

The complementary relationship between linguistics and IT The Extraction of Logical Connectors in the Thoughts of Blaise Pascal

Refka Daoud

LLTA research laboratory, University of Sfax
CPTC Laboratory, Multidisciplinary Center, Texts and Cultures University of Bourgogne-Franche-Comté

Abstract

*In this article, we are interested in the complementary relationship between linguistics and computer science to insist on the crossing of these two disciplines and subsequently on the concept of interdisciplinarity. To do this, we have taken as an example the extraction of information, and more precisely logical connectors. Our approach uses a terminology extraction program that applies to the work of Blaise Pascal, *Les Pensées*. We want to extract a few connectors from the Nooj platform.*

Keywords: Interdisciplinarity, linguistics, computing, complementarity, extraction of information, Nooj.

La relation de complémentarité entre linguistique et informatique L'extraction des connecteurs logiques dans les Pensées de Blaise Pascal

Refka Daoud

Laboratoire de recherche LLTA, Université de Sfax
Laboratoire CPTC, Centre Pluridisciplinaire, Textes et Cultures Université de Bourgogne-
Franche-Comté

Résumé

Dans cet article, nous nous intéressons à la relation de complémentarité entre linguistique et informatique afin d'insister sur le croisement de ces deux disciplines et par la suite sur le concept d'interdisciplinarité. Pour ce faire, nous avons pris comme exemple l'extraction des informations, et plus précisément des connecteurs logiques. Notre approche utilise un programme d'extraction terminologique qui s'applique à l'œuvre de Blaise Pascal, les Pensées. Nous tenons d'extraire quelques connecteurs à partir de la plateforme Nooj.

Mots clés : Interdisciplinarité, linguistique, informatique, complémentarité, extraction des informations, Nooj.

Introduction

Les notions de *pluridisciplinarité* et d'*interdisciplinarité* ont émergé dans la communauté scientifique et universitaire, au cours des années soixante. Elles traduisaient un besoin pratique et une aspiration vers l'idéal philosophique de l'unité.

Les débats sont nombreux sur cette question.

L'interdisciplinarité est l'«interaction existant entre deux ou plusieurs disciplines, qui peut aller de la simple communication des idées jusqu'à l'intégration mutuelle des concepts directeurs de l'épistémologie, de la terminologie, de la méthodologie, des procédés, des données et de l'orientation de la recherche et de l'enseignement s'y rapportant¹ »

Elle suppose donc un dialogue et un échange de connaissance, d'analyses, de méthodes entre deux ou plusieurs disciplines. Elle implique qu'il y ait des interactions et un enrichissement mutuel entre plusieurs spécialistes. L'interdisciplinarité est aussi le principe que l'on retrouve comme fondateur des *sciences cognitives*. Son but est d'élaborer un formalisme général pour permettre d'exprimer dans un langage commun, les concepts, les préoccupations et les contributions d'un certain nombre de disciplines. Elle facilite les contacts et les échanges entre des domaines de connaissance éloignés, elle met en place une dialectique et des échanges entre les disciplines et elle assure la créativité scientifique déjà évoquée par Vinck.

L'interdisciplinarité favorise l'enrichissement des informations collectées et par la suite une compréhension plus complète de la question étudiée. Elle constitue « un facteur important dans le développement des connaissances² ». Elle implique donc une interaction et une coopération active entre les disciplines.

Le concept d'*interdisciplinarité* s'applique de façon générale aux diverses branches de la connaissance. Nous allons nous concentrer sur la relation entre linguistique et informatique. Dès lors, cette approche fait collaborer étroitement des linguistes et des informaticiens, spécialistes du traitement automatique de la langue, de l'extraction d'information et de l'apprentissage automatique. Et pour mettre en valeur cette relation de complémentarité, notre travail portera sur le traitement automatique ou plus précisément sur l'extraction des connecteurs logiques dans les Pensées de Blaise Pascal.

I- La relation de complémentarité entre linguistique et informatique

L'informatique ouvre de nouveaux champs d'application de l'instrumentation informatique du langage humain. Les informaticiens cherchent à instrumenter les divers aspects du langage.

Les linguistes s'appuient sur des techniques informatiques pour faciliter leur tâche. Dès lors, l'informatique nous ouvre la voie à de nouveaux domaines de recherche.

Linguistique et informatique se complètent. Nous pouvons citer comme exemple les travaux de Chomsky : sa théorie linguistique « la grammaire générative ». Ses travaux sur la syntaxe des langues naturelles ont présenté une source d'inspiration pour traiter les phrases comme des expressions d'un langage formel.

C'est ce qui a favorisé la naissance de la linguistique informatique qui présente une approche interdisciplinaire étudiant le traitement automatique des Langues naturelles. Elle fait

¹Grand Dictionnaire Terminologique, de l'Office québécois de la langue française.

² Georges Gusdorf

collaborer des linguistes et des informaticiens. Nous notons la présence d'innombrables applications de la linguistique informatique dont nous citons la traduction automatique (qui nécessite des analyses lexicale, syntaxique, sémantique et pragmatique afin de déterminer le sens particulier d'un mot, dans le contexte où il apparaît), l'extraction de l'information...

En effet, tous les travaux en traitement automatique des langues s'appuient sur des éléments d'analyse linguistique. En effet, l'informatique aide les linguistes pour :

- chercher des exemples et des contre-exemples dans de grands corpus
- faire des statistiques sur les emplois de termes, de constructions syntaxiques, des rimes...
- calculer des occurrences de termes

Ces recherches ont favorisé la naissance de la linguistique du corpus : l'informatique nous permet l'utilisation de grands corpus de textes pour l'étude des langues. Elle facilite leur recueil, leur constitution et leur annotation.

De sa part, l'informatique a besoin du langage pour son propre compte, par exemple dans les systèmes d'aides, pour la recherche d'information dans des bases de documents textuels. Elle inclut la théorie de la compilation, les langages de programmation, l'intelligence artificielle et le traitement automatique des langues. Ces domaines croisent le chemin de la linguistique. Par exemple le traitement automatique est un domaine de recherche pluridisciplinaire qui fait collaborer linguistes, informaticiens, logiciens, lexicographes ou traducteurs. Il a pour objectif de développer des logiciels ou des programmes informatiques capables de traiter de façon automatique des données linguistiques. Pour ce faire, nous devons expliciter les règles de la langue, puis les représenter dans des formalismes opératoires et calculables, et enfin les implémenter à l'aide de programmes informatiques. Les deux disciplines travaillent ensemble, l'une dépend de l'autre afin d'enrichir les approches en favorisant la créativité.

Nous constatons que la linguistique informatique a pour but la découverte des faits linguistiques à l'aide de processus informatiques, autrement dit, elle permet d'explorer le langage à l'aide d'outils informatiques. Elle présente un champ interdisciplinaire basé sur une modélisation symbolique ou statistique du langage naturel, établie dans une perspective informatique.

II- L'extraction des connecteurs logiques

Avec la disponibilité croissante de textes aux formats numériques, le besoin d'extractions et d'organisation de l'information à partir de ces ressources s'agrandit graduellement. La linguistique informatique joue un rôle très important dans ce procédé.

L'extraction d'information « est une approche de l'analyse de données textuelles basées sur une analyse locale qui permet d'extraire automatiquement des éléments textuels correspondant à un besoin donné à partir d'un texte en langage naturel³. Ce domaine a connu un développement constant tout au long des années 90 [...]». Dès lors, l'extraction des occurrences permet d'envisager de quantifier différents aspects de ce fait linguistique. Parmi les outils de l'extraction, nous citons l'utilisation des patrons d'extraction syntaxiques, morphosyntaxiques ou lexico-syntaxiques.

Ce système de repérage et d'extraction automatique est intéressant pour nous les linguistes parce qu'il nous permet d'économiser du temps et des efforts et aussi de traiter plus de données textuelles.

³ Poibeau et Nazarenko, 1999

Nous allons prendre comme exemple les « connecteurs logiques », nous allons les extraire afin de montrer l'importance de ces connecteurs. Plus précisément, nous avons essayé de justifier l'idée selon laquelle les connecteurs ont pour fonction de « mettre les informations contenues dans un texte au service de l'intention argumentative globale de celui-ci ⁴».

Du point de vue morphologique, les connecteurs pragmatiques sont des *mots outils*: ils ont une fonction argumentative et/ou ils servent à structurer les énoncés. Nous proposons la définition suivante du mot «connecteur»: « le connecteur est un élément articulateur de la phrase⁵ qui lie deux ou plusieurs énoncés en créant un nouveau sens à partir des sens primitifs des énoncés ».

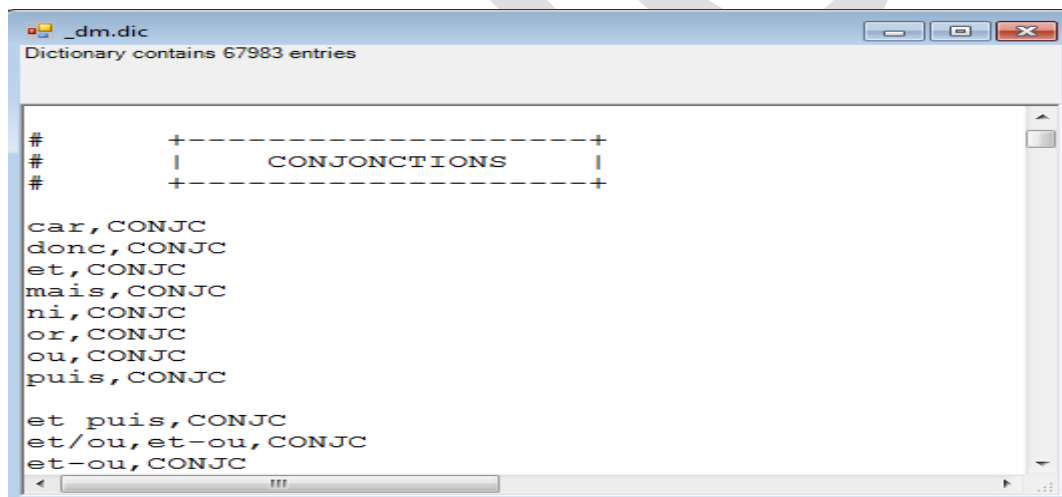
Pour mieux comprendre, voici la structure de base d'un connecteur ayant deux arguments:

❖ **P conn Q⁶**

Nous allons présenter un échantillon du traitement automatique des connecteurs «donc», «mais» et «parce que» à l'aide de NooJ4Web⁷ qui nous permet la recherche de mots, de lemmes et de leurs formes fléchies ou dérivées, de collocations complexes par exemple les noms composés et les expressions figées.

L'extraction des connecteurs «mais» et «donc» ne nécessite pas la création d'un dictionnaire puisque les données concernant ces connecteurs sont déjà stockées et implémentées dans la mémoire de NooJ sous forme de dictionnaire et de grammaire. Nous allons évoquer un échantillon de ce dictionnaire :

❖ **Un échantillon du dictionnaire_dm.dic**

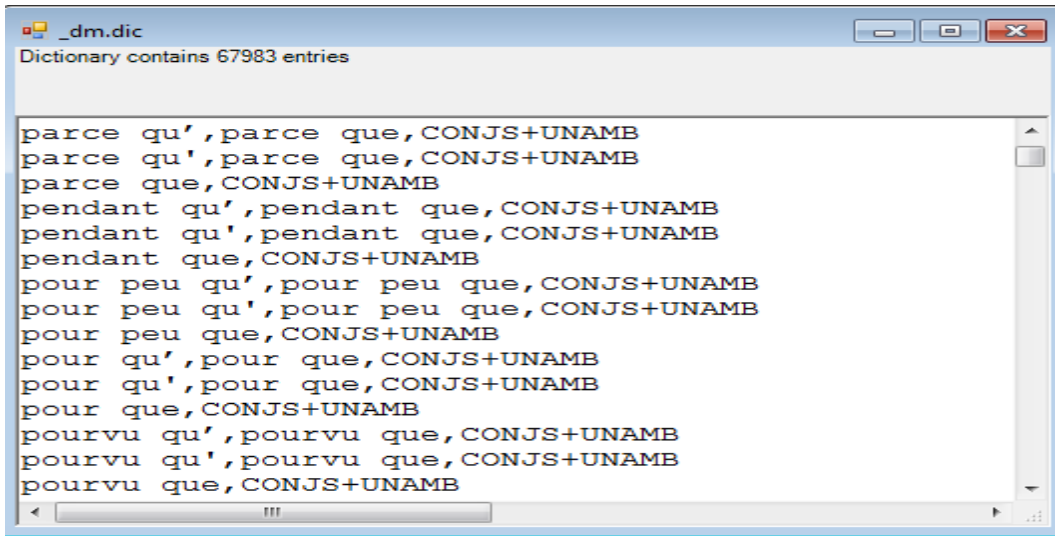


⁴Plantin, 1989 : fiche 10

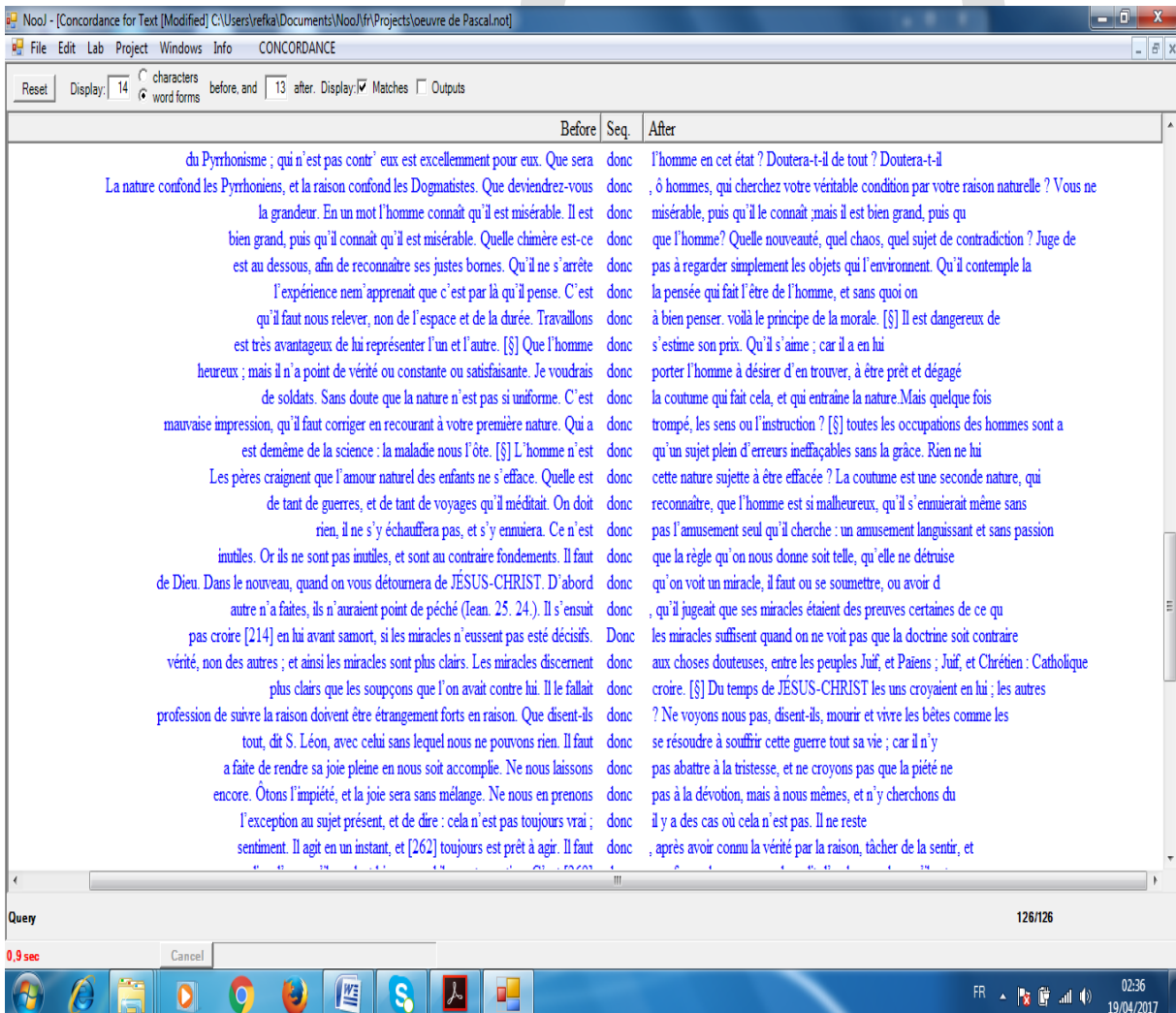
⁵Conjonction, adverbe, locution

⁶ P et Q étant les arguments du connecteur, conn signifiant « connecteur ».

⁷ Ce logiciel est développé par Marx Silberstein. Il est considéré comme une plateforme linguistique de développement, un système de recherche documentaire, un extracteur terminologique.



✓ **Table de concordances du connecteur « donc »**



Les résultats que nous venons d'obtenir sont affichés sous forme de concordances dans une grille de données faisant apparaître les séquences pertinentes entourées par les contextes dans lesquels elles apparaissent. La conjonction de coordination « donc » est un terme de logique qui permet d'évoquer la conclusion déduite d'une proposition, d'un principe ou d'un fait. Par exemple, parmi les déductions qui sont évoquées dans cette grille le fait que la pensée fait la grandeur de l'homme (« il est grand puisqu'il connaît qu'il est misérable », « c'est donc la pensée qui fait l'être de l'homme », « travaillons donc à bien penser, voilà le principe de la morale »...).

✓ **Table des concordances du connecteur « mais »**

Text	Before	Seq.	After
donc servir qu'à humilier la raison qui voudrait juger de tout ;	mais		non pas à combattre notre certitude, comme s'il n'y avait que la raison capable de
besoin, et que nous commissions toutes choses par instinct et par sentiment.	Mais		la nature nous a refusé ce bien, et elle ne nous a donné que très peu de
l'impuissance où nous sommes, d'arriver au bien par [159] nos efforts.	Mais		l'exemple ne nous instruit point. Il n'est jamais si parfaitement semblable, qu'il n'y
personne ne le pût perdre contre son gré. Ils l'ont compris,	mais		ils ne l'ont pu trouver ; et au lieu d'un bien solide et effectif, ils n
Les autres ont voulu y renoncer à la raison, et devenir bêtes. [162]	Mais		ils ne l'ont pu ni les uns ni les autres ; et la raison demeure toujours qui
'il est misérable. Il est donc misérable, puis qu'il le connaît ;	mais		il est bien grand, puis qu'il connaît qu'il est misérable. Quelle chimère est-ce donc
'est à dire une certaine portion de matière qui lui est propre.	Mais		pour comprendre ce qu'elle est, il faut qu'il la compare avec tout ce qui est
'égard de celui que les astres qui roulent dans le firmament embrassent.	Mais		[166] si notre vue s'arrête là, que l'imagination passe outre. Elle se lassera plutôt de concevoir
Qu'est-ce qu'un homme dans [167] l'infini ? Qui le peut comprendre ?	Mais		pour lui présenter un autre prodige aussi étonnant, qu'il recherche dans ce qu'il connaît les
de la nature. Je veux lui peindre non seulement l'univers visible,	mais		encore tout ce qu'il est capable de concevoir de l'immensité de la nature, dans l
et d'édifier une tour, qui s'élève jusqu'à l'infini.	Mais		tout notre édifice craque, et la terre s'ouvre jusqu'aux abîmes. [171] 83 Chapitre 24 JE puis bien conc
Il est vrai que c'est être misérable, que de se connaître misérable ;	mais		c'est aussi être grand, que de connaître qu'on est misérable. Ainsi toutes ses misères prouvent
été ; parce que sa condition n'était pas de l'être toujours.	Mais		on trouvait Persée si malheureux de n'être plus Roi, parce que sa condition était de l
être jamais avisé de s'affliger de n'avoir pas trois yeux ;	mais		on est inconsolable de n'en avoir qu'un. [173] [§] Nous avons un si grande idée de l
'homme n'est qu'un roseau le plus faible de la nature ;	mais		c'est un roseau pensant. Il ne faut pas que l'univers entier s'arme pour l
l'écraser. Une vapeur, une goutte d'eau suffit pour le tuer.	Mais		quand l'univers l'écraserait, l'homme serait encore plus noble que ce qui le tue ; parce
encore plus dangereux de lui laisser ignorer l'un et l'autre. [175]	Mais		il est très avantageux de lui représenter l'un et l'autre. [§] Que l'homme donc s
s'aime ; car il a en lui une nature capable de bien ;	mais		qu'il n'aime pas pour cela les bassesses qui y sont. Qu'il se méprise parce que
qui y sont. Qu'il se méprise parce que cette capacité est vide ;	mais		qu'il ne méprise pas pour cela cette capacité naturelle. Qu'il se haïsse ; qu'il s
en lui la capacité de connaître la vérité, et d'être heureux ;	mais		il n'a point de vérité ou constante ou satisfaisante. Je voudrais donc porter l'homme à
'est donc la coutume qui fait cela, et qui entraîne la nature.	Mais		quelque fois aussi la nature la surmonte, et retient l'homme dans son instinct, malgré toute la
'être estimé dans les villes où l'on ne fait que passer ;	mais		quand on y doit demeurer un peu de temps on s'en soucie. Combien de temps faut
moyens ; le seul avenir est notre objet. Ainsi nous ne vivons jamais ;	mais		nous espérons de vivre ; [181] et nous disposant toujours à être heureux, il est indubitable que nous ne
qui se mit dans son urètre. Rome même allait trembler sous lui.	Mais		ce petit gravier, qui n'était rien ailleurs, mis en cet endroit, le voilà mort, sa famille
est bon en effet de la suivre, puisque la mode en est ;	mais		comme si chacun savait certainement où est la raison et la justice. On se trouve déçu à
trop bas. La perspective l'assigne dans l'art de la peinture.	Mais		dans la vérité et dans la morale qui l'assignera. [§] Cette maîtresse d'erreur que l'on
serait règle infallible de vérité, si elle l'était infallible du mensonge.	Mais		estant le plus souvent fausse, elle ne donne aucune marque de sa qualité, marquant de même caractère

Le connecteur « mais » permet de réévaluer une information précédemment entretenue comme vraie. Il oppose deux arguments en les enchaînant. Nous allons présenter cet exemple

que nous avons dans notre grille de concordances « l'homme n'est qu'un roseau, le plus faible de la nature ; mais c'est un roseau pensant, il ne faut pas que l'univers entier s'arme pour l'écraser ». Pascal présente la condition de l'être humain, à savoir sa fragilité et sa mortalité. Nous avons affaire à un roseau qui peut mourir à tout instant, qui s'approche inévitablement de sa propre mort, qui est mourant en plus d'être mortel, et d'autre part, à une pensée qui sait tout cela. Le connecteur « mais » met en valeur l'opposition entre « faiblesse physique » et « pensée » : la force de l'être humain et sa dignité proviennent de sa pensée.

✓ **Table des concordances du connecteur « parce que »**

Text	Before	Seq.	After
qui a déjà été imprimée deux ou trois fois sur des copies assez peu correctes,	parce que		ces impressions ont été faites sans la participation de ceux qui donnent à y
réussir c'est de paraître honnête, fidèle, judicieux, et capable de servir utilement ses amis ;	parce que		les hommes n'aiment naturellement que ce qui leur peut être utile. Or quel
a que deux sortes de personnes ; ou ceux qui servent Dieu de tout leur coeur,	parce qu'		ils le connaissent ; ou ceux qui le cherchent de tout leur coeur, parce qu'ils
coeur, parce qu'ils le connaissent ; ou ceux qui le cherchent de tout leur coeur,	parce qu'		ils ne le connaissent pas encore. C'est donc pour les personnes qui cherch
qui ils méprisent pour ne les pas mépriser jusqu'à les abandonner dans leur folie. Mais	parce que		cette Religion nous oblige de les regarder toujours tant qu'ils seront en ce
monde en général. Sur ce fondement les impies prennent lieu de blasphémer la Religion Chrétienne,	parce qu'		ils la connaissent mal. Ils s'imaginent qu'elle consiste simplement en l'ador
tout à fait contraire. Et delà ils concluent que cette religion n'est pas véritable ;	parce que		si elle l'était il faudrait que Dieu [27] se manifestât aux hommes par des p
premier avènement de JÉSUS-CHRIST est prédit ; le temps du second ne l'est point ;	parce que		le premier devait être caché ; au lieu que le second doit être éclatant et tel
étaient les dépositaires des prophéties ; et ils le prouvaient aussi en ne le recevant point ;	parce qu'		en cela ils accomplissaient les prophéties. [§] Les Juifs avaient des miracle
des Juifs n'était pas vraie, quoiqu'elle eût les miracles, les prophéties, et la perpétuité ;	parce qu'		elle n'avait pas cet autre point de n'adorer et n'aimer que Dieu. La Religio
partie n'est donc pas égale. Il ne faut pas égaler et confondre ces choses	parce qu'		elles semblent être semblables par un bout, étant si différentes par l'autre.
se souviennent de lui quand il sera en son Royaume. [§] La Synagogue ne périssait point,	parce qu'		elle était la figure de l'Eglise ; mais parce qu'elle n'était que la figure, elle e
La Synagogue ne périssait point, parce qu'elle était la figure de l'Eglise ; mais	parce qu'		elle n'était que la figure, elle est tombée dans la servitude. La figure a subs
agréant vos parfums vous donnerait en récompense une terre abondante, Dieu l'aura pour vous,	parce que		vous avez eu pour lui, la même intention qu'un homme a pour celui à qui il
coeur (Jer. 31. 33.). Que les Juifs réprouveraient JÉSUS-CHRIST, et qu'ils seraient réprouvés de Dieu (Idem 32. 40.),	parce que		la vigne élue ne donnerait que du verjus (Is. 5. 2. 3. 4. etc.). Que le peupl
des miracles, et les Apôtres ensuite, et les premiers Saints en ont fait aussi beaucoup ;	parce que		les prophéties n'étant pas encore accomplies, et s'accomplissant par aux,
seconde destruction est sans promesse de rétablissement, sans [126] Prophètes, sans Rois, sans consolation, sans espérance ;	parce que		le sceptre est ôté pour jamais. Ce n'est pas avoir été captif que de l'avoir
trouve dans tous leurs 61 écrits, dans le Talmud, et dans les Rabbins, n'est que	parce que		JÉSUS-CHRIST n'a pas dompté les nations à main armée. JÉSUS-CHR
'est pas en cette sorte qu'il a voulu paraître dans son avènement de douceurs ;	parce que		tant d'hommes se rendants indignes de sa clémence, il a voulu les laisser c
à ceux qui le tentent, et qu'il se découvre à ceux qui le cherchent ;	parce que		les hommes sont tout ensemble indignes de Dieu, et capables de Dieu ; in
nations. Et de là vient qu'étant dans le désert, ils ne furent pas circoncis,	parce qu'		ils ne pouvaient se confondre avec les autres peuples ; et que depuis que J
de Dieu, et pour s'unir à lui. Il ne faut point séparer ces connaissances ;	parce qu'		étant séparées, elles sont non seulement inutiles, mais nuisibles. La connais
Mais la connaissance de Jésus-Christ nous exempte et de l'orgueil, et du désespoir ;	parce que		nous y trouvons Dieu, notre misère, et la voie unique de la réparer. Nous
sans connaître tout [150] ensemble et Dieu, et nos misères, et le remède de nos misères ;	parce que		JÉSUS-CHRIST n'est pas simplement Dieu, mais que c'est un Dieu répa
il y a un Dieu ; ou, s'ils y arrivent, c'est inutilement pour eux ;	parce qu'		ils se forment un moyen de communiquer sans médiateur avec ce Dieu qu'
consul ? Au contraire tout le monde trouvait qu'il était heureux de l'avoir été ;	parce que		sa condition n'était pas de l'être toujours. Mais on trouvait Persée si malh
de l'être toujours. Mais on trouvait Persée si malheureux de n'être plus Roi,	parce que		sa condition était de l'être toujours, qu'on trouvait étrange qu'il pût supp

Le connecteur « parce que » met en relation deux contenus sémantiques posés comme vrais et il explique l'un par rapport à l'autre. C'est un opérateur qui sert à constituer, à partir de deux idées p et q qu'il relie, une nouvelle idée, à savoir l'idée d'une relation de causalité entre p et q. Nous avons choisi un exemple de cette grille: le connecteur « parce que » met en

relation deux idées celle du « bonheur » et de « la recherche de Dieu ». L'homme cherche le remède de ses maux, mais en vain pour une simple raison il le cherche dans les « choses extérieures » tant dis que le remède est en « Dieu ».

Conclusion

Dans cette étude nous avons présenté le concept d'interdisciplinarité en tant que l'art qui fait travailler ensemble des personnes ou des équipes de diverses disciplines scientifiques. Nous avons pris l'exemple de la linguistique et de l'informatique afin de montrer que l'interdisciplinarité croise les démarches scientifiques de chaque discipline. Elle met en place un dialogue et des échanges entre les disciplines. Nous avons évoqué comme exemple la linguistique informatique, le traitement automatique des langues et l'extraction d'information en tant que des approches interdisciplinaires.

Références

Apostel, Léo, Benoist Jean-Marie, Bottomore Tom Burton, Boulding Kenneth Ewart Dufrenne Mikel, Eliade Mircea, Furtado. Celso, Gusdorf Georges, Krishna Daya Wolfgang J. Mommsen, Morin Edgar, Piattelli-Palmarini Massimo, Sinaceur Mohammed d Allai, Stanislav Nikolaevitch Smirn, (1983), *Interdisciplinarité et sciences humaines Volume I, Unesco.*

Bourguignon, André, (1996), « L'approche transdisciplinaire. De la pluridisciplinarité à la transdisciplinarité ».

Billaud, Jean-Paul, « De l'objet de l'interdisciplinarité à l'interdisciplinarité autour des objets », *UMR Ladyss-CNRS, Université Paris-X, 2000 avenue de la République.*

Calenge, Bertrand, (2002), « A la recherche de l'interdisciplinarité », *Bulletin des bibliothèques de France.*

Cori, Marcel et Marandin, Jean-Morie, (2001), « La linguistique au contact de l'informatique : de la construction des grammaires aux grammaires de construction. *Histoire Epistémologie, Langage, volume 23, N°1.*

Habert, Benoît. (2002), *Outils des linguistes/outiller la linguistique : par où, par qui commencer ? Intervention à la table ronde TAL et enseignement, TALN°2.*

Marie-Andrée Bertrand Gilles Bibeau Violaine Crevier Hélène David Hubert Doucet Lucie Gélinau Olivette Genest Jacques Hamel Sunita Kapila Diane Laflamme Yves Lenoir Frédéric Lesemann Creutzer Mathurin Robert Moher Basarab Nicolescu Édité par Lucie Gélinau avec la collaboration de Carole Mailloux, (Octobre 2002), « L'interdisciplinarité et la recherche sociale appliquée. Réflexions sur des expériences en cours. »

Morin, Edgar, (1994), « Sur l'interdisciplinarité », *Bulletin interactif du Centre International de Recherche et Etudes transdisciplinaires, n°2.*

Morin, Edgar, (1999), *Acquisition des patrons lexico-syntaxiques caractéristiques d'une relation sémantique, Traitement automatique des Langues.*

Patrick Charaudeau, "Pour une interdisciplinarité "focalisée" dans les sciences humaines et sociales", in *Questions de Communication*, (201, consulté le 28 mai 2017 sur le site de Patrick Charaudeau - Livres, articles, publications.

Poibeau, T. & Nazarenko, A. (1999), « L'extraction d'information, une nouvelle conception. »

Vinck, Dominique, (2000), *Pratiques de l'interdisciplinarité. Mutations des sciences, de l'industrie et de l'enseignement, PUG-COLLECTION : GENIE INDUSTRIEL.*

Vinck, Dominique, (2008), « Dispositifs et pratiques d'interdisciplinarité SPI/SHS en école d'ingénieurs », *HAL*.

IJHCS