

# Un jeu de règles permettant de déterminer si une relation causale est exprimée entre des propositions<sup>1</sup>

Cécile Grivaz

Département de Linguistique  
Université de Genève  
cecile@grivaz.net

## Résumé

*Les jugements humains concernant la présence ou l'absence d'une relation causale dans un texte peuvent diverger, surtout en l'absence d'un marqueur causal explicite. Clarifier cette tâche peut permettre l'annotation d'un corpus qui serait utile à la fois à l'étude d'occurrences causales et à l'évaluation de programmes chargés de reconnaître ces occurrences. Dans cet article, nous présentons des expériences qui permettent d'établir un manuel d'annotation pour un tel corpus. Ces expériences nous ont permis d'obtenir le jeu de règles cohérent que nous présentons ici.*

*Mots clés : causalité, linguistique informatique, jugement humain, annotation.*

## 1. Introduction

La causalité a été largement étudiée dans des travaux précédents issus de nombreux champs de recherche tels que les sciences cognitives, la philosophie, la linguistique, la physique et le droit. Plus de deux siècles se sont écoulés depuis le travail fondateur de Hume (1739, 1740), et pourtant il reste de nombreuses questions quant à la nature même de la causalité. Il existe, pour ne citer qu'elle, une récente conférence virtuelle organisée par le site Interdisciplines.org (2005, 2006) qui traite du domaine de la causalité à partir des approches diverses de la philosophie et des sciences cognitives et sociales. En linguistique, la causalité a joué un grand rôle dans les travaux de divers auteurs tels que Talmy (1988) pour la sémantique ou Sanders (2005) pour la pragmatique.

Bien que la causalité ait été largement étudiée, il est difficile de trouver des tests consensuels capables d'identifier facilement un

---

<sup>1</sup> Article rédigé dans le cadre du projet FNSRS 100012-113382 *Pragmatique lexicale et non lexicale de la causalité en français : aspects descriptifs, théoriques et expérimentaux*, et dans le cadre d'une thèse de l'Université de Genève sous la direction de J. Moeschler et M. Rajman.

passage de texte comme représentant une relation causale, et, comme nous le verrons plus tard, les jugements humains divergent pour de nombreuses occurrences. Par exemple, en l'absence de contexte, il est extrêmement difficile de déterminer si le passage (1) représente une relation causale. Le simple fait de reformuler la phrase avec un marqueur causal explicite comme dans (2) n'aide pas à désambigüiser, car il reste difficile de déterminer si la phrase (2) est synonyme de (1) ou non. Même si l'on utilise un exemple moins extrême comme (3), le jugement reste difficile, car on peut considérer que la cause de *se casser la jambe* n'est pas de skier mais un événement qui survient dans ce contexte tel qu'une collision ou une chute :

- (1) L'oreille de Jean est tombée, il a pris une douche.
- (2) La chute de l'oreille de Jean a été causée par le fait qu'il a pris une douche.
- (3) Il s'est cassé la jambe en skiant.

Il n'existe pas à notre connaissance de corpus français annoté avec des relations causales. Un tel corpus serait utile à des linguistes pour l'étude d'occurrences de relations causales. Mais nous nous plaçons surtout dans une optique de traitement automatique de la langue où une définition précise de ce qui est causal ou non permet d'évaluer un système capable de réaliser une telle annotation automatiquement.

Dans ce travail, nous décrivons une méthode qui permet de mieux définir le concept de causalité à partir de ses caractéristiques. Nous développons un jeu de règles qui transmet à des annotateurs notre vision raisonnée de la causalité et leur permet d'annoter précisément un corpus avec des relations causales. En particulier, nos consignes résolvent l'ambiguïté dans de nombreux cas difficiles. Ce travail était nécessaire, car il n'existe pas à notre connaissance de conditions nécessaires et suffisantes facilement utilisables pour déterminer si un passage de texte est causal ou non.

Notre manuel doit aider les annotateurs à identifier de façon cohérente des relations causales entre propositions même en l'absence de marqueur explicite tel que *parce que* ou *donc*. Nous présentons les résultats suivants :

- Nous décrivons une expérience qui permet d'éliciter des caractéristiques de la causalité utilisées consciemment pour le raisonnement (section 3).
- Nous décrivons une méthodologie pour affiner et tester nos premières règles grossières et nous détaillons les règles auxquelles nous avons abouti (section 4).
- Nous montrons que nos consignes d'annotations sont cohérentes puisqu'elles permettent un très bon accord entre la

majorité des annotateurs et nos prédictions ( $\kappa=0.84$ ) et qu'elles permettent bien d'identifier la causalité puisqu'elles satisfont les experts (4.7).

## 2. Approche générale

Notre tâche est de créer des consignes d'annotation qui soient suffisamment claires pour permettre à des annotateurs d'identifier de façon cohérente les occurrences de passages qui représentent une relation causale. Par exemple, étant donné un passage qui contient la phrase *Le verre est cassé, il est tombé de la table*, nous voulons qu'un annotateur puisse identifier qu'il s'agit de l'expression d'une relation causale, que la cause est représentée par *il est tombé de la table* et que l'effet est représenté par *le verre est cassé*. La cause et l'effet sont des éventualités qui peuvent être représentées soit par des propositions, comme ici, soit par des groupes nominaux comme *un glissement de terrain*. Nous ne nous intéressons ici qu'aux premières.

Hovy et ses collègues (2006) ont réalisé un travail très similaire à celui que nous présentons ici. Ils ont annoté un corpus en associant à chaque mot son concept dans une ontologie. Cette tâche de désambiguïsation du sens de chaque mot est difficile et peut déboucher sur un accord inter-annotateurs bas. Pour pallier cette difficulté, Hovy et ses collègues ont modifié et clarifié la tâche et testé leurs consignes sur un petit échantillon jusqu'à obtenir un accord inter-annotateurs de plus de 90%.

Nous avons réalisé un travail proche de celui d'Hovy et de ses collègues pour la tâche qui nous occupe. Nous avons écrit un premier jeu de consignes que nous avons ensuite affiné en nous basant sur les divergences entre annotateurs avec ce premier manuel. Nous avons écrit ce premier jeu en nous inspirant d'une part de travaux théoriques précédents et d'autre part de caractéristiques intuitives de la causalité. Nous décrivons comment nous avons obtenu ces caractéristiques intuitives dans la section 3. Ce processus nous a permis d'une part d'affiner et de formaliser notre compréhension de la causalité, et d'autre part d'obtenir un manuel d'annotation cohérent. Nous discutons plus en détail la cohérence de ces règles dans la section 4.7.

## 3. Caractéristiques intuitives

Nous décrivons maintenant une première expérience qui nous a permis de découvrir des caractéristiques intuitives de la causalité. Nous voulions savoir si les gens utilisaient certaines caractéristiques ou certains tests pour déterminer s'ils avaient affaire à une relation causale ou non. Nous voulions vérifier si ces caractéristiques rejoignaient celles décrites dans les travaux théoriques et surtout nous voulions disposer de ces caractéristiques pour nos consignes

d'annotation, soit si elles satisfaisaient des experts de la causalité, pour les intégrer à nos consignes, soit si elles ne permettaient pas de bien définir la causalité, pour en avertir explicitement les annotateurs qui pourraient être tentés de les utiliser.

L'expérience permettant d'éliciter des caractéristiques ou des tests de la causalité est la suivante. Nous avons demandé à des sujets de déterminer si des passages de texte en français contenaient l'expression d'une relation causale en justifiant systématiquement leur réponse, qu'elle soit positive ou négative. Ces justifications contiennent des caractéristiques qui sont utilisées pour le raisonnement causal.

Nous avons utilisé des passages constitués de quelques phrases tirées de divers textes du corpus BAF (RALI laboratory 1997). Ces passages contiennent les connecteurs *donc* ou *parce que* qui sont souvent causaux, ou le connecteur *mais*, typiquement non causal. Les passages sont donc de trois types en fonction du connecteur qu'ils contiennent. Nous avons soumis certains de ces passages tels quels aux sujets, et nous avons enlevé le connecteur de certains autres passages afin de les rendre moins clairs et de réduire le nombre de justifications se rapportant à la présence de connecteurs. Nous avons remplacé les connecteurs par une virgule lorsque c'était nécessaire. Chaque sujet devait analyser des passages de chacun des six types résultants : pour chacun des trois connecteurs, des passages contenant le connecteur et des passages dont on l'avait enlevé.

Nous avons fait passer cette expérience à 9 étudiants en linguistique qui n'avaient pas encore étudié la causalité et parlaient le français couramment. Chaque étudiant devait analyser 10 passages dans un ordre aléatoire différent. L'analyse prend environ 45 minutes.

La plupart des passages (72.2%) ont été jugés causaux, ce qui n'est pas surprenant, car 80.0% d'entre eux contiennent un connecteur causal ou en ont contenu un qui a été supprimé.

Nous avons groupé les justifications en 5 classes : reformulation, test linguistique, présence d'un marqueur explicite, présence d'une relation non causale et autre. Nous appelons *reformulations* des justifications qui n'apportent pas d'information supplémentaire sur le raisonnement causal et qui sont des reformulations de la consigne. Ce sont des réponses du type « une explication est donnée » pour les réponses positives ou du type « Je ne vois pas de relation causale » pour des réponses négatives. Ce type de justification est très courant pour les réponses positives (37.5%) et le plus courant pour les réponses négatives (41.4%).

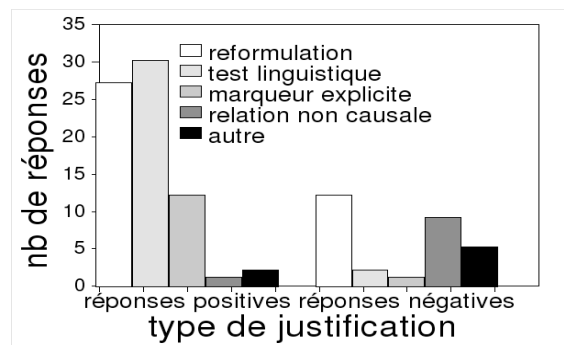


Figure 1 : Types de justifications pour les intuitions causales. Ce graphique montre la quantité de justifications de chaque type pour les passages considérés comme causaux à gauche et comme non causaux à droite

Nous appelons *tests linguistiques* les justifications dans lesquelles le sujet place un marqueur causal explicite dans le passage qu'il analyse et constate que le résultat est une phrase valide. Ces marqueurs explicites peuvent être des connecteurs comme *parce que* ou *donc* (dans ce cas, les sujets ont souvent remplacé les connecteurs qui avaient été supprimés du texte). Il peut aussi s'agir de prédicats comme *est le fruit de*, *entraîne*, *permet* ou *est la cause de*, en quel cas les éventualités exprimées sous forme de propositions doivent être nominalisées, ce qui est généralement fait avec le groupe *le fait que*. La différence entre ce type de justification et la reformulation est qu'ici le texte à analyser est repris au moins en partie dans la justification. Ce type de justification est le plus courant pour les réponses positives (40.0%) et ne représente que 6.9% des réponses négatives.

Un autre type de justification est celui qui consiste à indiquer la présence d'un *marqueur explicite* présent dans le texte pour justifier qu'il s'y trouve une relation causale ou à relever la présence d'un marqueur non causal pour justifier l'absence de relation causale. On trouve par exemple « *Donc* apparaît. *Donc* introduit une conséquence » ou « Le *mais* exprime une nuance, une restriction dans ce cas précis ». Ce type de justification représente 16.7% des réponses positives et 3.4% des réponses négatives, alors que 30% des passages contenaient un connecteur (20% causal et 10% non causal).

Parfois les sujets ont analysé le texte et y ont détecté la présence d'une relation non causale. Ils ont alors justifié leur réponse négative par la présence de cette autre relation. On trouve par exemple « c'est une description, la seconde phrase apporte seulement une précision ».

Cette catégorie représente 31.0% des réponses négatives. Elle est mentionnée pour une réponse positive, mais dans ce cas, le sujet a changé d'avis après avoir donné la justification « La seconde proposition détaille la première mais n'introduit pas une cause liée par un connecteur. Mais si on met *parce que* entre les deux ça peut jouer. »

Finalement, quelques justifications n'entrent dans aucune de ces catégories. Il s'agit principalement de points d'interrogation et d'une justification qu'il nous a été impossible de comprendre.

La figure 1 montre le nombre de réponses positives ou négatives pour la causalité contenant une justification de chaque type. Les réponses contenant plusieurs justifications contribuent à chacune de ces justifications. Le type de justification le plus fréquent pour les réponses positives est le test linguistique suivi de près par la reformulation. Pour les réponses négatives, c'est la reformulation qui est la plus présente. Nous pensons que cette forte représentation de la reformulation indique que les sujets ont peu d'accès à leur raisonnement causal. Ils savent que quelque chose est causal ou non, mais ils ne savent pas comment ils le savent.

On peut s'étonner de ne pas trouver de raisonnements se rapportant aux éventualités elles-mêmes, mais uniquement des justifications de type linguistique. En particulier, nous nous attendions à trouver la contrefactualité comme justification : si la cause ne s'était pas passée, alors la conséquence ne se serait pas passée non plus. Nous attribuons ce phénomène au contexte de l'expérience. En effet, il s'agit d'une expérience de linguistique, sur du texte et avec des étudiants en linguistique. Il serait intéressant de savoir si le phénomène se reproduit dans une expérience présentée comme psychologique, avec de la vidéo et des ingénieurs.

Tous ces tests satisfont des experts de la causalité, ils représentent bien des caractéristiques de la causalité. Leur intégration au manuel d'annotation se fait de la façon suivante. La reformulation se traduit par l'intuition causale, qui est le test le plus important. Comme nous le verrons plus tard, les autres tests servent surtout à clarifier l'intuition ou à éliminer des cas non causaux. Les tests linguistiques sont utilisés directement. Nous ne nous sommes pas intéressés aux marqueurs explicites parce que notre but était surtout de pouvoir désambiguïser les cas implicites. Nous n'avons pas non plus demandé aux annotateurs d'identifier s'il y avait une relation non causale, parce que nous pensons qu'identifier précisément une autre relation était une perte de temps, quand il suffisait d'identifier qu'il ne s'agissait pas d'une relation causale et parce que certaines relations non causales comme la relation temporelle ne sont pas incompatibles avec

une relation causale.

#### 4. Les consignes d'annotation

Dans cette section, nous présentons deux expériences. La première nous a permis de tester et d'affiner nos premières consignes. La seconde nous a permis de tester nos consignes finales. Nous décrivons les consignes que nous avons obtenues ainsi que la façon dont nous avons dû modifier le jeu initial pour obtenir des consignes plus claires et plus cohérentes.

Les consignes consistent en un jeu de caractéristiques que l'annotateur doit identifier. Certaines caractéristiques sont nécessaires à la causalité, certaines permettent d'éliminer des cas non causaux et certaines permettent d'aider à désambiguïser quelques cas. Comme il n'y a pas de conditions nécessaires et suffisantes pour identifier les expressions causales, l'annotateur devra toujours se fier à son intuition pour décider, en définitive, si le cas est causal ou non. Les consignes permettent d'augmenter l'accord sur certains cas ambigus et doivent aider à guider l'intuition de l'annotateur. Nous reviendrons plus en détails sur ces consignes dans les paragraphes suivants.

Nous nous sommes basée sur deux critères pour affiner le jeu de consignes initial : les divergences entre des experts de la causalité et les divergences entre nos prédictions et les annotations de la majorité des annotateurs. Nous mesurons l'accord inter-annotateurs grâce au score  $\kappa$  introduit par Cohen (1960).

Nous testons les consignes à l'aide d'un programme *ad hoc* qui présente des énoncés et permet aux annotateurs d'identifier certaines caractéristiques de la causalité et de déterminer si l'énoncé est causal ou non. Les énoncés sont présentés avec la cause potentielle et l'effet potentiel surlignés, de sorte que l'utilisateur doit uniquement identifier si l'énoncé est causal ou non, mais ne doit pas identifier les segments qui représentent la cause ou la conséquence.

Les sujets de la première expérience étaient des étudiants en linguistique qui n'avaient pas encore étudié la causalité et qui parlaient le français couramment. Nous leur avons demandé de décider si 24 énoncés représentaient des relations causales ou non. Les énoncés étaient des phrases artificielles ambiguës causalement. La figure 2 représente un exemple de consigne avec un énoncé.



Merci d'utiliser CARMA, vous êtes Cecile.

**changer d'identifiant**

Lisez, s'il vous plaît le segment de texte suivant. Ce segment représente peut-être une relation causale. La cause potentielle est indiquée en jaune et la conséquence potentielle en noir. Cochez, s'il vous plaît, toutes les affirmations qui vous semblent correspondre au texte. Pour cocher ou non la dernière affirmation ("vous pensez qu'une relation causale est exprimée") basez-vous, s'il vous plaît, sur les caractéristiques que vous venez de cocher ou non.

**Paul a pris ses médicaments. Il va guérir.**

- Paul a pris ses médicaments se passe avant il va guérir.
- Paul a pris ses médicaments se passe immédiatement avant il va guérir.
- Paul a pris ses médicaments et il va guérir, se passent au moins en partie simultanément.
- On peut construire une chaîne causale entre Paul a pris ses médicaments et il va guérir.
- On peut dire il va guérir parce que Paul a pris ses médicaments.
- On peut dire Paul a pris ses médicaments donc il va guérir.
- Si Paul a pris ses médicaments ne s'était pas passé, alors il va guérir ne se serait pas passé non plus
- Paul a pris ses médicaments et il va guérir sont deux paraphrases qui représentent le même événement
- Je pense que dans ce texte, il est exprimé que Paul a pris ses médicaments est la cause de il va guérir.

**Envoyer**

Figure 2 : Consigne avec un énoncé pour la première expérience

L'accord inter-annotateurs moyen entre chaque paire d'annotateurs était un très médiocre  $\kappa=0.3$  pour la causalité. Le score  $\kappa$  entre nos annotations et celles de la majorité des étudiants est de  $\kappa=0.43$ . Nous avons calculé ce score en ignorant les exemples qui sont ambigus à cause d'un manque de contexte et qui ne pourraient être interprétés comme causaux ou non qu'en tenant compte de plus d'informations contextuelles, comme, par exemple :

(4) Après qu'il a fait un long voyage au Canada, il s'est marié.

Nous avons également ignoré les énoncés qui n'avaient pas permis d'obtenir une réponse de la majorité des annotateurs (des énoncés que deux annotateurs avaient considérés comme causaux et que deux autres annotateurs avaient considérés comme non causaux). Ce score de 0.43 est médiocre mais nous a permis de clarifier nos consignes.

Nous avons également fait passer la même expérience à deux experts de la causalité. Nous avons obtenu un accord inter-annotateurs médiocre de  $\kappa=0.32$ . Néanmoins nous avons pu utiliser les divergences entre les experts pour clarifier la notion de causalité.

Nous nous sommes servie des résultats de cette première expérience pour affiner et clarifier nos consignes. Nous avons également modifié l'interface du programme d'annotation, en particulier nous avons ajouté du contexte aux énoncés à annoter et nous avons demandé aux annotateurs d'identifier moins de caractéristiques ; nous leur avons seulement conseillé d'en utiliser certaines pour désambiguïser les cas difficiles. De façon générale, nous avons voulu simplifier le processus afin de ne pas faire perdre de temps aux annotateurs pour identifier des caractéristiques quand la causalité était suffisamment claire.



Nous allons maintenant présenter les nouvelles consignes qui permettent d'obtenir un bon  $\kappa$  entre la majorité des annotateurs et nos prédictions. Dans la section 4.7, nous discuterons des résultats du test de ces consignes.

#### 4.1. Tests linguistiques

Dans la première expérience, les résultats des tests linguistiques avec *parce que* et *donc* étaient corrélés avec l'intuition des annotateurs. En effet, dans 84% des réponses positives pour la causalité, les annotateurs avaient également considéré que le test avec *parce que* était positif. Ils avaient considéré que le test de *parce que* fonctionnait dans 13% des cas qu'ils considéraient comme non causaux. De même, une réponse positive avec le test de *donc* est associée à 64% des réponses causales et à 11% des réponses non causales.

Ces tests étaient également très présents dans les justifications de l'expérience qui permettait de trouver des caractéristiques intuitives de la causalité, ce qui indique qu'ils sont intuitifs.

Le problème de *parce que* et de *donc* est qu'ils sont causalement ambigus.

En effet, *parce que* et *donc* peuvent avoir des usages d'acte de langage comme dans *Dépêche-toi, parce qu'on va être en retard* où la seconde phrase n'est pas la cause de la première, mais de son énonciation. Nous ne voulions pas que les annotateurs résolvent ce type d'ellipse, mais qu'une phrase comme la précédente soit annotée comme non causale. De même dans *George est sorti parce que sa veste n'est pas sur la chaise*, la seconde phrase n'est pas la cause de la première mais de la croyance de celle-ci. Nous ne voulions pas que *la veste de George n'est pas sur la chaise* soit annoté comme la cause de *George est sorti*. Néanmoins, dans ce cas, la cause de la croyance est la conséquence de ce qui est cru. En effet, *George est sorti* est la cause de *sa veste n'est pas sur la chaise*. Nous voulions donc que les annotateurs annotent cette phrase comme causale, mais dans l'ordre inverse de celui (conséquence-cause) que demande habituellement le connecteur *parce que*.

Dans les dernières consignes, nous avons détaillé les usages épistémiques et d'acte de langage de *parce que* et de *donc*, dans le manuel d'annotation. Nous avons demandé aux annotateurs de systématiquement indiquer si le passage qu'ils analysaient était positif pour les tests linguistiques. En particulier, nous leur avons demandé d'essayer d'insérer *parce que* et *donc*, et, en accommodant les éventuels problèmes syntaxiques, de décider si la phrase résultante est sémantiquement correcte. Nous avons demandé d'annoter systématiquement le test avec *c'est parce que*. En effet, *c'est parce que*,

n'a pas, à notre connaissance, d'usage épistémique ou d'acte de langage. Si nous avons demandé cette annotation systématique des tests linguistiques, c'est parce que nous estimons qu'ils permettent de réaliser très rapidement un test discriminatif et intuitif de la causalité.

#### 4.2. *Contrefactualité*

La propriété contrefactuelle est une propriété très importante de la causalité (voir par exemple Reboul 2005). La contrefactualité est le fait que si la cause ne s'était pas passée, alors la conséquence n'aurait pas eu lieu non plus. C'est une des propriétés qui différencient la causalité du *si ... alors* logique.

Cette propriété permet d'éliminer facilement certains exemples non causaux comme *Mon bus va bientôt partir, je viens de finir mon petit déjeuner*. En effet, ces événements ne sont pas liés causalement, car le bus partirait bientôt même si je ne venais pas de finir mon petit déjeuner. C'est pourquoi nous avons demandé aux annotateurs d'identifier systématiquement cette propriété dans la première expérience. Bien qu'elle ait été assez discriminative (associée à 78% des réponses positives et 16% des réponses négatives), elle n'est pas associée à toutes les réponses positives. En effet, par exemple avec la phrase *Jean s'est cassé la jambe en skiant*, certains annotateurs ont considéré que la propriété contrefactuelle ne s'appliquait pas, alors même qu'ils considéraient la phrase comme causale. Ils argumentaient que s'il n'avait pas skié, Jean aurait pu quand même se casser la jambe, par exemple en glissant sur une peau de banane. C'est pourquoi nous avons présenté et commenté cet exemple dans le nouveau manuel d'annotation. Nous avons également ajouté *probablement* dans la formulation de la propriété contrefactuelle, dans l'expérience suivante : *si la cause ne s'était pas passée, la conséquence ne se serait probablement pas passée non plus*.

Nous avons demandé aux annotateurs d'identifier systématiquement la propriété contrefactuelle dans les passages qu'ils annotaient, parce que nous pensons qu'elle permet de rapidement discriminer certains cas non causaux et qu'elle peut être identifiée rapidement.

#### 4.3. *Asymétrie temporelle*

L'asymétrie temporelle est une caractéristique habituelle de la causalité. Elle a été identifiée par Hume (1739, 1740). Dans beaucoup de cas de causalité, comme par exemple *Jean est tombé, Max l'a poussé* la cause arrive avant l'effet. Dans les cas de causalité directe comme *le verre atteignit le sol et se cassa*, la cause précède immédiatement l'effet. Il y a également des cas où la cause et l'effet arrivent au moins en partie simultanément comme dans *il s'est fatigué en conduisant*.

Dans la première expérience, nous demandions aux annotateurs d'identifier précisément l'ordre temporel des éventualités qu'ils analysaient. Nous leur demandions de déterminer si la cause potentielle arrivait avant, immédiatement avant ou en partie simultanément avec la conséquence. Nous voulions ainsi clarifier leur intuition par une analyse minutieuse de la temporalité. Néanmoins cette analyse prenait du temps et s'est avérée peu utile. C'est pourquoi dans la nouvelle expérience nous leur avons seulement demandé d'annoter comme non causaux des cas dans lesquels la cause se passait après l'effet. En effet, il est bien plus facile de discriminer des cas où la cause potentielle se passerait après l'effet potentiel, que d'analyser finement les autres cas.

#### 4.4. Asymétrie

Dans la première expérience, nous avons demandé aux annotateurs d'identifier des couples d'éventualités qui potentiellement causaux seraient des paraphrases l'une de l'autre. Nous voulions éliminer des cas comme *c'est un triangle, il a trois côtés*. Cette propriété a été très peu annotée (14% de tous les exemples). Nous l'avons donc généralisée dans le nouveau manuel : c'est la propriété d'asymétrie.

En effet les relations causales sont asymétriques (Hume 1739, 1740) pas seulement temporellement mais également causalement. Si une éventualité est la cause d'une autre éventualité, la seconde ne peut que dans des cas anecdotiques être la cause de la première. Cette propriété est présente même dans les cas où la cause et l'effet sont simultanés. Par exemple, dans *Jean se fatigue en conduisant*, le fait que Jean conduise est la cause de sa fatigue, mais sa fatigue ne peut pas être la cause du fait qu'il conduise. Il est possible d'imaginer des cas de causalité circulaire, mais nous pensons que ces cas sont rares et que les deux sens de la causalité seraient alors indiqués de façon séparée dans le texte en raison de la qualité inhabituelle du phénomène. Si un tel cas devait se produire, alors les deux occurrences de relations causales devraient être annotées séparément et chacune seulement dans le sens qu'elle exprime comme dans (5) :

- (5) Elle souriait parce qu'elle le voyait sourire et plus elle souriait, plus il souriait.

Dans notre seconde expérience nous avons demandé aux annotateurs, pour chaque paire d'éventualités, d'indiquer si le texte les présentait dans l'ordre cause-conséquence ou conséquence-cause. Nous leur avons demandé d'annoter comme non causaux les exemples pour lesquels il était difficile de choisir l'ordre ou pour lesquels les deux ordres semblaient possibles. En particulier, nous voulions éviter que des annotateurs considèrent comme causale une

phrase comme *c'est un triangle, il a trois côtés*, qui avait largement été annotée comme causale lors de l'expérience précédente, mais qui ne correspond pas à notre vision de la causalité.

#### 4.5. Chaînes causales

Une chaîne causale est une chaîne formée de causes et d'effets directs qui peut être associée à une relation causale (étudiée par exemple dans Moeschler 2003). Par exemple, pour la relation causale *Jean est tombé, Max l'a poussé*, on peut former la chaîne causale suivante :

(6) Max pousse Jean → Jean est déséquilibré → Jean tombe.

Dans la première expérience, nous avons demandé aux annotateurs d'identifier systématiquement si une chaîne causale pouvait être formée entre les éventualités qu'ils analysaient dans l'espoir qu'identifier les suites de causes et de conséquences directes les aide à clarifier leur intuition. La première expérience a démontré que cette caractéristique était très discriminative puisqu'elle était associée à 82% des réponses positives et à seulement 4% des réponses négatives. Néanmoins, nous ne savons pas si elle a permis de clarifier l'intuition des annotateurs ou s'il leur était nécessaire d'avoir une intuition sur le caractère causal du passage avant de pouvoir former la chaîne causale. S'il est nécessaire d'avoir une intuition sur la causalité du passage pour former la chaîne et si cette intuition n'est pas remise en cause par cet exercice, alors l'identification d'une chaîne causale est redondante avec l'intuition et constitue une perte de temps pour les annotateurs d'autant plus grande que le processus est relativement coûteux, puisqu'il faut imaginer un certain nombre d'éventualités causes et conséquences entre les éventualités que l'on cherche à analyser.

C'est pourquoi, nous n'avons pas demandé cette identification systématique dans la nouvelle expérience. Nous l'avons néanmoins proposée comme une façon de clarifier l'intuition sur des cas ambigus, en particulier si les éventualités analysées étaient éloignées temporellement.

#### 4.6. Autres règles d'annotation

Nous avons encore dû clarifier deux cas dans le manuel d'annotation. D'abord, des phrases comme *il s'est cassé la jambe en skiant* ont généré des désaccords entre les experts de la causalité. Ce problème est dû au fait que *skier* n'est en fait pas la cause de *se casser la jambe*. La cause de *se casser la jambe* est un événement non précisé dans le contexte de *skier*, tel que *tomber* ou *subir une collision*. Considérer ou non le contexte de l'éventualité cause comme causal est un choix arbitraire. Notre manuel prescrit d'annoter de tels contextes comme causaux,

afin de définir une acceptation précise mais large de la notion de causalité.

Des discussions avec les experts indiquent que les jugements causaux deviennent plus difficiles si l'éventualité potentiellement causale est une éventualité négative comme dans *Comme aucune étude semblable n'a été faite auparavant, ce rapport sera fort pertinent* ou que sa classe aspectuelle est celle d'un état comme dans *Marjorie est majeure, elle peut voter*. Nous prescrivons, dans des cas semblables, de chercher à clarifier l'intuition en construisant des chaînes causales ou en essayant de remonter à une loi générale comme *si on est majeur, on peut voter*. Notre manuel prescrit également, dans ces cas-là, une vérification systématique de l'asymétrie, par exemple avec le test de *c'est parce que*. En effet *C'est parce que Marjorie est majeure qu'elle peut voter* est parfait alors que *C'est parce qu'elle peut voter que Marjorie est majeure* est moins bon.

#### 4.7. Test du nouveau manuel

Nous présentons maintenant les résultats d'une expérience qui permet de tester notre manuel d'annotation. Pour cette expérience, nous avons changé l'interface du programme d'annotation. La nouvelle interface permet de présenter plus de contexte pour les passages à analyser et le nombre de caractéristiques à annoter systématiquement est réduit par rapport à la première expérience, comme on l'a vu précédemment. Une autre différence est que la cause potentielle et l'effet potentiel ne sont pas marqués différemment, et qu'au contraire, l'annotateur doit décider, s'il s'agit bien d'une relation causale, si le texte présente les éventualités dans l'ordre cause-conséquence ou conséquence-cause. Cette modification permet de diminuer le nombre de passages à tester (avec l'ancienne interface, tester les deux sens aurait nécessité de présenter le passage deux fois à l'annotateur, avec les éventualités conséquence et cause inversées). Ce changement permet aussi d'attirer l'attention des annotateurs sur le problème de l'asymétrie. Finalement, nous avons ajouté la possibilité de répondre *ne sait pas* à la question de la causalité, ainsi qu'aux questions portant sur la caractéristique à identifier, alors que dans la première expérience, les annotateurs étaient forcés de choisir une réponse.

Les 15 passages à analyser sont tirés du roman de Jules Verne *De la terre à la lune*. Chaque passage est constitué d'un paragraphe et contient deux propositions mises en évidence qui représentent les éventualités à analyser pour déterminer si elles sont en relation causale. Nous avons éliminé les passages les plus évidemment non causaux, afin de faire gagner du temps aux annotateurs et d'obtenir des résultats plus significatifs. Chacun des 15 passages a été analysé

par 4 étudiants en linguistique qui n'avaient pas étudié la causalité et qui parlaient le français couramment.

Pour cette expérience et pour la causalité, nous avons obtenu un score  $\kappa$  moyen entre les annotateurs de 0.38. L'amélioration du score  $\kappa$  est donc très faible par rapport à l'expérience précédente pour laquelle il était de 0.30. Cette légère amélioration peut être expliquée par le fait que la nouvelle expérience utilise des passages extraits d'un texte existant et non des énoncés délibérément ambigus causalement.

Cette expérience présente néanmoins une très nette amélioration du score  $\kappa$  entre nos annotations et celles de la majorité des annotateurs. En effet le nouveau score est de 0.84 contre 0.43 précédemment. Comme précédemment, ce score ignore les passages pour lesquels il n'y a pas eu de majorité pour les réponses des annotateurs (il s'agit de deux des quinze énoncés). Parmi les treize exemples à majorité claire, la majorité diffère de nos annotations seulement une fois où 3 des annotateurs considéraient comme causal un passage qui nous semblait non causal (le quatrième le considérait comme non causal).

Nous pensons que cet accord très haut entre nos annotations et celles de la majorité des annotateurs indique que nous avons pu transmettre notre vision de la causalité de façon claire et cohérente. Nous avons pu traduire par notre manuel notre compréhension de la causalité avec suffisamment de clarté pour que les annotateurs puissent obtenir des résultats proches des nôtres. Nous avons donc développé une méthodologie claire et cohérente pour identifier un type précis de relation dans un texte. L'accord des experts et le fait que le manuel est basé sur des travaux théoriques nous permettent d'affirmer que ces relations sont bien des relations causales.

## **5. Rapport avec les travaux précédents**

Nous discutons ici différents travaux précédents qui ont influencé le nôtre. A notre connaissance, il n'existe pas de corpus français annoté avec des relations causales. Il existe cependant, en anglais, un corpus annoté avec les relations de la Théorie des Structures Rhétoriques (RST). Ces relations comprennent des relations causales dont nous discuterons plus en détail dans la section 5.1.

Inui (2005) a réalisé en japonais un travail proche du nôtre en annotant un corpus avec des relations causales. Nous discuterons de ce travail dans la section 5.2. Finalement, dans la section 5.3, nous discuterons des rapports entre notre méthodologie et celle de Hovy et al. (2006) qui a inspiré ce travail.

### 5.1. Carlson et ses collègues

Il existe en anglais un important corpus annoté avec des arbres syntaxiques (Marcus et al. 1994). C'est ce corpus que Carlson et ses collègues (2002) ont annoté avec des relations issues de la Théorie des Structures Rhétoriques (RST). Une de ces relations est la relation *cause* qui est subdivisée en trois sous-relations : *cause*, *result (résultat)* et *consequence (conséquence)*. La différence entre les relations *cause* et *result* est l'importance donnée dans le texte soit à l'effet soit à la cause. La relation *consequence* se distingue des autres relations par le fait qu'elle représente un lien causal moins direct. Nous considérons la relation annotée par Carlson et ses collègues *reason (raison)* comme un type de relation causale. Elle se différencie des autres en cela qu'ici la conséquence est réalisée par un agent animé. Les instructions d'annotation de Carlson et de ses collègues ne donnent pas de définition du concept de cause. Néanmoins, il est demandé aux annotateurs de toujours choisir la relation la moins générale possible de sorte que, par exemple, une relation causale et temporelle devra être annotée comme causale (plus précis) et non pas comme temporelle (plus général).

En comparaison avec le travail de Carlson et de ses collègues, notre travail est plus spécifique. En effet nous ne cherchons à identifier précisément que la relation causale. Nous ne cherchons pas non plus à analyser les relations causales plus finement en les subdivisant en sous-catégories, et nous ne faisons pas, par exemple, de différence entre *cause* et *raison*. Mais nous cherchons à obtenir un manuel beaucoup plus détaillé permettant de discriminer très précisément ce qui est causal de ce qui ne l'est pas.

### 5.2. Inui

Inui (2005) a annoté un corpus japonais avec des relations causales. Son travail est plus proche du nôtre que celui de Carlson et de ses collègues en cela qu'il n'a annoté que les relations causales. Le corpus d'Inui est constitué de 750 articles de journaux qui traitent du domaine social. Les relations causales qu'il annoté peuvent être signalées ou non par un marqueur causal. Inui n'a pas seulement annoté les relations causales entre propositions mais également entre groupes nominaux, comme dans *le tremblement de terre a provoqué un glissement de terrain*. Pour identifier précisément ces relations, Inui s'est servi d'un ensemble de tests. Les annotateurs devaient essayer les différents tests linguistiques sur les passages qu'ils analysaient. Si l'un des tests donnait une phrase sémantiquement et syntaxiquement correcte, le passage était annoté comme causal. Sinon, les annotateurs devaient essayer les autres tests, et considérer comme non causaux les

passages qui ne passaient aucun des tests linguistiques. En plus d'annoter les passages comme causaux ou non, Inui a ajouté une étiquette qui représente le degré de nécessité de la relation causale annotée. Cette étiquette indique si la relation causale entre les deux éventualités existe habituellement ou non.

Notre travail est différent de celui de Inui sur trois points. D'abord, nous ne nous sommes intéressée qu'aux relations causales entre des éventualités représentées par des propositions et non des groupes nominaux. Ensuite, nous avons seulement cherché à annoter la causalité, sans nous intéresser à son degré de nécessité. Enfin, nous nous sommes intéressée plus précisément aux consignes d'annotations pour lesquelles nous n'avons pas seulement utilisé des tests linguistiques qui peuvent être ambigus, mais également d'autres caractéristiques de la causalité. Il faut néanmoins noter qu'il est possible que les tests linguistiques soient moins ambigus en japonais qu'en français.

### *5.3. Hovy et ses collègues*

Hovy et ses collègues (2006) ont réalisé une grande tâche d'annotation de désambiguïsation de sens de mots sur le corpus de la Penn TreeBank (Marcus et al. 1994). Il s'agit de lier chaque mot à son sens dans une ontologie. Cette tâche peut donner un accord inter-annotateurs bas. Hovy et ses collègues s'étaient donné pour but d'obtenir un accord inter-annotateurs supérieur à 90%. Ils ont donc utilisé un système itératif dans lequel les consignes étaient testées sur des échantillons de 50 phrases et modifiées jusqu'à obtenir un accord haut. Les modifications dans les consignes n'étaient pas seulement des clarifications, comme c'est le cas dans notre travail. Il s'agissait aussi de définir les groupes de sens optimaux auxquels rattacher les mots. Ainsi, à chaque itération, les consignes étaient clarifiées, et les groupes de sens étaient modifiés jusqu'à ce que l'accord soit suffisamment haut. A ce moment, Hovy et ses collègues ont fait coïncider leurs groupes de sens à des sens dans une ontologie. Ce sont ensuite ces sens qui ont été annotés.

Une grande différence entre ce travail et le nôtre est que nous n'avons pas adapté la tâche, nous avons seulement cherché à la clarifier. Ceci explique peut être pourquoi nous n'obtenons pas un bon accord inter-annotateurs malgré nos efforts. Nous sommes convaincue qu'en augmentant les itérations et en clarifiant encore le manuel, notre méthodologie, permettrait d'obtenir un accord haut. Néanmoins, nous pensons que le nombre d'itérations nécessaires est élevé parce que nous ne voulons pas adapter la tâche et que nous voulons satisfaire des experts de la causalité. C'est pourquoi nous



nous contentons pour l'instant d'un manuel qui permet de transmettre notre vision de la causalité à la majorité des annotateurs.

## **6. Conclusion et perspectives**

Dans ce travail, nous avons décrit une première expérience qui visait à découvrir des caractéristiques intuitives de la causalité. Nous avons utilisé ces caractéristiques ainsi que des travaux théoriques et l'avis d'experts pour écrire des consignes d'annotation de la causalité que nous avons ensuite testées et affinées. Nous avons présenté en détail les consignes qui permettent à la majorité des annotateurs d'imiter notre jugement raisonné. Nous avons montré que les consignes sont cohérentes puisqu'elles permettent un accord haut entre la majorité des annotateurs et nos prédictions et qu'elles représentent bien la causalité et pas une autre relation puisqu'elles sont validées par des experts.

Nous croyons que notre méthodologie, qui est largement inspirée de celle de Hovy et de ses collègues (2006), est applicable à de nombreuses tâches qui ne permettent pas immédiatement un accord inter-annotateurs élevé. Nous croyons également que le fait de chercher d'abord des caractéristiques intuitives dans de telles tâches est utile à la compréhension des processus de raisonnement qui leur sont liés, et surtout à l'élaboration de consignes d'annotation qui tiennent compte de ce raisonnement intuitif. Ces consignes ne devront intégrer le raisonnement intuitif que s'il satisfait les experts. Si, au contraire, le raisonnement intuitif n'est pas satisfaisant, alors les consignes devront en tenir compte et en avertir les annotateurs qui pourraient être tentés de l'utiliser. Finalement, nous sommes persuadée que les discussions qui suivent des annotations divergentes par des experts peuvent être la base d'une meilleure compréhension et surtout d'une meilleure formalisation de la tâche que l'on cherche à définir.

Néanmoins, notre manuel d'annotation ne permet pas d'obtenir un accord suffisamment élevé entre les annotateurs. Il nécessiterait des révisions supplémentaires pour permettre l'annotation d'un corpus réellement utile. Il reste plusieurs points à clarifier. En particulier, nous aimerions explorer la notion de subjectivité. En effet, la causalité est une notion subjective et c'est le point de vue du texte lui-même et non pas de l'annotateur qui doit être analysé. Nous devons également chercher à annoter précisément quelle partie du texte représente la cause et quelle partie représente l'effet. Les limites de ces segments ne seront probablement pas tout de suite consensuelles et nécessiteront des clarifications.

Enfin, cet article s'inscrit dans un travail de thèse de linguistique

informatique. Dans le cadre de cette thèse, nous souhaitons développer un programme capable de détecter la présence de relations causales dans des textes en français. La création de ce manuel d'annotation nous a permis, d'une part, de clarifier la tâche, et, d'autre part, d'évaluer la qualité de notre système en comparant ses résultats aux réponses correctes définies par le manuel.

### Bibliographie

- CARLSON L., MARCU D. & OKUROWSKI M. (2002), « Building a discourse-tagged corpus in the framework of rhetorical structure theory », *Proceedings of the Second SIGdial Workshop on Discourse and Dialogue, Association for Computational Linguistics*, 1-10.
- COHEN J. (1960), « A coefficient of agreement for nominal scales », *Psychological Bulletin* 20, 37-46.
- HOVY E.H., MARCUS M.P., PALMER M., RAMSHAW L.A. & WEISCHEDEL R.M. (2006), « Ontonotes : The 90% Solution », in MOORE R.C., BILMES J.A., CHU-CARROLL J. & SANDERSON M. (eds), *HLT-NAACL, The Association for Computational Linguistics*, 57-60.
- HUME D. (1739-1740), *A Treatise of Human Nature*, London, Millar.
- INUI T. (2005), « Creating an annotated corpus for the analysis of causal relations », *COE-LKR2005*, <http://www.mibel.cs.tsukuba.ac.jp/~inui/paper/coe2005.pdf>.
- MARCUS M.P., SANTORINI B. & MARCINKIEWICZ M.A. (1994), « Building a large annotated corpus of English : The Penn TreeBank », *Computational Linguistics* 19/2, 313-330.
- MOESCHLER J. (2003), « Causality, lexicon, and discourse meaning », *Rivista di Linguistica* 15/2, 343-369.
- RALI LABORATORY (1997), *Corpus de bitextes anglais-Français*, Université de Montréal.
- REBOUL A. (2005), « Similarities and differences between human and non-human causal cognition », [www.interdisciplines.org/causality](http://www.interdisciplines.org/causality).
- SANDERS T. (2005), « Coherence, causality and cognitive complexity in discourse », in LE DRAOULEC A., AURNAGUE M., BRAS M. & VIEU L. (eds), *Proceedings of SEM-05, First International Symposium on the Exploration and Modelling of Meaning*, Biarritz, 105-114.
- TALMY L. (1988), « Force dynamics in language and cognition », *Cognitive Science* 12/1, 49-100.

### Annexe 1 : les consignes d'annotation

Note : on parle ici d'événements de façon générale pour désigner aussi bien des états que des événements. Nous nous intéressons ici uniquement à des exemples causaux entre événements représentés par des propositions.

### ***Quelques caractéristiques de la causalité***

Une cause se passe toujours avant une conséquence. Dans certains cas il est difficile de juger de la temporalité, comme dans *Jean a tapé le verre contre la table et le verre s'est cassé* ou *Jacques s'est fatigué en conduisant*, mais la cause ne peut en tous cas jamais se passer après la conséquence.

On vous propose trois tests linguistiques avec *C'est parce que*, *parce que* et *donc*. En général, *c'est parce que* est le test le plus fiable, si le test de *c'est parce que* ne fonctionne pas, alors il n'y a pas de relation causale. Les deux autres tests peuvent aider, mais ils peuvent être plus ambigus. Par exemple, dans *Jean est malade parce que je ne l'ai pas vu de la journée* la deuxième partie n'est pas la cause de la première.

Une autre caractéristique de la causalité est la propriété contrefactuelle : si la cause ne s'était pas passée alors la conséquence ne se serait pas passée non plus. Quand on utilise ce test il faut faire attention à ce qui se serait probablement passé. Par exemple, si on se demande si *Jean s'est cassé la jambe en skiant* est causal, il faut se demander ce qui se serait probablement passé si Jean n'était pas allé skier (il ne se serait pas cassé la jambe), même s'il aurait pu se casser la jambe d'une autre façon, par exemple en glissant sur une peau de banane.

### ***Cas ambigus et quelques trucs supplémentaires***

#### **Tests supplémentaires**

Il est parfois utile d'imaginer une suite de causes et de conséquences directes entre les événements que l'on teste pour voir s'il s'agit d'une causalité indirecte (on considère toujours un cas de causalité indirecte comme causal). Par exemple si on teste *Jean est à l'hôpital, Jacques l'a poussé* on peut trouver la suite Jacques pousse Jean → Jean tombe → Jean se blesse → Jacques appelle les secours → les secours constatent que Jean est blessé → les secours emmènent Jean à l'hôpital → Jean est à l'hôpital. On parle alors de *chaîne causale*.

La causalité est asymétrique. Si on a deux événements, l'un peut causer l'autre, ou il peut ne pas y avoir de relation causale, mais les deux ne peuvent pas se causer l'un l'autre. Si on n'arrive pas à déterminer parmi les deux événements, lequel est la cause et lequel est la conséquence, on a souvent affaire à un cas non causal.

#### ***Cas non causaux typiques***

La causalité n'est pas le *si ... alors* logique. Par exemple *si George Bush est le président des États Unis, alors l'Allemagne est en Europe* est

logiquement correct, mais pas causal. On peut le déterminer par l'impossibilité de construire une chaîne causale, ou par le fait qu'il n'y a pas d'asymétrie. *C'est un triangle, il a trois côtés* pose le même problème, même si le raisonnement contrefactuel s'applique et que la phrase passe le test de *parce que* et de *donc*. Le test de *c'est parce que* (*c'est parce qu'il a trois côtés que c'est un triangle*) est plus difficile. Mais ici, c'est surtout la symétrie qui devrait faire suspecter un cas non causal.

### **Cas ambigus typiques**

Le contexte peut être la cause. Par exemple *Il s'est cassé le bras en faisant du patin à roulettes*. Si on essaie de faire une chaîne causale, on devra sortir un sous-événement (par exemple foncer dans un mur, ou tomber) de l'événement contexte (faire du patin à roulettes). Ce cas est ambigu. Pour cette application nous considérerons les événements contextes comme causaux, et *Il s'est cassé le bras en faisant du patin à roulettes* devrait être positif pour la causalité.

Cas d'actes de langage et épistémiques. On considérera les cas du type *dépêche-toi parce qu'on va être en retard* comme non causaux. En fait il y a une relation causale entre *on va être en retard* et *je te demande de te dépêcher*, mais on ne résoudra pas ce type d'ellipse ici. De même dans *il va faire jour, il est 7h30 l'heure cause la croyance qu'il va faire jour*, pas le fait qu'il va faire jour et on ne résout pas non plus l'ellipse dans ce cas. Par contre, le cas similaire *Georges est sorti parce que sa veste n'est pas sur la chaise* est causal, mais dans l'autre sens. On ne résout toujours pas l'ellipse *je crois que* mais ici la cause de la croyance est la conséquence du fait que l'on croit être vrai, donc on a un cas causal : *Georges est sorti cause sa veste n'est pas sur la chaise*.

Les événements négatifs et les états sont souvent difficiles. Par exemple *Marjorie est majeure, elle peut voter* ou *Comme aucune étude n'a été faite auparavant sur ce sujet, ce rapport sera fort pertinent*. Dans ce genre de cas, il est souvent utile d'essayer de faire une chaîne causale (p.ex. il n'y a pas d'études → les spécialistes n'ont pas de données → les spécialistes veulent des données → ce rapport sera pertinent) ou de chercher à remonter à une loi générale (p.ex. si on est majeur, on peut voter). Dans tous les cas, il faut bien vérifier l'asymétrie (*C'est parce qu'elle est majeure que Marjorie peut voter* marche parfaitement alors que *c'est parce que Marjorie peut voter qu'elle est majeure* est bizarre).

### **Annexe 2 : données des expériences d'annotation**

Nous présentons maintenant les textes qui devaient être analysés pour les expériences d'annotation. Pour la première expérience, nous représentons la cause potentielle en gras et la conséquence potentielle

en italique. Pour la deuxième expérience, où les annotateurs devaient choisir lequel des deux événements était la cause et lequel la conséquence, nous présentons le premier événement en italique et le second en gras.

*Phrases causalement ambiguës pour la première expérience d'annotations*

- (1) **Paul a pris ses médicaments**, *il va guérir* : test positif pour la causalité.
- (2) *Il fait jour maintenant*, **il faisait nuit il y a quelques heures** : test négatif pour la causalité.
- (3) *Si George Bush est président des États-Unis en 2007*, **alors l'Allemagne est en Europe** : autre test négatif pour la causalité mais avec un *si...* *alors* pour détecter des sujets qui confondraient la causalité et le *si alors* logique.
- (4) *J'ai mis le poulet dans le four*, **je suis en train de cuisiner** : on teste ici la relation partie-tout qui peut être perçue comme causale ou non.
- (5) **En signant l'accord**, *il a pris de gros risques* : une autre relation partie-tout qui semble plus causale que la précédente.
- (6) *Il s'est cassé la jambe* **en skiant** : exemple vraisemblablement causal, mais difficile à classer temporellement. Est-ce que se casser la jambe arrive après ou pendant skier ?
- (7) *Il s'est fatigué* **en roulant toute la journée** : même type de problème que l'exemple précédent, mais avec deux activités plutôt qu'une activité et un événement. Ce type d'exemple est difficile à classer temporellement, ce qui pourrait avoir un impact sur la causalité.
- (8) **Frédérique est majeure**, *elle peut voter* : autre exemple avec deux états. La capacité causale d'un état est moins évidente que celle d'un événement, et il y a recouvrement temporel.
- (9) **Le gendarme court**, *il est essoufflé* : *idem*, mais avec des activités.
- (10) *Après qu'il a fait un long voyage au Canada*, **il s'est marié** : une relation temporelle marquée et causalement ambiguë sans contexte. L'expérience doit montrer si les sujets enrichissent la relation temporelle avec une relation causale.
- (11) **Il a ouvert son courrier** puis *il est sorti* : même problème que la phrase précédente.
- (12) *George est sorti*, **sa veste n'est pas sur le porte-manteau** : les tests linguistiques avec *parce que* ou *donc* fonctionnent mais il s'agit d'un usage épistémique et *sa veste n'est pas sur le porte-manteau* n'est pas la cause de *George est sorti*.
- (13) *Dépêche-toi!* **On va être en retard** : *idem* mais avec un acte de

parole.

(14) *Le Concorde ne s'arrêtera pas vraiment, il ne sortira jamais de l'imaginaire des hommes* : cette phrase contenait un *parce que* explicite. La causalité est difficile à cerner ici à cause de l'emploi argumentatif.

(15) *C'est un triangle, il a trois côtés* : ici, le raisonnement contrefactuel donne un résultat positif, mais on considère l'exemple comme non causal.

(16) *J'ai mis le champagne au frigo* pour qu'il soit bien frais pour le boire : il y a bien un rapport causal mais il est donné comme raison d'un acte.

(17) *Il faut accorder du mérite au gouvernement, il essaie de déterminer ce qui est bon pour le pays* : un exemple similaire avait engendré un accord très bas dans une expérience précédente, probablement à cause de la présence de *falloir*.

(18) *Aucune étude semblable n'a été faite auparavant. Cette recherche sera fort pertinente* : comme le précédent, cet exemple avait engendré un accord bas dans une expérience précédente. Peut-être à cause de la négation dans l'événement cause potentiel.

(19) *Véronique s'est coincé le doigt, elle a mal au pied* : la structure suggère une interprétation causale, mais incongrue.

(20) *Le poulet a cuit, je l'avais mis dans le frigo* : *idem*.

(21) *L'oreille de Joël est tombée, il a pris une douche* : *idem*.

(22) *George est allé dans le désert, il a été pris dans une tempête de neige* : *idem* mais avec l'ordre inversé.

(23) *Marjorie est allée à la piscine, elle a perdu la tête* : *idem*.

(24) *Marjorie est allée à la piscine, elle a perdu la tête* : *idem* mais avec un autre ordre pour voir si l'ordre importe.

*Exemples d'extraits de De la Terre à la lune de Jules Verne pour la dernière expérience d'annotation.*

(1) Bref, il ne fut plus permis, même au moins lettré des Yankees, d'ignorer un seul des faits relatifs à son satellite, ni à la plus bornée des vieilles *mistress* d'admettre encore de superstitieuses erreurs à son endroit. **La science leur arrivait sous toutes les formes** elle les pénétrait par les yeux et les oreilles ; impossible d'être un âne...en astronomie.

(2) Jusqu'alors, bien des gens ignoraient comment on avait pu calculer la distance qui sépare la Lune de la Terre. **On profita de la circonstance pour leur apprendre que cette distance s'obtenait par la mesure de la parallaxe de la Lune.** Si le mot parallaxe semblait les étonner, on leur disait que c'était l'angle formé par deux lignes droites menées de

chaque extrémité du rayon terrestre jusqu'à la Lune. Doutaient-ils de la perfection de cette méthode, on leur prouvait immédiatement que, non seulement cette distance moyenne était bien de deux cent trente-quatre mille trois cent quarante-sept milles (– 94,330 lieues), mais encore que les astronomes ne se trompaient pas de soixante-dix milles (– 30 lieues).

(3) Jusqu'alors, bien des gens ignoraient comment on avait pu calculer la distance qui sépare la Lune de la Terre. On profita de la circonstance pour leur apprendre que cette distance s'obtenait par la mesure de la parallaxe de la Lune. Si le mot parallaxe semblait les étonner, on leur disait que c'était l'angle formé par deux lignes droites menées de chaque extrémité du rayon terrestre jusqu'à la Lune. *Doutaient-ils de la perfection de cette méthode, on leur prouvait immédiatement que, non seulement cette distance moyenne était bien de deux cent trente-quatre mille trois cent quarante-sept milles (– 94,330 lieues), mais encore que les astronomes ne se trompaient pas de soixante-dix milles (– 30 lieues).*

(4) *A ceux qui n'étaient pas familiarisés avec les mouvements de la Lune, les journaux démontraient quotidiennement qu'elle possède deux mouvements distincts, le premier dit de rotation sur un axe, le second dit de révolution autour de la Terre, s'accomplissant tous les deux dans un temps égal, soit vingt-sept jours et un tiers [C'est la durée de la révolution sidérale, c'est-à-dire le temps que la Lune met à revenir à une même étoile].*

(5) *Le mouvement de rotation est celui qui crée le jour et la nuit à la surface de la Lune ; seulement il n'y a qu'un jour, il n'y a qu'une nuit par mois lunaire, et ils durent chacun trois cent cinquante-quatre heures et un tiers. Mais, heureusement pour elle, la face tournée vers le globe terrestre est éclairée par lui avec une intensité égale à la lumière de quatorze Lunes. Quant à l'autre face, toujours invisible, elle a naturellement trois cent cinquante-quatre heures d'une nuit absolue, tempérée seulement par cette « pâle clarté qui tombe des étoiles ». Ce phénomène est uniquement dû à cette particularité que les mouvements de rotation et de révolution s'accomplissent dans un temps rigoureusement égal, phénomène commun, suivant Cassini et Herschell, aux satellites de Jupiter, et très probablement à tous les autres satellites.*

