

QUELQUES COMBINAISONS DE CONNECTEURS DISCURSIFS

Pascal AMSILI

Laboratoire de Linguistique Formelle

Université Paris Diderot

pascal.amsili@linguist.univ-paris-diderot.fr

Grégoire WINTERSTEIN

Division of Linguistics and Multilingual Studies

Nanyang Technological University

gregoire.winterstein@linguist.univ-paris-diderot.fr

RESUME

Ce travail s'intéresse aux discours dans lesquels deux mêmes unités de discours sont reliées par plusieurs connecteurs. Nous abordons la question en étudiant deux cas particuliers. Le premier met en jeu les marqueurs mais et aussi, respectivement marqueurs des relations de discours CONTRASTE et PARALLELE. Pour ces deux marqueurs, nous montrons que bien que leur sémantique paraisse contradictoire toute approche sémantique du discours doit pouvoir accommoder le fait que ces relations peuvent coexister et relier les mêmes unités discursives. Le deuxième cas d'étude concerne les marqueurs en effet et et. Nous montrons comment la sémantique de et ne peut se composer qu'avec une seule des lectures possibles de en effet, ce qui a pour effet de désambiguïser la relation de discours marquée par ce dernier.

ABSTRACT

This work is concerned with discourses in which two discourse units are connected with several connectives. We focus on two specific cases. In the first case, we consider the French connectives mais and aussi, which are used to denote the discourse relations CONTRAST and PARALLEL. We show that these two connectives are capable of connecting the same discourse units, even though their semantic descriptions suggest that they are incompatible, and that this fact has to be taken into account in any formal approach to discourse. The second case involves the connectives en effet and et. We show that the semantic properties of et are such that it can only combine with one of the possible interpretations of en effet, which, as a consequence, disambiguates en effet when it is combined with it.

1 Introduction

Nous nous intéressons dans cet article à la façon dont plusieurs relations de discours (ou plusieurs connecteurs qui marquent des relations de discours) peuvent se combiner, ou simplement coexister, ce qui conduit à des effets particuliers que nous allons illustrer avec quelques exemples.

1.1 Relation de discours, connecteurs de discours

Les relations de discours sous-tendent une grande partie des approches formelles du discours telles que la RST (Mann & Thompson, 1988) ou la SDRT (Asher, 1993; Asher & Lascarides, 2003). Ces approches s'intéressent en effet aux relations rhétoriques qui s'établissent entre les unités minimales du discours, c'est-à-dire, en première approximation, les propositions qui forment le discours soit par leur juxtaposition ou par des relations de coordination (parataxe) soit par des relations syntaxiques de subordination (hypotaxe).

Dans l'exemple (1a), le segment *il faisait noir* est présenté comme une *justification* du premier segment *Jean a allumé la lumière*. Cette relation est inférée par les interprètes sur la base du contenu des segments, de connaissances du monde, et d'attentes générales sur la cohérence des discours. Dans d'autres cas, comme dans la variante (1b), un connecteur (ici la conjonction *car*) est explicitement introduit, forçant en quelque sorte une interprétation qui restait implicite dans le premier cas.

- (1) a. Jean a allumé la lumière, il faisait noir.
- b. Jean a allumé la lumière, car il faisait noir.

Depuis les travaux de Hobbs (1985), la notion de relation de discours a fait l'objet d'une littérature abondante, par exemple sur la distinction entre relations subordonnantes et coordonnantes (Polanyi, 1985) ainsi que sur l'inventaire des relations et leur organisation hiérarchique (Grosz & Sidner, 1986; Mann & Thompson, 1988; Asher, 1993).

De nombreux travaux ont aussi porté sur les effets sémantiques des relations de discours. En effet, la structuration rhétorique des discours ne fonctionne pas comme une couche autonome qui viendrait se surimposer sur les contenus sémantiques des segments du discours: dans de nombreux cas, la structure rhétorique force une interprétation particulière des segments, et ce phénomène a été abondamment illustré, en particulier par Asher & Lascarides (2003), à propos de la résolution des anaphores, de

l'interprétation des temps verbaux, des présuppositions, du bridging, de l'utilisation de constructions elliptiques... Avec tous ces phénomènes, on observe que la reconnaissance par les interprètes de la relation de discours en jeu est nécessaire à l'interprétation correcte des contenus sémantiques des segments.

Mais en retour, on observe aussi, et ce fait constitue l'une des difficultés majeures pour la modélisation des relations de discours, que dans certains cas c'est l'interprétation spécifique de certains segments qui mène à l'identification de la relation rhétorique qui les relie. C'est le cas de la fameuse « *push-fall law* » invoquée à propos de l'exemple (2) devenu classique (Asher, 1993):

- (2) a. Max fell. John helped him up.
b. Max fell. John pushed him.

Dans le premier cas, la relation déduite par l'interprète est une relation de séquence temporelle (le deuxième événement décrit se produit après le premier, en réponse), et cette analyse est basée sur les contenus lexicaux des deux segments, ainsi que sur des connaissances du monde ; dans le second cas la relation privilégiée est une explication, le second événement étant cette fois présenté comme précédant le premier (et le causant), là encore sur la base d'hypothèses sur le monde.

La prise en compte des relations rhétoriques est donc, dans de nombreux cas, une dimension à prendre en compte pour aboutir à une interprétation sémantique correcte des discours, et se trouve confrontée à deux types de difficultés: d'une part l'absence de connecteur explicite rend nécessaire d'inférer la relation de discours sur la base du contenu des segments en relations, il y a sur ce plan potentiellement une circularité ; d'autre part, quand un connecteur de discours est présent, il peut être ambigu (au sens où il peut véhiculer plusieurs relations de discours), et laisser donc entier le problème évoqué à l'instant.

1.2 Relations multiples

À ces difficultés bien connues et qui ont fait l'objet de nombreux travaux, nous voulons en ajouter une, qui est à notre connaissance bien souvent laissée de côté: il s'agit de la coexistence de multiples relations entre deux segments donnés, qu'elles soient ou non marquées par de multiples connecteurs de discours. C'est surtout à ce dernier cas que nous nous intéresserons dans ce texte, en prenant quelques exemples.

Cette question de la multiplicité des relations conduit à deux problématiques distinctes : d'une part, on peut se demander dans quelle mesure les relations de discours, telles qu'elles sont définies dans les différentes théories, sont compatibles entre elles, et dans quelle mesure leur combinaison est possible, et prévue. D'autre part, sur le plan de l'interface

syntaxe-sémantique, on peut se demander quels mécanismes régissent la combinaison de plusieurs connecteurs discursifs, et dans quelle mesure on peut parler de compositionnalité à leur propos.

Évoquons rapidement le point de vue de quelques-unes des approches les plus influentes sur la question, en l'occurrence la RST, la SDRT, et l'approche adoptée dans la réalisation du Penn Discourse TreeBank (PDTB).

La **RST** se présente comme une théorie descriptive de l'organisation textuelle, et doit par conséquent être vue plus comme un outil à destination de l'analyste que comme un modèle de l'organisation du discours, capable de prédire les observations empiriques sur le discours. On notera par ailleurs que la RST se fonde peu sur les connecteurs qui ne sont pas considérés, dans cette théorie, comme des indicateurs fiables.

L'analyse structurelle selon la RST est censée être soumise à plusieurs contraintes (dont la justification n'est pas réellement fournie par Mann & Thompson (1988)) qui rendent délicate la prise en compte de combinaisons de relations discursives. Parmi ces contraintes, on trouve en effet une contrainte d'**unicité**, selon laquelle il n'y a pas de situation où deux relations distinctes s'appliquent à la même paire de segments textuels¹.

On en déduit que du point de vue de la RST, la coexistence de deux relations distinctes entre les mêmes segments textuels constitue une exception qui ne semble justifiée que dans des cas très rares (parmi lesquels figurent les textes dont le parallélisme est l'élément structurant principal).

Le cadre de la **SDRT** prévoit la possibilité d'avoir plusieurs relations entre deux mêmes segments. La multiplicité des relations est par ailleurs conçue comme un aspect désirable de l'interprétation d'un discours. Cette hypothèse est une conséquence directe du **Principe de Maximisation de la Cohérence** qui stipule (entre autres) que si l'interprétation *I* d'un discours est plus cohérente qu'une interprétation *I'*, alors l'ensemble des relations de discours vérifiées dans *I'* se retrouvent dans *I*. La contrainte majeure qui pèse sur la coexistence de relations est celle de cohérence: on préférera des discours au sein desquels les relations inférées ne se contredisent pas.

¹ Les autres contraintes, pas non plus vraiment justifiées, semblent relativement raisonnables, même s'il n'est pas établi qu'elles s'appliquent toujours en toute circonstance: **complétude** et **connexité** (aucun pan d'un discours cohérent ne devrait échapper à l'analyse, et tous les pans du discours devraient être reliés d'une façon ou d'une autre), et **adjacence** (les segments en relation sont adjacents) (Mann & Thompson, 1988: p. 248).

Dans le cadre du projet d'annotation de relations de discours *Penn Discourse TreeBank* (Prasad et al., 2008)², la question des relations multiples entre deux segments de discours est abordée d'un point de vue pratique. De manière générale, lorsque deux connecteurs sont présents au sein d'un même segment, ils sont annotés séparément en indiquant de manière indépendante pour chacun d'entre eux quels sont ses arguments.

Cependant, comme il est souligné dans le manuel d'annotation, ce choix ignore la possibilité qu'un des deux connecteurs doive être interprété par rapport à l'autre. Webber et al. (2003) montrent entre autres que cette possibilité est réelle. Ainsi en (3) deux connecteurs coexistent dans le même segment de discours: *parce que* et *par exemple*.

- (3) Deux variables peuvent être fortement corrélées sans pour autant être dans une relation causale, *parce que par exemple*, elles dépendent toutes deux d'un facteur non étudié. (*Psychologie de la personnalité*, M. Hansenne, p. 38)

Ici le connecteur *par exemple* n'introduit pas une exemplification du premier segment de discours de (3) mais bien un exemple d'*explication* telle que *parce que* est à même d'introduire. L'interprétation de *par exemple* dépend donc crucialement de la présence de *parce que*. On peut noter que l'ordre de la combinaison est pertinent: la combinaison inverse *par exemple parce que* ne serait pas analysée de la même façon.

Le seul cas pour lequel des relations multiples sont annotées conjointement entre des mêmes segments de discours dans le PDTB concerne le cas particulier du connecteur vide \emptyset , c'est-à-dire le connecteur que l'on suppose relier deux segments de discours comme ceux de **Erreur ! Source du renvoi introuvable.** qui ne sont par ailleurs reliés par aucun connecteur explicite. Dans ce cas précis, une unique annotation est effectuée en indiquant l'ensemble des relations associées au connecteur vide. Toutefois cette annotation ne permet pas non plus d'indiquer d'éventuelles dépendances entre les connecteurs comme on a pu observer en (3). En résumé, même si le projet PDTB prévoit la possibilité de l'existence de plusieurs relations entre deux mêmes segments, il ne prévoit pas que deux connecteurs distincts marquent deux relations distinctes portant sur deux segments de discours identiques.

² Le PDTB est le résultat d'un vaste projet d'annotation discursive. Il ajoute aux annotations du Penn TreeBank une couche d'annotation identifiant les relations de discours présentes dans les textes du corpus ainsi que les arguments de ces relations.

1.3 Plan de l'article

Nous nous proposons d'étudier différents cas qui mettent en jeu la présence de deux connecteurs distincts au sein d'un même segment discursif. Chacun des connecteurs véhicule sa propre relation de discours (au moins), ce qui impose de prendre au sérieux la question de la combinaison des relations au niveau sémantique. On se demandera notamment comment des relations potentiellement contradictoires peuvent coexister et comment la sémantique d'une relation peut influencer l'interprétation d'un autre connecteur. Afin de rendre notre approche aussi explicite que possible, nous nous proposons, lorsque cela est pertinent, d'utiliser le cadre de la SDRT pour représenter nos analyses.

En premier lieu nous étudions la combinaison des relations de CONTRASTE et de PARALLELE, respectivement marquées par les éléments *mais* et *aussi*. Étant donnée la description de ces deux relations, leur coexistence dans un même discours peut paraître étonnante, et nous montrons qu'il est cependant nécessaire de les faire cohabiter.

Nous nous intéressons ensuite au cas de la combinaison de la conjonction *et* avec le marqueur *en effet*. L'intérêt de cette combinaison réside dans le fait que la relation marquée par *en effet* est ambiguë et que l'ajout de *et* a pour effet de désambigüiser l'emploi de *en effet*.

2 La combinaison *mais... aussi*

La combinaison des marqueurs *mais* et *aussi* est extrêmement fréquente. Une recherche au sein du French Treebank (Abeillé et al., 2003) met au jour 84 occurrences de segments contenant la combinaison *mais... aussi*³. Pour compléter ce décompte nous avons cherché à comparer le degré d'associativité de *mais* et *aussi* à celui de la combinaison d'autres connecteurs (en nous basant sur l'éventail de connecteurs retenus dans la base Lexconn (Roze et al., 2012)). Pour mesurer ce degré d'associativité nous nous sommes appuyés sur les algorithmes de recherche de collocations proposés dans la bibliothèque `nltk` (Bird et al., 2009). En associant à chacune des phrases du FTB la liste des connecteurs qui s'y trouvent, nous avons cherché au sein de l'ensemble de ces listes les paires de connecteurs qui se trouvent statistiquement le plus souvent ensemble en utilisant trois métriques pour évaluer la significativité de l'association (ratio de vraisemblance, χ^2 et test *t* de Student). Pour chacune de ces métriques,

³ On a pris soin ici et dans le décompte précédent d'écarter les exemples mettant en jeu la combinaison *non seulement . . . mais aussi*.

l'association de *mais* et *aussi* s'avère être la plus significative parmi l'ensemble des associations de connecteurs observées.⁴

Ces observations, ajoutées au caractère a priori contradictoire de *mais* et de *aussi* justifient de s'arrêter sur leur combinaison.

2.1 Propriétés de la relation de contraste

On considère que lorsqu'il y a une relation de CONTRASTE entre deux segments de discours, l'accent est mis sur les différences qui existent entre leurs contenus. Cela ne signifie évidemment pas que tous les segments de discours, dès lors qu'ils sont différents, peuvent entretenir une relation de discours: comme pour d'autres relations coordonnantes, pour qu'un CONTRASTE puisse être attesté entre deux segments, il faut qu'ils soient *comparables* (Lang, 1984), c'est-à-dire qu'ils doivent être à la fois similaires et différents (voir plus loin pour plus de détails concernant la mesure pertinente de similarité).

La différence entre les segments peut se manifester de plusieurs façons: il peut s'agir d'une différence sémantique, que ce soit à un niveau lexical (*grand/petit* dans (4a)), ou non lexical (4b), mais il peut s'agir aussi d'une différence pragmatique (4c).

- (4) a. Paul est grand mais Marie est petite.
b. Paul est venu mais pas Marie.
c. Paul fume, mais il est en bonne santé.

Cette distinction entre différence sémantique et différence pragmatique reprend une classification fermement établie dans l'étude des marqueurs contrastifs (depuis les travaux de Lakoff (1971)). Dans le premier cas, on parle d'**opposition** (sémantique), alors que dans le second cas, on parle d'usage **concessif**, et dans ce cas le deuxième segment s'oppose à une inférence qu'on pourrait tirer du premier segment et pas au contenu principal du premier segment.

En (4) on considère que *mais* est un marqueur monotone de la relation de CONTRASTE, c'est-à-dire que sa présence marque de manière certaine que c'est cette relation qui relie les deux arguments de *mais*⁵.

Les expressions linguistiques qui servent (ou peuvent servir) à marquer la relation de CONTRASTE sont essentiellement des connecteurs, qui peuvent être (syntactiquement) coordonnants (*mais, pourtant, cependant...*), ou (syntactiquement) subordonnants (*bien que, malgré que...*). Comme pour la

⁴ Parmi les autres combinaisons ressortant de manière systématique et susceptibles de mettre en jeu une forme de compositionnalité on trouve: *et surtout, sinon du moins* et *sans pour autant*.

⁵ Voir (Winterstein, 2012) pour une argumentation en faveur d'un traitement unifié de *mais* dans les différents cas de (4).

plupart des relations de discours, on trouve aussi de nombreux cas où la relation de contraste n'est pas marquée par un connecteur.

Dans l'exemple (5) l'emploi de *mais* est par exemple un cas de contraste basé sur une opposition sémantique.

(5) Le kilo de bœuf, pour le chanceux qui parvenait à se le procurer, valait officiellement moins de 10 roubles dans les magasins d'État [...] mais plusieurs centaines de roubles sur les marchés paysans. (*French TreeBank*)

Les segments reliés par la relation de CONTRASTE ont en commun d'être relatifs au prix du bœuf, et se distinguent sur la hauteur du prix en question. Ils sont donc bien *comparables* au sens de Lang (1984).

Au contraire en (6) (également tiré du FTB), le CONTRASTE est un cas de concession: la perte annoncée en octobre paraît incompatible avec l'optimisme affiché dans le second conjoint.

(6) En octobre, le groupe annonce une perte semestrielle de 68 millions de francs, mais se veut optimiste pour l'ensemble de l'année.

Un test fréquemment invoqué pour repérer les usages concessifs de *mais* consiste à pouvoir lui substituer le marqueur *pourtant* sans modification de sens. C'est bien ce qui semble être le cas en (7) (minimalement modifié pour avoir un segment droit syntaxiquement bien formé).

(7) En octobre, le groupe annonce une perte semestrielle de 68 millions de francs, *pourtant* il se veut optimiste pour l'ensemble de l'année.

A un niveau intuitif, il est assez facile de caractériser sémantiquement la relation de CONTRASTE, comme on le fait par exemple en RST (Mann & Thompson, 1988): les situations présentées dans les deux segments doivent présenter de nombreuses similitudes, et aussi quelques différences, de manière à ce que l'accent puisse être mis à la fois sur la comparabilité et sur les différences entre les situations.

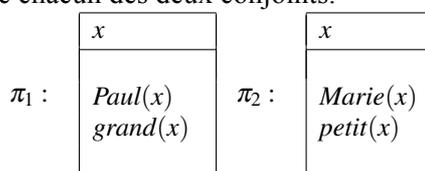
Pour aller plus loin dans la caractérisation sémantique, on peut proposer la description suivante en SDRT (tirée de (Asher & Lascarides, 2003) qui renvoie à (Asher, 1993) et à (Asher et al., 2001) pour les définitions relatives à l'établissement des thèmes contrastifs).

- CONTRASTE est véridique (c.à.d. que ses deux arguments α et β doivent être vrais dans le contexte d'énonciation).
- α et β doivent avoir des structures sémantiques similaires. Plus spécifiquement, il doit exister un isomorphisme partiel entre les DRSs de α et celles de β . Plus cette correspondance est proche d'un isomorphisme, meilleure sera la relation de CONTRASTE.
- Il doit exister un thème contrastif entre α et β . Plus ce thème est contrastif, meilleure sera la relation. Le degré de contraste est défini en termes de degrés de différence entre les propositions qui marquent les thèmes des nœuds de la structure sémantique. La différence

maximale est entre des propositions p et q telles que $p \sim \neg q$, c.à.d. telles qu'un constituant nie une conséquence par défaut de l'autre (dans la logique non-monotone utilisée en SDRT).

La contrainte d'isomorphisme assure que les segments ont suffisamment en commun pour pouvoir être comparés, et le thème contrastif est une mesure du degré de différenciation des deux segments.

Pour illustrer ces contraintes, prenons l'exemple de (4-a). Le cœur de la relation est donné par la construction du thème contrastif. On indique ci-dessous la structure de chacun des deux conjoints.



L'isomorphisme de structure apparaît relativement clairement: chacune des SDRS introduit un unique référent de discours sur lequel sont prédiquées deux propriétés. On peut donc établir une correspondance entre les constituants de chacune des structures.

Deux constituants d'une SDRS seront dits contrastifs s'ils ont des polarités différentes, au sens de Asher (1993: p. 305), c'est-à-dire qu'ils s'avèrent plausiblement complémentaires, modulo le remplacement de référents de discours. Dans un exemple comme (4-a) le degré de contraste est bon: les éléments contrastifs sont *Paul* et *Marie* d'une part et *grand* et *petit* d'autre part: ils occupent la même position dans la structure et sont distincts deux à deux (par exemple aucun n'implique l'autre).

On peut donc bien établir que d'une part les conjoints ont des structures parallèles et que d'autre part les éléments présents dans ces structures sont sémantiquement distincts: la relation de CONTRASTE est licite.

C'est cela qui explique que la variante suivante de (4-a), où le marqueur de contraste est absent, est préférentiellement interprétée comme un CONTRASTE (8-a), alors que la variante sans marqueur explicite de (4-c) n'a pas cette propriété (8-b).

- (8) a. Paul est grand, Marie est petite.
 b. ? Paul fume, il est en bonne santé.

Dans le cas de (4-c), l'isomorphisme de structure est moins net: la structure abstraite qui subsume les deux conjoints est moins riche que celle de (4-a). Par ailleurs, le degré de contraste sera établi au niveau inférentiel. Ces différences peuvent expliquer qu'en (4-c) le marqueur *mais* s'avère nécessaire pour marquer la relation de CONTRASTE, alors qu'en (4-a) les

structures des conjoints suffiraient à elles seules pour établir la présence de cette relation⁶.

2.2 Propriétés de la relation de parallèle

Le parallélisme est une relation de discours qui intervient lorsque deux ou plusieurs segments de discours sont présentés comme comprenant des éléments identiques ou similaires.

Les expressions linguistiques les plus couramment utilisées pour manifester cette relation ne sont pas des connecteurs au sens étroit du terme (syntaxiquement ce ne sont pas des conjonctions de coordination ou de subordination), mais plutôt des adverbes, appelés additifs, tels que *aussi*, *de même*, *autant*, *également*... Les exemples de (9)-(11) (issus du FTB) illustrent différentes possibilités de réalisation de la relation PARALLELE.

- (9) BSN a atteint son plus haut cours historique à la fin de l'année et progressé de 40 %. L'Oréal lui **aussi** a "surperformé" le CAC en gagnant 53 % sur douze mois.
- (10) Aucune mesure concrète n'est proposée pour l'accès prioritaire des conjoints fonctionnaires dans la zone d'accueil. **De même**, la bourse d'emploi interministérielle, envisagée pour favoriser la mobilité des agents, se heurte aussi à des blocages statutaires.
- (11) Pour Unilever, ce rachat entre dans la stratégie d'implantations dans le Sud-Est asiatique, notamment dans le domaine de produits alimentaires, le thé. Unilever est **également** présent en Thaïlande dans le domaine des lessives et des articles de toilette.

Tout comme la relation de CONTRASTE (et de manière plus intuitive) PARALLELE impose que ses deux arguments possèdent une structure commune. À la différence du CONTRASTE, cette similarité de structure doit s'accompagner d'une similarité sémantique. En SDRT, on décrit cette relation de la façon suivante.

- PARALLELE est véridique (c.à.d. que ses deux arguments α et β doivent être vrais dans le contexte d'énonciation).
- α et β doivent avoir des structures sémantiques similaires. Plus spécifiquement, il doit exister un isomorphisme partiel entre les structures DRSs de α et celles de β . Plus cette correspondance est proche d'un isomorphisme, meilleure sera la relation de PARALLELE.

⁶ Comme l'a fait remarquer un relecteur, dans (4-c), la conjonction *et* peut aussi être utilisée, avec une interprétation contrastive. Il n'en reste pas moins que le marquage explicite de la relation semble nécessaire, alors que ce n'est pas le cas dans (4-a). Par ailleurs l'emploi de *et* ajouterait une composante de surprise absente avec *mais*. Voir par exemple (Blakemore & Carston, 2005) à ce sujet.

- Il doit exister un thème commun entre α et β . Plus ce thème est informatif, meilleure sera la relation.

La contrainte d'isomorphisme est identique à celle de CONTRASTE, la relation de PARALLELE se distinguant en imposant l'existence d'un thème commun plutôt qu'un thème contrastif entre les deux segments.

La construction du thème commun se fait en abstrayant le contenu commun des segments mis en jeu. Plus le thème commun construit est riche et meilleure sera la relation de PARALLELE. Le thème commun est une DRS construite à partir de trois opérations élémentaires qui éliminent des structures au sein d'une DRS:

- Suppression d'un marqueur de discours (c.à.d. d'une variable au sein d'une DRS).
- Suppression d'une condition atomique.
- Renommage systématique d'un marqueur de discours lié.

Dans le cas de (12), l'établissement de l'isomorphisme structurel est identique au cas de (4-a).

(12) Paul est grand. Marie est également grande.

L'établissement du thème commun donne la DRS K où seule la condition établissant l'identité de la personne "grande" a disparu.

$$K : \begin{array}{|c|} \hline x \\ \hline grand(x) \\ \hline \end{array}$$

K est relativement informatif et conserve une partie significative des informations des deux unités de discours considérées, la relation de PARALLELE est donc ici appropriée.

On peut noter que bien que la relation soit appropriée, et l'isomorphisme important, on ne peut pas, comme c'était le cas dans le cas du CONTRASTE, se passer du marqueur de la relation de parallèle:

(13) # Paul est grand. Marie est grande.

La raison de cette difficulté est différente de celle de l'exemple (8-b) évoqué plus haut: dans ce dernier cas, la difficulté venait de la difficulté de repérer dans les contenus présentés des éléments formels à mettre en contraste. Dans le cas présent, le parallélisme est évident, et il semble bien que ce soit précisément pour cette raison que le marqueur additif est obligatoire. Voir (Amsili & Beyssade, 2010; Eckard & Fränkel, 2012).

2.3 mais... aussi

La présentation des deux relations de la section précédente suggère que les relations CONTRASTE et PARALLELE sont incompatibles. En effet si les deux relations imposent une similarité de structure entre leurs arguments, leurs contraintes sont opposées au niveau sémantique: CONTRASTE impose d'avoir une différence alors que PARALLELE recherche une similarité maximale.

Dans cette section nous montrons qu'au niveau empirique il est nécessaire de considérer que les deux relations peuvent lier les mêmes segments de discours et nous examinons la façon dont cette cohabitation est représentable dans le cadre de la SDRT.

2.3.1 Données

Si on admet que les connecteurs *mais* et *aussi* marquent de façon non ambiguë la présence des relations de CONTRASTE et PARALLELE, alors les cas de (14) (tirés du French Treebank) suggèrent fortement que les deux relations peuvent se combiner.

- (14) a. C'est le rôle de l'Etat, amplement souligné par Keynes **mais aussi** par certains de ses prédécesseurs, que de rétablir la confiance lors d'une récession et de prévenir ainsi des enchaînements cumulatifs à la baisse.
- b. Ils prennent beaucoup de risques, de véritables coups de poker qui rapportent gros, à l'image de leur ancêtre l'amiral Miaoulis, héros de la guerre d'indépendance de 1821, **mais aussi** armateur qui a forcé le blocus de Nelson pour commercer avec Marseille.
- c. Le côté sentimental a donc joué quelque peu, **mais** le prix de la saillie de Laudanum (25000 francs) est **aussi** un gage de la qualité d'étalon de ce pur-sang de vingt-trois ans.
- d. On peut en douter, **mais** on peut **aussi** s'en préoccuper puisqu'une large part de cet avenir est prévisible et que les problèmes se posent en termes de mondialité

On peut se poser la question de savoir quels sont les segments qui entrent dans les relations marquées par *mais* et *aussi*. Si l'identité de l'argument gauche ne fait pas de doute (par exemple en (14-d) ce sera *On peut en douter*), celle de l'argument droit est moins claire. Le paramètre de variation est relatif à l'inclusion, ou non, du connecteur tiers dans l'argument droit de la relation considérée. En termes plus concrets, dans le cas de (14-d), et pour la relation dénotée par le connecteur *mais*, les options pour l'argument droit sont:

- β =On peut s'en préoccuper
- β' =On peut aussi s'en préoccuper

De la même façon, pour la relation dénotée par *aussi* en (14-d) les options sont:

- β =On peut s'en préoccuper
- β'' =Mais on peut s'en préoccuper

Si le choix de β'' pour le PARALLELE semble peu convaincant (le segment ne forme pas une clause autonome syntaxiquement), l'alternative ouverte par β' l'est déjà plus. Or si l'on veut parler de composition de relations de discours il est nécessaire de montrer que cette composition se fait sur des relations portant sur des segments identiques.

Nous voulons démontrer que ce choix n'importe au final que très peu: même si le segment considéré par la relation CONTRASTE est β'' , les contenus pris en compte pour l'établissement de la relation de contraste correspondent au contenu principal du segment droit, qui dans ce cas précis est identique à celui de β . Pour cela on peut s'appuyer en partie sur la démonstration de Winterstein (2013) qui montre que le contenu pris en compte par *mais* dans sa sémantique se limite au contenu principal et ne tient pas compte d'autres niveaux du sens, notamment les implicatures de quantité ou les présuppositions.

Dans le cas qui nous intéresse la seule différence entre β et β'' tient à la présupposition introduite par l'emploi de *aussi*.⁷ Le contenu de cette présupposition est glosé en β'_{psp} .

(15) β'_{psp} : On peut faire autre chose que s'en préoccuper.

Au niveau informatif, le contenu de β'_{psp} est impliqué par celui de α . L'établissement du thème contrastif devant porter sur les différences entre les éléments pris en compte, c'est donc dans tous les cas sur la partie de β qui n'est pas impliquée par α que ce thème va se construire. Dans le cas étudié, ce sera donc essentiellement le fait qu'on puisse faire quelque chose d'autre que *douter* (et non que *s'en préoccuper*) qui permettra l'établissement du contraste (i.e. ce qui correspond à l'unique différence

⁷ Différents travaux, p.ex. (Amsili & Winterstein, 2012), montrent que la sémantique de *aussi* ne peut se réduire à sa présupposition et que cet adverbe encode également une composante véhiculant la similarité de son hôte et de l'antécédent de sa présupposition. Dans le cas de (14) les segments mis en jeu sont lexicalement identiques, ce qui rend cette condition supplémentaire triviale.

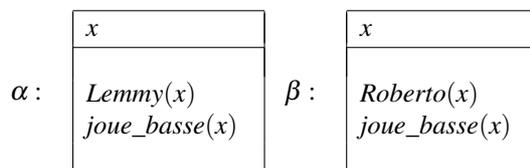
entre α et β). En résumé, le contenu sémantique pertinent pour l'établissement du CONTRASTE correspond dans tous les cas à celui de β .⁸

2.3.2 Formalisation en SDRT

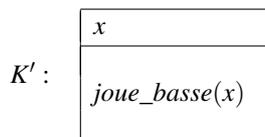
Nous nous intéressons maintenant à la façon dont les contraintes stipulées dans le cadre de la SDRT gèrent les cas de (14). Pour simplifier notre présentation, nous nous basons sur l'exemple construit (16).

(16) Lemmy joue de la basse, mais Roberto en joue aussi.

Les SDRS correspondant à chacun des conjoints de (16) sont présentées ci-dessous:



L'isomorphisme structural est ici évident de la même manière que pour (4-a). Pour la relation de PARALLELE l'établissement d'un thème commun est également simple: le thème obtenu est K' :



Il faut maintenant montrer comment obtenir un thème contrastif à partir des mêmes éléments. Étant donné la nature des éléments, il semble difficile de considérer qu'on est en présence d'une opposition sémantique: loin d'avoir des prédicats antonymes ou même distincts, les éléments en présence sont lexicalement identiques (et c'est ce qui favorise la relation de PARALLELE).

Nous nous proposons alors de suivre l'hypothèse de Winterstein (2012) qui considère que le cas de (16) est un cas d'emploi concessif de *mais*.

⁸ Un relecteur fait remarquer le cas intéressant de (i) :

(i) Jean aime Marie, (mais) il la déteste #(aussi).

En (i), il apparaît que *mais* est optionnel alors que la présence de *aussi* est très nettement préférée. Cette observation a potentiellement des conséquences intéressantes sur la portée des relations de discours marquées par ces connecteurs. Toutefois, pour des raisons de place, nous devons reporter cette discussion à des travaux ultérieurs. Nous nous contentons ici de souligner que dans certains cas, les deux relations peuvent relier deux unités de discours identiques et avoir des interprétations compatibles.

L'argument principal à l'appui de cette hypothèse consiste à considérer que, pris isolément, le segment α active l'inférence *Lemmy est le seul à jouer de la basse*. Cette activation est expliquée en termes probabilistes en considérant qu'en assertant α , le locuteur augmente la probabilité que Lemmy soit le seul joueur de basse. C'est alors cette inférence qui est directement niée par le second conjoint. Le thème contrastif construit ici est donc bien construit sur le mode d'une concession, i.e. à un niveau pragmatique.

En résumé, la combinaison des relations de PARALLELE et de CONTRASTE dans les exemples (14) et (16) est rendue possible par les faits suivants:

- Les deux relations reposent sur un isomorphisme de structure
- La relation de PARALLELE est établie sur des critères de similarité lexicale.
- La relation de CONTRASTE repose sur un contraste d'ordre pragmatique et non purement sémantique, ce qui permet d'éviter le conflit avec les contraintes propres à la relation de PARALLELE.

Une des conséquences de cette hypothèse est que l'on prédit que l'absence de *mais* entraînera la disparition de la relation de CONTRASTE. En effet, comme déjà noté, lorsque la relation marquée par *mais* est d'ordre concessif, elle s'avère difficilement inférable lorsque le connecteur est omis, contrairement aux cas d'opposition sémantique. Cette prédiction semble vérifiée, par exemple si on ôte *mais* en (14-d):

- (17) On peut en douter, on peut aussi s'en préoccuper puisqu'une large part de cet avenir est prévisible et que les problèmes se posent en termes de mondialité.

3 La combinaison *et... en effet*

Nous nous intéressons maintenant à la combinaison du marqueur *en effet* avec la conjonction *et*. Contrairement au cas de *mais* et *aussi*, cette combinaison a pour effet de désambigüiser l'interprétation de *en effet*, alors que les données de la section précédente montraient comment deux relations distinctes pouvaient cohabiter.

Nous commençons par brièvement présenter les interprétations de chacun des connecteurs avant de nous pencher sur leur combinaison.

3.1 Le connecteur *et*

À première vue, il semblerait que la conjonction *et*, de par son apparente pauvreté sémantique, ne se comporte pas comme un marqueur fort d'une

relation entre ses conjoints. Jayez & Winterstein (2013) montrent cependant que lorsque *et* connecte des unités de discours, la conjonction encode une contrainte forte sur l'interprétation de son argument droit. La proposition est couchée dans le cadre d'une approche argumentative du discours et stipule que:

1. Les deux conjoints reliés par *et* doivent pouvoir être compris comme des arguments en faveur d'une même conclusion.
2. Une fois que le premier conjoint a été asserté, le deuxième conjoint doit garder un effet argumentatif, c.à.d. que ses effets ne peuvent être pris en charge entièrement par le conjoint de gauche.

C'est la seconde contrainte qui permet d'expliquer l'asymétrie de (18).

- (18) a. Paul habite en France et il vit à Paris.
b. ? Paul vit à Paris et il habite en France.

En (18-b), contrairement à (18-a), le deuxième conjoint est impliqué par le premier, et quels que soient ses effets argumentatifs, ils ont déjà été pris en compte par l'assertion du segment de gauche. Au contraire en (18-a) le deuxième conjoint est plus informatif, et il peut donc apporter des effets argumentatifs supplémentaires.⁹

3.2 Le connecteur *en effet*

On peut distinguer trois lectures du marqueur *en effet* (voir par exemple Danlos (2012)):

1. Une lecture absolue, employée en dialogue:

⁹ Comme le fait justement remarquer un relecteur, cette asymétrie peut être formalisée en stipulant un principe de non-redondance de l'information qui empêche l'assertion d'un conjoint déjà impliqué par le premier. Pour autant, cela n'expliquerait pas la différence qu'il peut exister entre (i) et (ii), tirés de (Jayez & Winterstein, 2012), similaires à des exemples proposés notamment dans (Asher & Lascarides, 2003):

(i) Paul a oublié son passeport et il a été bloqué à la frontière.

(ii) Paul a été bloqué à la frontière et il a oublié son passeport.

Dans le modèle proposé par Jayez & Winterstein, ces différences d'interprétation sont intégrées en supposant que (i) et (ii) sont articulés par rapport à des buts argumentatifs différents: l'ordre et le contenu des conjoints de (i) sont compatibles avec une lecture de RÉSULTAT et J&W supposent que cette relation est abduite préférentiellement à d'autres relations. En (ii), RÉSULTAT n'est pas accessible du fait de l'ordre des conjoints et une autre stratégie d'abduction est mise en œuvre.

- (19) a. Est-ce que Max est un crétin?
 b. En effet !
2. Une lecture de CONFIRMATION, c'est-à-dire une lecture qui montre que le segment hôte de *en effet* vient confirmer l'information présentée dans son argument gauche (souvent présenté comme attribué à une source d'information différente du locuteur, cf. Danlos (2012)).
- (20) Arthur avait dit qu'il amènerait à boire. En effet, il a amené une caisse de bière.
3. Une lecture de JUSTIFICATION, où le segment hôte de *en effet* apporte une raison de croire l'argument gauche du marqueur:
- (21) Lemmy s'attend à une mauvaise surprise pour le Bac. En effet, il a réalisé après les épreuves qu'il avait fait un hors sujet.

On s'intéresse ici aux deux dernières lectures, et au fait que dans certains cas la contribution de *en effet* est ambiguë. Ainsi en (22) il semble que les deux lectures soient accessibles.

- (22) Paul a décidé que son fils serait jockey. Il l'a en effet déjà inscrit à des cours d'équitation.

On peut interpréter le deuxième segment de (22) soit comme apportant un argument pour étayer la thèse du locuteur selon laquelle Paul veut que son fils devienne jockey (JUSTIFICATION), soit comme la marque que le locuteur reconnaît que les intentions de Paul sont réelles (CONFIRMATION).

3.3 Combinaison: désambiguïsation

L'exemple de (22) montre que la lecture *en effet* peut effectivement être ambiguë. Il a été observé que cette ambiguïté est levée lorsqu'*en effet* est combiné avec le marqueur *et* qui force alors une lecture de CONFIRMATION.¹⁰

On peut par exemple modifier (22) comme en (23):

- (23) Paul a décidé que son fils serait jockey. Et en effet, il l'a déjà inscrit à des cours d'équitation.

En (23), la seule interprétation possible pour le second segment est de le voir comme une confirmation des dires de Paul par le locuteur. On remarque par ailleurs que lorsque l'emploi de *en effet* dénote de façon claire une relation de JUSTIFICATION, l'ajout de *et* produit un énoncé dégradé.

¹⁰ Danlos (2012) prétend même que la lecture de CONFIRMATION est impossible sans la présence de *et*. Il semble effectivement y avoir une nette préférence à employer cette conjonction lorsqu'*en effet* apparaît à l'initiale, comme en (20), mais dès qu'*en effet* est employé en incise, comme en (22), cette préférence disparaît naturellement.

(24) # Lemmy s'attend à une mauvaise surprise pour le Bac. Et en effet, il a réalisé après les épreuves qu'il avait fait un hors sujet.

Pour expliquer ces faits, nous allons montrer que la sémantique de *et* n'est pas compatible avec la relation de JUSTIFICATION. Si on se place dans le même cadre argumentatif que celui supposé dans la description de *et*, l'explication est assez simple. En termes argumentatifs, une relation de JUSTIFICATION peut se gloser de la façon suivante:

(25) β justifie α ssi β se présente comme un argument pour α .

Si l'on cherche ensuite à combiner cette description avec celle proposée pour *et*, on ajoute les contraintes suivantes:

(26) a. α et β doivent se présenter comme des arguments pour une même conclusion, γ

b. β doit conserver un effet argumentatif en faveur de γ une fois que α a été asserté.

La combinaison des contraintes (25) et (26-a) permet d'identifier γ : puisque α et β doivent partager leurs conclusions, et qu'en vertu de (25), β vise α , alors $\gamma = \alpha$. La contrainte (26-b) doit alors s'interpréter de la façon suivante: il faut pouvoir interpréter β comme un argument en faveur de α après avoir supposé que α est vrai. Mais si α est supposé vrai, alors il est impossible pour β d'avoir un quelconque effet argumentatif, et la contrainte (26-b) ne peut pas être validée.

Par conséquent la seule lecture de *en effet* compatible avec *et* est la lecture de CONFIRMATION, ce qui est un effet direct de la combinaison des marqueurs en présence.¹¹

4 Conclusion

Les deux études de cas ci-dessus illustrent deux façons d'envisager la composition de connecteurs de discours (et des relations qu'ils marquent). D'une part, deux connecteurs peuvent être reliés par deux relations de

¹¹ Comme le fait remarquer un relecteur, cette conclusion découle également de l'hypothèse de Txurruka (2003) qui postule que *et* impose à ses conjoints d'entretenir des relations de discours coordonnantes et exclut les relations subordonnantes. Dans la mesure où CONFIRMATION est coordonnante et pas JUSTIFICATION, on obtient bien le résultat attendu. Cependant, Jayez & Winterstein (2013) montrent qu'avec certains connecteurs, des relations subordonnantes paraissent possibles avec *et*: cf. (ii).

(ii) Les oreillons sont de retour et c'est parce que vos enfants ne sont pas vaccinés.

Le lecteur intéressé est invité à se référer l'article en question pour plus de détails sur les différences entre *en effet* et *parce que*.

discours qui partagent une partie de leur sémantique tout en véhiculant des contenus *a priori* contradictoires par ailleurs. Cette possibilité est permise par le fait que les contributions des connecteurs peuvent se situer à des niveaux de sens différents. D'autre part, lorsque la relation marquée par un connecteur est ambiguë, la présence d'un second connecteur dans le même segment peut avoir pour effet de lever l'ambiguïté. Dans un des cas étudiés ici, c'est la contribution de *et* qui s'avère compatible avec une seule des relations marquées par *en effet*.

L'étude présentée ici est basée sur des exemples d'interaction entre connecteurs multiples. Une question plus générale sur la compositionnalité des relations de discours serait d'étudier la façon dont les contraintes propres à un connecteur se composent avec les propriétés de ses arguments, notamment du point de vue de l'optionnalité des marqueurs de discours. En effet, si l'on admet, comme c'est le cas en SDRT par exemple, qu'une relation de discours peut exister sans marqueur explicite, on est en droit de se demander quelle est alors l'utilité d'ajouter un marqueur dans ces cas précis.

Pour reprendre le cas de *mais*, la SDRT prédit que ce marqueur est optionnel lorsque la relation de CONTRASTE est sémantique. Ainsi en (27) l'emploi de *mais* sera jugé optionnel.¹²

(27) Paul aime le football. (Mais) il n'aime pas le basket-ball.

La théorie de la pertinence (Sperber & Wilson, 1986; Blakemore, 2002) apporte une piste de justification pour l'emploi de *mais* dans un cas comme (27): la présence d'un marqueur explicite sert à économiser les efforts d'interprétation de l'allocutaire, notamment du point de vue de l'établissement de la relation de discours. Cette hypothèse paraît intuitivement justifiée lorsque la relation de discours n'est pas facilement accessible, mais l'est moins lorsque la relation est par ailleurs patente. Considérons à nouveau le cas de (8-b).

(28) ? Paul fume, il est en bonne santé. =(8-b)

Le caractère dégradé de (8-b) peut être imputé au fait que nos connaissances du monde nous indiquent que les deux segments en jeu sont

¹² Nous excluons la lecture rectificative de *mais*, cf. (i) qui est souvent en concurrence avec la lecture contrastive qui nous intéresse ici.

(i) Paul n'aime pas le football, mais le basket-ball.

Dans tous les cas il semble difficile d'invoquer une lecture rectificative de *mais* dans le cas de (27) étant donné que cette lecture n'est normalement activée que dans le cas où le premier conjoint contient une négation (Anscombe & Ducrot, 1977; Jasinskaja, 2012). Dans le cas de (27), c'est dans le deuxième argument de *mais* que l'on trouve cette négation, et non dans le premier.

contradictoires: les personnes qui fument sont généralement en mauvaise santé. Les conditions d'emploi de la relation de CONTRASTE sont bien remplies. Pourtant, pour que la relation soit perçue et que l'exemple soit acceptable, l'usage d'un marqueur approprié est nécessaire. Dans ce type de cas, le rôle du marqueur ne peut alors pas être d'alléger les efforts d'interprétation en vue d'établir la relation de discours. C'est au contraire parce que les conditions d'établissement de cette relation sont saillants qu'il semble qu'il soit nécessaire de la marquer par le biais d'un marqueur approprié.

En résumé, un connecteur de discours peut contraindre l'interprétation de ses arguments par rapport à la sémantique de la relation qu'il dénote. Mais par ailleurs il est aussi possible que deux segments de discours entretiennent des relations qui correspondent aux contraintes imposées par un connecteur. Dans ce cas, alors qu'on pourrait attendre que le connecteur devienne optionnel, on rencontre des cas où l'usage d'un connecteur est précisément obligatoire. Il semble que le champ d'investigation ouvert par la prise en considération de relations de discours multiples, avec ou sans connecteur exprimé, nécessite que l'on poursuive des études dans la continuité des ébauches que nous avons présentées dans cet article.

Remerciements

Les auteurs souhaitent remercier les relecteurs dont le travail de relecture très détaillé et constructif a contribué à l'amélioration de cet article.

Références

- ABEILLE A., CLEMENT L., TOUSSENEL F. (2003). Building a treebank for French. In: A. ABEILLE (ed.), *Treebanks*, Dordrecht: Kluwer.
- AMSILI P., BEYSSADE C. (2010). Obligatory Presuppositions in Discourse. In: A. BENZ, P. KUEHNLEIN, Candace SIDNER (eds.), *Constraints in Discourse 2*, Amsterdam & Philadelphia: Benjamins Publishers, Pragmatics & Beyond, 105–123.
- AMSILI P., WINTERSTEIN G. (2012). Les déclencheurs de présupposition additifs. In: *Langages* 186 (2), 85–100.
- ANSCOMBRE J.C., DUCROT O. (1977). Deux *mais* en français. In: *Lingua* 43, 23–40.
- ASHER N. (1993). *Reference to abstract objects in discourse*. Dordrecht: Kluwer.
- ASHER N., HARDT D., BUSQUETS J. (2001). Discourse Parallelism, Ellipsis and Ambiguity. In: *Journal of Semantics* 18 (1).
- ASHER N., LASCARIDES A. (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.
- BIRD S., LOPER E., KLEIN E. (2009). *Natural Language Processing with Python*. Beijing: O'Reilly Media Inc.
- BLAKEMORE D. (2002). *Relevance and Linguistic Meaning. The semantics and pragmatics of discourse markers*. Cambridge: Cambridge University Press.

- BLAKEMORE D., CARSTON R. (2005). The pragmatics of sentential coordination with 'and'. In: *Lingua* 115, 569–589.
- DANLOS L. (2012). Formalisation des conditions d'emploi des connecteurs en réalité et (et) en effet. In: *Actes du 2ème Congrès Mondial de Linguistique Française*. Lyon, France, 493–508.
- ECKARD R., FRÄNKEL M. (2012). Particle, *Maximize Presupposition* and Discourse Management. In: *Lingua* 122 (15), 1801–1818.
- GROSZ B., SIDNER C. (1986). Attention, Intentions, and the structure of discourse. In: *Computational Linguistics* 12, 175–204.
- HOBBS J. (1985). *On the Coherence and Structure of Discourse*. Stanford University: Center for the Study of Language and Information. Report No. CSLI-85-37.
- JASINSKAJA K. (2012). Correction by Adversative and Additive markers. In: *Lingua* 122 (15), 1899–1918.
- JAYEZ J., WINTERSTEIN G. (2013). Additivity and Probability. In: *Lingua* 132, 85–102.
- LAKOFF R. (1971). If's, And's and Buts about conjunction. In: C. FILLMORE, D. LANGENDOEN (eds.), *Studies in Linguistic Semantics*, New York: de Gruyter, 114–149.
- LANG E. (1984). *The semantics of coordination*. Amsterdam: John Benjamins.
- MANN W., THOMPSON S. (1988). Rhetorical Structure Theory : Toward a functional theory of text organization. In: *Text* 8 (3), 243–281.
- POLANYI L. (1985). *A theory of discourse structure and discourse reference*, t. 21 de *CLS*. Chicago Linguistics Society.
- PRASAD R., DINESH N., LEE A., MILTSAKAKI E., ROBALDO L., JOSHI A., WEBBER B. (2008). The Penn Discourse Treebank 2.0. In: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)*. Marrakech, Morocco.
- ROZE C., DANLOS L., MULLER P. (2012). LEXCONN: A French Lexicon of Discourse Connectives. In: *Discours* 10.
- SPERBER D., WILSON D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.
- TXURRUKA I. (2003). The natural language conjunction 'and'. In: *Linguistics and Philosophy* 26, 255–285.
- WEBBER B., JOSHI A., STONE M., KNOTT A. (2003). Anaphora and discourse structure. In: *Computational Linguistics* 29 (4), 545–587.
- WINTERSTEIN G. (2012). What *but*-sentences argue for: a modern argumentative analysis of *but*. In: *Lingua* 122 (15) 1864–1885.
- (2013). The independence of quantity implicatures and adversative relations. In: *Lingua* 132, 67–84.