas d'Élaboration d'entité puisqu'elles ne caractérisent pas toutes des états descriptifs d'entités. Ce sont assez souvent des événements comme pour 9 (événement futur) et 14 (événement passé) dans l'exemple (3) ainsi que 16 dans (11). Il faudrait donc, comme précédemment, restreindre l'éventuelle règle recueillant de nouveaux indices aux états. Malheureusement, cela ne suffirait pas. En effet, dans ce cas la syntaxe ne nous permet pas de distinguer les états descriptifs, qui introduisent une *Élaboration d'entité*, des autres états, qui donnent lieu à des *Arrière-plans* ou des *Explications* comme dans *Son cocher, qui était ivre, s'assoupit tout à coup*, voire d'autres relations encore. Et bien entendu, dans le cas de propositions indépendantes complètes la syntaxe ne peut être d'aucun secours.

Nous laissons donc pour des travaux futurs l'examen plus précis de ces deux familles de cas qui devra déterminer quels sont les moyens mis en œuvre pour signaler et reconnaître la relation d'Élaboration d'entité.

### 3.3 Rôle structurel et topiques d'entités

En ce qui concerne le rôle de la relation d'Élaboration d'entité dans la structure globale du discours, nous avons vu plus haut que cette relation est subordonnante, ce qui correspond bien avec le point de vue pris dans d'autres théories pour les

<sup>9</sup> Nous excluons de la même façon les relatives attributives (*J'ai les cheveux qui grisonnent*) et les relatives substantives (*Qui m'aime me suive*).



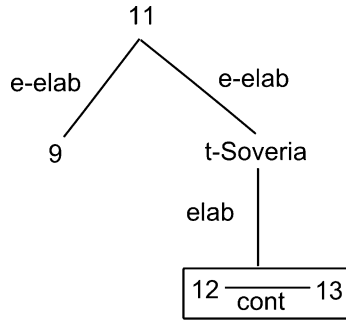


Figure 1.

relations correspondantes, notamment en RST où *Object-attribute elaboration* est une relation de type noyau-satellite.

En SDRT, une continuation d'une relation subordonnante – par exemple, pour la relation d'Élaboration, un cas  $Elab(a,b)$  et  $Elab(a,c)$  – donne lieu à la création d'un segment complexe regroupant les différents sous-segments ayant la même fonction rhétorique, relié directement au point d'attachement – ici  $b$  et  $c$  forment le segment complexe  $[b,c]$  attaché directement à  $a$ , ce que l'on notera en fait par  $Elab(a,[b,c])$  et *Continuation*( $b,c$ ).

La relation d'Élaboration d'entité n'échappe pas à ce principe, et de fait, pour la première phrase de (1) on obtient a priori la structure suivante :  $E-Elab(8,[4,5])$  et  $E-Elab(5,[6,7])$ . Mais ce principe appliqué à la seconde phrase nous donne  $E-Elab(11,[9,12,13])$  alors que 9 est une élaboration de *Chelles* et 12 et 13 une élaboration de *Soveria* et que donc le lien de continuation entre 9 et 12 est inexistant contrairement à celui entre 12 et 13.

Il semble nécessaire de prendre en compte cette différence et de garder par conséquent trace du lien entre les différents sous-segments, c'est-à-dire l'entité dont il est question dans l'Élaboration d'entité. Nous proposons donc, d'une part, de bloquer le regroupement des élaborations d'entités qui ne traitent pas de la même entité, ce qui revient à contraindre sémantiquement la relation de *Continuation* dans les cas de continuation d'Élaboration d'entité (nous reviendrons sur ce point en section 4). Et d'autre part, dans ces mêmes cas de continuation d'Élaboration d'entité, nous proposons d'introduire un nouveau constituant implicite, un *topique d'entité*, afin de mettre en évidence l'entité élaborée et de regrouper les segments l'élaborant. Ce topique s'insère dans la structure entre le constituant élaboré auquel il s'attache par la relation *Élaboration* et le constituant complexe élaborant auquel il est attaché par la relation *E-Elab*. Dans le cas de la seconde phrase de (1), la continuation entre 12 et 13 nous force à insérer entre le segment 11 et le segment complexe  $[12-13]$  un topique d'entité pour *Soveria*, étiqueté «t-Soveria»; nous aurons donc :  $E-Elab(11,9)$ ,  $E-Elab(11,t-Soveria)$  et  $Elab(t-Soveria,[12,13])$ . Cette structure est illustrée dans la Figure 1.



Les topiques d'entité ont la caractéristique principale des autres topiques utilisés jusqu'à présent en SDRT, comme les topiques narratifs: ils généralisent l'ensemble des sous-segments qui leur sont attachés par la relation d'*Élaboration* (parfois nommée *topique* pour mettre en évidence que le constituant topique est un constituant implicite). C'est du fait de ce rôle généralisateur des topiques qu'il ne nous a pas semblé opportun d'introduire systématiquement un topique d'entité pour toute relation d'*Élaboration d'entité*, y compris donc les cas ne formant pas de segment complexe, comme 9 dans (1).

D'autre part, tout comme les topiques narratifs, les topiques d'entité sont des constituants, ils décrivent donc une éventualité principale. Pour les topiques d'entités, cette éventualité est nécessairement de type «état descriptif» (cf. plus haut). Cet état est caractérisé par un prédicat généralisant ceux qui caractérisent les états descriptifs des constituants élémentaires du segment complexe; à la limite, il correspond simplement à l'état d'existence de l'entité concernée, et le topique d'entité représente alors simplement un acte de référence à l'entité.

La structuration hiérarchique obtenue par l'utilisation de topiques d'entité met en évidence une différence supplémentaire entre les relations d'*Élaboration d'entité* et d'*Arrière-plan*. L'arrière-plan d'un événement peut bien entendu être lui-même complexe, c'est d'ailleurs une des justifications de la nature subordonnante de cette relation. Mais il n'y a ni arrière-plan d'un arrière-plan,<sup>10</sup> ni sous-regroupements à l'intérieur de cet arrière-plan. En d'autres termes, un arrière-plan décrit la situation dans laquelle se déroule l'événement en avant-plan; cette situation peut être complexe et donc décrite par plusieurs propositions, mais cette situation est unique. Ceci est en contraste avec le fait qu'un même segment peut donner lieu à des élaborations d'entités qui soit n'ont aucun rapport entre elles (comme dans la seconde phrase de (1)), soit ont au contraire un rapport hiérarchique, comme dans la première phrase de (1), pour laquelle on obtient au final: *E-Elab(8,t-PaulMariani)*, *Elab(t-PaulMariani,[4,5])*, *E-Elab(5,t-FrançoisDoubin)* et *Elab(t-FrançoisDoubin,[6,7])*. En regroupant l'analyse de cette première phrase et celle de la seconde phrase plus haut, et en retenant l'hypothèse où 11 est une continuation de l'élaboration de l'entité «Paul Mariani» commencée en 4 et 5, on obtient la structure de tout l'exemple (1) sur la Figure 2.

#### 4 PASSAGES OU DISCOURS DESCRIPTIFS

Afin d'éviter certaines confusions pour les annotateurs, il convient de compléter cette étude de la relation d'*Élaboration d'entité* par l'examen de cas où il semblerait à première vue qu'elle est impliquée. Cet examen nous permet également de

<sup>10</sup> Nous avons récemment été amenés à considérer des cas où des *Arrières-plans* sur des éléments déjà situés à l'arrière-plan semblent possibles. L'examen approfondi de ce cas de figure reste à faire et il est trop tôt pour présumer de son traitement et d'une modification de la relation d'*Arrière-plan*.



ancée par Knott

arratif, ont pour  
dans ce cas, une  
t de perspective  
te, on se retrouve  
entité, sans passer  
exemple (12) ainsi

- On la rencontre  
[Réputée et  
caille,] [c'est une  
(Wikipédia)  
are often mass

However this jewel  
nous considérons  
ous sommes là en  
que l'attachement  
le segment 2 de  
ation d'entité mais  
semble-t-il, Mann  
s. Ils considèrent  
re subordonnante  
pliquée les amène



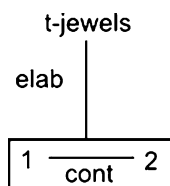


Figure 3.

à conclure, de façon erronée selon nous, que les *Élaborations d'entités* ne peuvent s'intercaler dans des séquences d'autres relations rhétoriques. Or, l'usage abondant des appositions dans les faits divers des journaux montre que cette hypothèse n'est pas valable. Elle n'est pas valable non plus dans le cas d'*Élaboration d'entité* dans des propositions indépendantes, comme le montre l'exemple (14), où une *Élaboration d'entité* intervient en 6–8, s'intercalant dans un *Arrière-plan* (9 est attaché à 5 par *Arrière-plan*).

- (14) [...] [Je m'apprêtais à en faire autant pour le bas de mon pantalon]<sub>3</sub> [(je suis extrêmement soigneux de mes effets),]<sub>4</sub> [quand apparut mon ami Axelsen.]<sub>5</sub> [Mon ami Axelsen est un jeune peintre norvégien,]<sub>6</sub> [plein de talent et de sentimentalité.]<sub>7</sub> [Il a du talent à jeun et de la sentimentalité le reste du temps.]<sub>8</sub> [A ce moment, la sentimentalité dominait.]<sub>9</sub> (*Une mort bizarre*, Alphonse Allais)

Le traitement spécial de la relation d'*Élaboration d'entité* proposé par Knott *et al.* (2001) est donc injustifié selon nous. Nous proposons plutôt dans les cas comme (12) et (13) l'utilisation d'une relation de *Continuation de description*, c'est-à-dire une relation de *Continuation* renforcée, dont la sémantique est celle d'une continuation d'états descriptifs d'une même entité, et qui requiert la construction d'un topique d'entité implicite comme décrit ci-dessus en section 3.3.<sup>11</sup> Nous aurions donc pour l'exemple (13) la structure illustrée par la Figure 3. Il est clair dans cette structure qu'on ne peut pas attacher 3 à 1, qui n'est pas sur la *frontière droite*.

## 5 DISCUSSION

Dans cet article, nous avons approfondi l'étude de la relation discursive d'*Élaboration*. Dans le cadre de cette étude nous avons été amenés à considérer plus particulièrement les élaborations d'entités pour lesquelles nous avons proposé une nouvelle relation de discours. Cette proposition nous a ensuite permis de traiter un type de discours peu étudié jusqu'à présent en SDRT, les passages descriptifs. Avoir clarifié la distinction entre *Élaboration* et *Élaboration d'entité*, et

<sup>11</sup> Il conviendrait alors d'utiliser également ci-dessus une autre relation que la simple relation de *Continuation* qui a une sémantique vide. Nous laissons l'étude de cette nouvelle relation pour un autre article.



avoir mis ces relations en perspective par rapport à la relation d'*Arrière-plan* devrait permettre une annotation plus aisée et en conséquence augmenter l'accord entre annotateurs. Nous démarrons actuellement l'expérimentation visant à confirmer cette hypothèse.

Du point de vue théorique, une piste à suivre pour poursuivre ces recherches est de clarifier la relation entre les segments qui sont reliés par *Élaboration d'entité* et les présuppositions. Tout comme le contenu des présuppositions, le contenu des segments en apposition et des relatives non restrictives ont la propriété de projection, c'est-à-dire que ces contenus (et donc nombre de segments reliés par *Élaboration d'entité*) échappent à la portée de la négation, du conditionnel, ou de l'interrogative. Le contenu du type de segments dépendants que nous avons examiné ici a été assimilé aux implicatures conventionnelles (Potts, 2005). Beaver (2006) souligne que les implicatures conventionnelles sont très similaires aux présuppositions car elles satisfont tous les tests détectant les présuppositions, en particulier les tests de projection. Il y a certes des différences entre les déclencheurs de présupposition traditionnels comme les descriptions définies et le contenu des segments en apposition ou des relatives non-restrictives. Néanmoins, le comportement similaire des présuppositions, des appositions et relatives non restrictives vis à vis de la projection confirme, d'une part, si besoin était le bien-fondé d'avoir considéré ces dernières comme introduisant des segments, et nous semble, d'autre part, être une piste prometteuse de recherche pour généraliser nos observations. De plus, il apparaît que la relation d'*Élaboration d'entité* pourrait contribuer à l'analyse de certaines présuppositions. En SDRT, le traitement des présuppositions est réalisé à l'aide des relations discursives de *Conséquence* ou d'*Arrière-plan*. Bien que ce traitement soit en général satisfaisant, il existe des présuppositions qui apportent de nouvelles informations tout en étant liées au contexte discursif. C'est le cas des présuppositions introduites par des descriptions définies comme dans l'apposition de l'exemple (15).

- (15) Mikhaïl Saakachvili, le jeune et bouillant président géorgien, avait besoin d'action pour sauver son régime. (*Libération*)

La description définie '*le jeune et bouillant président géorgien*' génère une présupposition «il y a un unique jeune et bouillant président géorgien», mais cette personne est clairement identifiée avec '*Mikhaïl Saakachvili*'. C'est un type de liage classique dans l'approche de Van der Sandt (1992). Cependant cette présupposition apporte également des informations nouvelles au contexte. Elle ne peut donc pas constituer une *Conséquence* (même défaisable) du contexte discursif. La présupposition ne peut par conséquent pas être liée dans le modèle d'Asher et Lascarides (1998, 2003). En effet, ce n'est pas un *Arrière-plan* non plus car l'individu décrit par la présupposition est identifié avec '*Mikhaïl Saakachvili*'. Le lien entre ces deux constituants n'est pas de nature temporelle ou événementielle, c'est plutôt un lien d'identité entre objets. Ce lien peut être saisi par la relation d'*Élaboration d'entité* introduite dans cet article. Cette nouvelle relation permettrait donc à la



SDRT de traiter certains exemples classiques de liage de Van der Sandt de façon plus satisfaisante que ce qui est proposé dans Asher et Lascarides (1998, 2003).

Par ailleurs, nous n'avons que rapidement évoqué les travaux d'Alistair Knott et ses collègues (2001) qui préconisent de restreindre l'usage des relations de discours à la modélisation de la cohérence locale, et d'opter pour une structure globale basée sur les continuations et changements de topiques d'entité, et sur les règles de gestion de ces derniers. Ce traitement remet en question l'hypothèse de la frontière droite du discours, puisque les changements de topique d'entité (en d'autres termes, les relations d'Élaboration d'entité) ne seraient pas gouvernés par cette règle qui serait confinée à la cohérence locale. Un tel traitement ne nous semble *a priori* adéquat que pour le genre descriptif (consacré à la description d'un objet, d'un paysage, etc.). Dans les textes narratifs ou argumentatifs, comme par exemple les dépêches d'agence, la structure globale est souvent basée à l'inverse sur des relations rhétoriques, et ce n'est que localement qu'il y a enrichissement informationnel sur certaines entités. Nous prenons néanmoins la position de Knott *et al.* au sérieux et les propositions faites dans notre article devront être mises à l'épreuve plus systématiquement. Les données proposées par Knott et ses collègues sont intéressantes, mais les contextes discursifs proposés sont trop courts et leur acceptabilité est parfois discutable; nous envisageons donc de recueillir plus de données de ce type, notamment au sein du projet d'annotation discursive ANNODIS. Ceci est possible car avoir clarifié ce qu'est la relation d'Élaboration d'entité dans la SDRT nous permet d'annoter plus précisément les textes sans pour autant importer toutes les hypothèses théoriques de la SDRT.

### Remerciements

Nous tenons à remercier tous les membres du projet ANNODIS et en particulier le groupe d'analyse ascendante (Marianne Vergez-Couret, Anne Le Draoulec, Myriam Bras, Anna Berdah, Farah Benamara-Zitoun et Philippe Muller) dont le travail régulier constitue un préalable à la rédaction de cet article; les participants au Projet RCEA pour leurs élaborations interactives et enfin les trois relecteurs de la revue *JFLS* pour leurs relectures attentives et leurs commentaires constructifs. Nous restons bien sûr responsables des problèmes résiduels de cet article.

*Adresses pour correspondance:*

Laurent Prévot

Laboratoire Parole et Langage

Université de Provence

29 av. R. Schuman

13621 AIX-EN-PROVENCE Cedex 1

France

e-mail: laurent.prevot@lpl-aix.fr

Laure Vieu

Istituto di Scienze e Tecnologie della Cognizione



Consiglio Nazionale delle Ricerche  
Via alla Cascata 56C  
Povo  
38100 Trento  
Italie  
e-mail: [vieu@irit.fr](mailto:vieu@irit.fr)

Nicholas Asher  
Institut de Recherche en Informatique de Toulouse  
Université Paul Sabatier  
118, route de Narbonne  
31062 TOULOUSE Cedex 09  
France  
e-mail: [asher@irit.fr](mailto:asher@irit.fr)

#### REFERENCES

- ANNODIS, Annotation Discursive, un corpus de référence annoté discursivement, Projet ANR [2008–2010].
- Asher, N. (2000). Truth conditional discourse semantics for parentheticals. *Journal of Semantics*, 17: 31–50.
- Asher, N. and Lascarides, A. (1998). The semantics and pragmatics of presupposition. *Journal of Semantics*, 15: 239–300.
- Asher, N. et Lascarides, A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.
- Asher, N., Prévot, L. et Vieu, L. (2007). Setting the background in discourse. *Discours* 1: on-line journal, <http://discours.revues.org/index301.html>
- Asher, N. et Vieu, L. (2005). Subordinating and coordinating discourse relations. *Lingua*, 115(4): 591–610.
- Beaver, D. (2006). The new mythology of Conventional Implicature. Présentation invitée à l'occasion du soixantième anniversaire de Robert van der Sandt. Accessible sur le site de cette manifestation : <http://ncs.ruhosting.nl/robshop/>
- Bonami, O. et Godard, D. (2007). Quelle syntaxe, incidemment, pour les adverbes incidents? Dans: *Bulletin de la Société de Linguistique de Paris*, 102.1: 255–284.
- Carlson, G. N. (1977). *Reference to Kinds in English*. Thèse de doctorat. Publié en 1980 par Garland Press, New York.
- Carlson, L., Marcu, D. et Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of Rhetorical Structure Theory. Dans: J. van Kuppevelt and R. Smith (dir.), *Current and New Directions in Discourse and Dialogue*. Dordrecht: Kluwer Academic Publishers, pp. 85–112.
- Combettes, B. (1998). *Les constructions détachées en français*. Gap: Ophrys.
- Deulofeu, H.-J. (2003). L'approche macrosyntaxique en syntaxe: un nouveau modèle de rasoir d'Occam contre les notions inutiles? *Scolia*, 16: 77–95.
- Fabricius-Hansen, C. et Behrens, B. (2001). Elaboration and related discourse relations viewed from an interlingual perspective. Dans: C. Smith (dir.), *Proceedings of the Third Workshop on Text Structure*, University of Austin, Texas, October 2000. SPRIK Report No. 13, University of Oslo, 2001.



- Fernald, T. B. (2000). *Predicates and Temporal Arguments*. Oxford: Oxford University Press.
- Geurts, B. (1995). *Presupposing*. Thèse de doctorat, Université de Stuttgart.
- Halliday, M. A. K. (1994). *An Introduction to Functional Grammar*, second edition. London: Edward Arnold.
- Higginbotham, J. (2000). On events in linguistic semantics. Dans: J. Higginbotham, F. Pianesi et A. C. Varzi (dir.), *Speaking of Events*. Oxford: Oxford University Press. Oxford University Press, pp. 49–79.
- Hovy, E. H. et Maier, E. (1991). Parsimonious or profligate: how many and which discourse structure relations? Manuscrit non publié mais accessible à l'adresse suivante: [www.isi.edu/natural-language/people/hovy/papers/93discproc.pdf](http://www.isi.edu/natural-language/people/hovy/papers/93discproc.pdf)
- Knott, A., Oberlander, J., O'Donnell, M. et Mellish, C. (2001). Beyond elaboration: the interaction of relations and focus in coherent text. Dans: T. Sanders, J. Schilperoord and W. Spooren (dir.), *Text representation: linguistic and psycholinguistic aspects*. Amsterdam/Philadelphia: John Benjamins, pp. 181–196.
- Kratzer, A. (1995). Stage-level and individual-level predicates. Dans: G. Carlson and F. Pelletier (dir.), *The Generic Book*. University of Chicago Press, pp. 125–175.
- Kratzer, A. (2003). *The Event Argument and the Semantics of Verbs*. (Manuscrit non publié).
- Maienborn, C. (2004). A Pragmatic Explanation of the Stage Level/Individual Level Contrast in Combination with Locatives. Dans: B. Agbayani, V. Samiian et B. Tucker (dir.), *Proceedings of the Western Conference on Linguistics (WECOL 2003)*, volume 15, Fresno: CSU, pp. 158–170.
- Mann, W. C. et Thompson, S. A. (1987). *Rhetorical Structure Theory: a theory of text organisation*. Technical report ISI/RS-87-190, Information Sciences Institute.
- Marandin, J.-M. (1998). *Grammaire de l'incidence*. (Manuscrit non publié).
- Milner, J.-C. (1978). *De la Syntaxe à l'interprétation*. Paris: Le Seuil.
- Parsons, T. (1990). *Events in the Semantics of English*. Cambridge (Mass.): MIT Press.
- Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford: Oxford University Press.
- Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., Joshi, A., Robaldo, L. et Webber, B. (2007). *The Penn Discourse TreeBank 2.0. Annotation Manual*. IRCS Technical Report, Institute for Research in Cognitive Science, University of Pennsylvania.
- Reese, B., Denis, P., Asher, N., Baldridge, J. et Hunter, J. (2007). *Reference Manual for the Analysis and Annotation of Rhetorical Structure*. Technical report under the auspices of National Science Foundation Grant IIS-0535154, University of Texas at Austin.
- van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, 9.4: 333–377.
- Vieu, L., Bras, M., Asher, N. et Aurnague, M. (2005). Locating adverbials in discourse. *Journal of French Language Studies* 15: 173–193.
- Vieu, L. et Prévot, L. (2004). Background in Segmented Discourse Representation. Dans: L. Danlos et N. Asher (dir.), *Actes de l'atelier de travail Segmented Discourse Representation Theory*. 11ème conférence Traitement Automatique des Langues Naturelles (TALN), 485–494, Fès, Maroc, Juin.
- Wolf, F. et Gibson, E. (2005). Representing discourse coherence: A corpus-based study. *Computational Linguistics* 31(2): 249–287.