



Syntaxe et sémantique en corpus. Du temps et de la modalité en français oral

Lotfi Abouda

► To cite this version:

Lotfi Abouda. Syntaxe et sémantique en corpus. Du temps et de la modalité en français oral. Sciences de l'Homme et Société. Université d'Orléans (UO), 2015. tel-04375996

HAL Id: tel-04375996

<https://shs.hal.science/tel-04375996>

Submitted on 6 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Habilitation à Diriger des Recherches
Discipline : Sciences du Langage - Linguistique

Syntaxe et Sémantique en corpus
Du temps et de la modalité en français oral

Dossier scientifique présenté par

Lotfi ABOUDA

le 28 novembre 2015

Vol. 1 Dossier de synthèse

JURY

Gabriel BERGOUNIOUX Professeur à l'Université d'Orléans
Marie Luisa DONAIRE (rapporteure), Professeure à l'Universidad de Oviedo
Annette GERSTENBERG Professeure à la Freie Universität Berlin
Laurent GOSSELIN (rapporteur), Professeur à l'université de Rouen
Pierre Patrick HAILLET (rapporteur), Professeur à l'Université de Cergy-Pontoise

Sommaire

Sommaire	7
Introduction	10
1. Du statut catégoriel du conditionnel à l'hypothèse de l'identification syntaxique des modes : à la recherche du paradigme	14
1.1. Temps ou Mode ?	15
1.1.1. Les arguments classiques	15
1.1.1.1. Arguments diachroniques	15
1.1.1.2. Arguments morphologiques	17
1.1.1.3. Arguments analogiques	21
1.1.2. Arguments syntaxiques	23
1.2. La visibilité syntaxique des modes	33
1.2.1. Le modèle théorique	33
1.2.2. De la visibilité des modes, et de l'invisibilité des temps en syntaxe	34
1.2.2.1. Domaine d'influence	36
1.2.2.2. L'impossible modal vs l'inacceptable temporel	38
1.2.2.3. Alternance, concordance, liberté	40
2. Une vision de l'indicatif : propositions	44
2.1. Préliminaires	46
2.1.1. Paradigmes de paradigmes	46
2.1.2. Mode vs temps VS Modalité vs temporalité	47
2.1.3. Traitement de la polysémie temporelle	49
2.2. Deux groupes de formes temporelles	51
2.2.1. Morphologie de l'indicatif	51
2.2.2. Arguments pour une bipartition	54
2.2.2.1. Une piste ancienne mais actuelle	54
2.2.2.2. De l'opposition PR/IMP	55
2.2.2.3. Portrait-robot d'une proposition	62
2.3. De la polyphonie à la polychronie	78
2.3.1. Polyphonie linguistique	79
2.3.2. De la polyphonie à la polychronie : un modèle de traitement de la polysémie temporelle ?	82
2.3.2.1. Le problème et un diagnostic	83

2.3.2.2.	Polychronie	84
2.3.2.3.	Articulation avec les mécanismes de résolution des conflits	85
3.	Retour aux données.....	88
3.1.	La linguistique par l'exemple	91
3.1.1.	<i>Pierre, Marie et la pomme</i>	91
3.1.2.	De Stendhal à Martin	95
3.2.	La linguistique en corpus	99
3.2.1.	<i>Les ESLOs</i>	99
3.2.1.1.	De l'écrit à l'oral	101
3.2.1.2.	De la représentativité au concept de données situées	104
3.2.1.3.	Des corpus fantômes aux corpus disponibles	109
3.2.2.	<i>Etude de la temporalité en corpus</i>	111
3.2.2.1.	Le présent de l'omelette	111
3.2.2.2.	Distribution futur simple / futur périphrastique	113
3.2.2.2.1.	Constitution du sous-corpus d'étude	114
3.2.2.2.2.	TXM	118
3.2.2.2.3.	Annotation	121
	Conclusion générale.....	135
	Annexe 1 : Corpus futur.....	140
	Sélection des locuteurs des entretiens d'ESLO1	140
	Sélection des locuteurs des entretiens d'ESLO2	142
	Annexe 2 : Extrait d'un tableau CSV en cours d'annotation (avant réinjection sous TXM).....	145
	Annexe 3 : Exemple de requête sous TXM (après réinjection de l'annotation)	146
	Bibliographie.....	147
	Liste des travaux	159
	Index des noms	161
	Index des notions	163

Syntaxe et Sémantique en corpus. Du temps et de la modalité en français oral

Introduction

C'est sans doute pour préserver la liberté de cet espace singulier qu'est le mémoire de synthèse d'une HDR, que le genre de ce discours reste si peu défini. Même en mémorisant en détail tout un parcours de chercheur sur plus de vingt ans, en percevant les thèmes et leur organisation, les méthodologies et leur évolution, on peut rester longtemps dans l'expectative, ne sachant par où commencer, incertain sur les façons de dire, le ton à adopter, le degré d'abstraction à privilégier...

Exercice difficile, car il oblige à réviser sur vingt années le choix d'objets et de méthodes, les hypothèses heureuses et les égarements. Il oblige aussi à essayer de mettre en rapport l'enseignant-chercheur avec sa réalité sociale, les conditions et la nature de son travail, les rencontres stimulantes et le temps perdu, les publications et les travaux inachevés.

Exercice nécessaire, car il contraint, dans une époque où tout s'accélère, de s'arrêter un instant, pour tenter d'établir un bilan, à mi-parcours, afin de mieux se projeter dans l'avenir.

Un tel bilan sera forcément sélectif. Il tentera de restituer ici trois étapes importantes de mon parcours depuis mes recherches doctorales.

Ma thèse de doctorat, soutenue en octobre 1997, avait pour titre *Syntaxe et Sémantique du conditionnel en français moderne*. Au cours de ce travail, différentes questions liées aux phénomènes de la temporalité et de la modalité ont pu être posées, à commencer par la question du statut catégoriel de l'objet de recherche, i.e. le conditionnel. Pour résumer la problématique en une alternative : s'agit-il d'un mode ou d'un temps ?

Question classique, mais toujours d'actualité, comme peuvent en témoigner les flottements qu'on peut observer dans certaines grammaires scolaires, même si, aujourd'hui, suivant en cela la quasi-totalité des linguistes, nombre de grammaires de référence, y compris le *Bon Usage* – il serait à ce propos instructif de suivre le traitement que propose cet ouvrage au fil de ses éditions –, acceptent de le considérer comme une forme de l'indicatif. Cette question sera abordée dans la première partie de cette synthèse, où, après la présentation de ma contribution à ce débat (§ 1.1.2), il sera question (§ 1.2.2) de présenter l'hypothèse, plus générale, de la visibilité syntaxique des modes verbaux.

Or, si le conditionnel est une forme de l'indicatif, la question se pose de sa place dans le système de l'indicatif, ses rapports avec les autres formes, notamment, sur ce point la morphologie et la diachronie sont explicites, l'imparfait et le futur.

Mes recherches sur le conditionnel m'ont conduit à des travaux sur le futur et l'imparfait, qui, à leur tour, m'ont amené à l'étude du présent... Ce cheminement, qui sera résumé dans la deuxième partie de ce travail, s'il n'est pas habituel – le parcours classique commencerait par le présent pour terminer par le conditionnel, à l'image de ce que l'on peut observer dans les travaux généralistes sur les temps de l'indicatif – me semble à la fois plus riche et paradoxalement plus simple. J'aurai l'occasion de montrer (§ 2.2) que dans la distribution des marques morphologiques, l'approche illustrée dans cette présentation permet de restituer, à travers l'absence de marque du présent, non pas l'effet résultant de la concaténation d'un seul zéro mais de deux. Le présent est doublement neutre, dans sa dimension chronologique (ni passé, ni futur) et sur le plan énonciatif (le locuteur n'introduit pas de distance subjective par rapport à son énoncé). De cette dernière propriété, appréhendée dans le cadre d'un double paradigme temporel, des conséquences seront tirées sur le statut des autres formes temporelles sans remettre en cause leur appartenance catégorielle à l'indicatif.

La détermination de la valeur en langue d'une forme morphologique donnée, en l'occurrence le conditionnel, ne peut se faire en dehors de sa disposition dans le système, ce qui oblige à des recherches sur les formes verbales adjacentes. Mais elle ne peut pas non plus être envisagée en dehors d'une autre relation paradigmatique, autrement plus complexe parce que plus abstraite et dégagée de tout support morphologique, qui lie la valeur en langue aux différents emplois en contexte de la forme considérée.

Chacune des formes verbales étudiées est susceptible d'une multitude d'emplois, que la tradition grammaticale qualifie de particuliers (le présent narratif ou des recettes, l'imparfait d'atténuation, le futur conjectural etc.). A ce phénomène, très général, de polysémie (ou de distorsion) temporelle, qui constitue un défi majeur pour la plupart des modèles descriptifs et explicatifs de la temporalité, une analyse classique ne propose que des solutions particulières et généralement *ad hoc*, où, très souvent, l'unité du signifié linguistique ne peut être préservée, ou n'est conservée que par artifice.

C'est dans ce cadre que s'inscrit la proposition avancée dans ma thèse et développée dans des publications ultérieures (§ 2.3), celle d'une théorie polychronique qui se voudrait un prolongement de la polyphonie dans le domaine de la temporalité. Si, dans ses versions les plus radicales, le modèle polychronique ne peut plus être conservé, je tenterai dans cette synthèse d'en proposer une version, aux contours plus limités, qui se voudrait complémentaire des mécanismes de résolution des conflits (Gosselin, 1996).

Au-delà de la nécessité d'un cadre théorique explicatif global pour la polysémie temporelle, il me paraît désormais indispensable d'asseoir le débat scientifique sur les données (§ 3) : comment distinguer les emplois, les hiérarchiser, les quantifier, les décrire avec le plus de précision possible ?

Avec l'émergence de corpus oraux (3.2), dont ESLO (§ 3.2.1), la confrontation théorique peut être justifiée par des données authentiques disponibles en masse, qualifiées par un appareil de métadonnées qui permet de les situer aux plans diaphasique et diatratique. Avec les instruments mis à disposition par la communauté du traitement automatique du langage (TAL), grâce au développement d'un nombre d'outils d'interrogation et de traitements dont l'efficacité et la robustesse s'accroissent, il est possible d'exploiter de façon ergonomique et économique les ressources à disposition.

Le champ sémantique de la temporalité peut tirer profit des perspectives offertes par la technologie. Mes premières recherches menées dans ce cadre sur le présent (§ 3.2.2.1) et sur les deux formes synthétique et analytique du futur (§ 3.2.2.2) constituent un encouragement en ce sens.

On remarquera dans cette trajectoire une constance vis-à-vis d'une thématique générale, celle de la temporalité. Mais les approches méthodologiques successivement adoptées ont considérablement évolué, en partant de considérations strictement théoriques pour arriver aux données. S'il m'est possible d'assumer cette démarche, voire de la revendiquer, on peut surtout y voir un parcours, personnel et social, qui affirme la nécessité d'un retour aux données et s'en donne les moyens.

1. Du statut catégoriel du conditionnel à l'hypothèse de l'identification syntaxique des modes : à la recherche du paradigme

L'admission officielle du conditionnel au sein de l'indicatif a été à la fois très discutée et très tardive. Figurant encore comme un mode distinct (en plus de l'indicatif, du subjonctif et de l'impératif) dans la circulaire du 25 juillet 1975, il n'apparaîtra dans la liste des temps de l'indicatif, sans autres commentaires, que dans la Terminologie grammaticale de 1997 (rééditée en 1998)¹. C'est sans doute à cause de cette admission tardive et quelque peu discrète² que le conditionnel est encore souvent considéré comme un mode dans la plupart des grammaires scolaires.

1.1. Temps ou Mode ?

Si, en contraste avec la situation institutionnelle et didactique, les linguistes ont été nombreux, et depuis longtemps, à considérer le conditionnel comme une forme de l'indicatif (Brunot (1905 & 1922) Damourette & Pichon (1911-1940) ; G. Guillaume (1929) ; R.-L. Wagner (1939)...), il ne m'avait pas semblé inutile, au moment d'entamer mes recherches doctorales, de revenir sur la question. D'une part, parce que les arguments habituellement avancés à l'appui du traitement du conditionnel comme une forme de l'indicatif n'étaient pas aussi concluants, et, d'autre part, il était central de statuer sur la catégorie du conditionnel pour identifier le paradigme, temporel ou modal, dans lequel il puise en partie ses spécificités.

1.1.1. Les arguments classiques

Ainsi que j'ai essayé de le montrer dans Abouda (1997a et 1997b), l'examen détaillé des arguments, que j'ai pu qualifier de classiques, traditionnellement invoqués à l'appui du traitement du conditionnel comme un temps de l'indicatif, montre qu'ils ne sont pas exempts de certaines failles.

1.1.1.1. Arguments diachroniques

Si, à l'exception notable de Lanly (1958 & 2002)³, tous les historiens de la langue s'accordent sur l'origine étymologique du conditionnel⁴, ils divergent sur la manière d'interpréter la genèse de cette forme et sur le rôle qu'il conviendrait de lui attribuer dans

¹ Voir à ce propos D. Hisquin (2000).

² La comparaison des différentes éditions du *Bon Usage* (de 1936, date de la première édition à 2011, date de la 15^e édition) est à ce propos particulièrement instructive.

³ Combattant point par point, sur le plan phonétique, syntaxique et sémantique, la thèse majoritaire, Lanly considère que le conditionnel a pour origine le subjonctif imparfait latin. Même si ses analyses n'ont pas été relayées par d'autres travaux, il me semble que les arguments qu'il avance méritent d'être connus et examinés.

⁴ Le conditionnel est issu en roman de la périphrase du latin populaire composée de l'infinitif d'un verbe et de l'auxiliaire *avoir* à l'imparfait (*legere habebam* = 'j'avais à lire'). « De cette périphrase qui contenait un imparfait, notent Georges Le Bidois & Robert Le Bidois (1935-1938 : 457), sortit, à l'époque gallo-romaine, une nouvelle forme verbale, *lirais*, où se reconnaît, grâce à la désinence, la présence de l'imparfait (*av*)*ais* associé à un infinitif. »

l'identification du statut catégoriel du conditionnel. Pour certains, comme Wagner et Pinchon (1962)⁵, le conditionnel serait un temps de l'indicatif, au même titre que le futur, puisque les deux formes sont issues selon un mécanisme morphologique similaire. Sauf que la similitude des mécanismes de génération peut ne pas suffire à garantir l'appartenance du conditionnel à l'indicatif, ne serait-ce que parce qu'on peut penser que futur et conditionnel constituent tous les deux un mode à part, ainsi que le fait Yvon (1951) qui regroupe les deux formes au sein d'un mode qu'il baptise *suppositif*... Pour d'autres, la question ne pouvait être tranchée diachroniquement qu'au niveau des emplois. S'ensuivit un débat qui opposera deux groupes de grammairiens, cherchant chacun à identifier les attestations historiques les plus anciennes qui viendraient appuyer ses vues. Ainsi, pour Brunot & Bruneau (1956)⁶ et Georges Le Bidois et Robert Le Bidois (1935-1938), le conditionnel serait un temps parce qu'il serait d'abord apparu dans les propositions subordonnées dans la dépendance d'un verbe employé au passé. Il servirait à l'origine uniquement à engager l'avenir par rapport à un moment du passé choisi comme repère et serait initialement dépourvue de toute valeur modale. Ce n'est qu'à un stade plus avancé qu'à cette valeur temporelle primitive se serait surajoutée une valeur modale, d'abord hypothétique. Pour Sechehaye et Lerch⁷, au contraire, le conditionnel serait un mode puisque dès la périphrase latine dont il est issu, il existait une nuance modale, très nette, de nécessité dans le passé. Pour ceux-ci, le propre du conditionnel à l'origine, puisqu'il serait d'abord apparu dans les 'principales' des phrases hypothétiques, était de rejeter l'idée de l'obligation ou du devoir dans le passé relatif au présent du locuteur. L'emploi dit temporel du conditionnel ne serait apparu que plus tard, lorsque la nuance modale, s'affaiblissant progressivement, aurait cédé la place à un résidu simplement temporel, rendant ainsi possible l'apparition du conditionnel dans les subordonnées dépendantes d'un verbe principal au passé.

Discutant en détail les arguments de chacune des deux approches, Wagner finira par les renvoyer dos à dos, en refusant les termes mêmes du débat. Il écrira (1939 : 42) :

Mais qu'on pose le type *Se + impft indicatif*... forme en *-rais* comme contemporain des origines de la langue, c'est-à-dire comme aussi ancien que l'emploi de la forme en *-rais* en complétive dépendant d'un verbe principal au passé, note-t-il (1939 : 42), et le problème change dans sa donnée même. On n'a plus, en effet, à considérer la forme en *-rais* (forme de

⁵ Ils écrivent (1962 : 303) : « Historiquement, cette forme est de la même nature que le futur. Toutes deux sont issues, en roman, d'une périphrase composée de l'infinitif d'un verbe et du présent ou de l'imparfait de l'auxiliaire AVOIR. Si l'on fait du futur un temps de l'indicatif, comme il est naturel, il est normal de faire également du conditionnel un temps. Si l'on faisait du conditionnel un mode, il faudrait alors en faire un aussi du futur. Ces deux formes, solidaires, se définissent l'une par rapport à l'autre et toutes deux, en opposition avec le subjonctif, actualisent dans la durée le procès. Elles appartiennent donc à l'indicatif. »

⁶ Avant d'être totalement refondé par Brunot et Bruneau en 1933, l'ouvrage avait d'abord été publié sous le seul nom de Brunot en 1899.

⁷ Cités dans Wagner (1939).

création romane) comme le sujet d'une des plus étonnantes évolutions (de temps en mode ou inversement) qui soit ; elle s'impose à nous, au contraire, sous les traits d'une forme qui contenait en elle, de par sa nature même, la possibilité de deux emplois parallèles ».

Même si, au moment de rédiger ma thèse, je n'avais pas connaissance des travaux sur la grammaticalisation⁸, domaine qui connaîtra quelques années plus tard un essor considérable⁹, je n'avais pas souscrit au point de vue de Wagner qui récusait par principe le phénomène, que de très nombreux travaux finiront par identifier comme l'un des mécanismes majeurs de la grammaticalisation, i.e. la recatégorisation. M'appuyant alors sur quelques remarques glanées dans le champ de la temporalité¹⁰, j'ai fini par conclure (1997a : 268) qu'on ne pouvait pas « exclure d'une façon apriorique le passage d'une forme donnée d'une catégorie à une autre, que cette catégorie soit d'ailleurs grammaticale ou syntaxique ». Restait à dater précisément les emplois du conditionnel, tâche qui, en plus d'être hors de portée dans le cadre d'une recherche doctorale synchronique, était particulièrement délicate et difficile. Avec l'arrivée de l'ère numérique, qui interviendra quelques années plus tard, la numérisation de manuscrits anciens et leur mise à disposition ergonomique a permis le retour de la question de la datation des emplois du conditionnel, redevenue aujourd'hui d'actualité comme en témoigne le nombre considérable de travaux récents¹¹ et les projets de recherche en cours...

En bref, l'argumentation diachronique, au-delà de tous les questionnements méthodologiques qu'elle soulève et des incertitudes liées à la reconstruction théorique d'une forme dont on n'avait pas d'attestations, ne m'avait pas semblé décisive pour identifier le statut catégoriel synchronique du conditionnel, d'autant que celui-ci pouvait avoir subi une recatégorisation qui l'aurait fait quitter un espace catégoriel pour un autre.

1.1.1.2. *Arguments morphologiques*

Sans revenir dans ce cadre sur l'épineuse question du découpage morphologique des verbes au futur et au conditionnel¹², qui, quelle que soit l'analyse retenue, ne remettra pas en

⁸ Notamment Fleischman (1983) et (1992), Bybee, pagliuca & Perkins (1991), Bybee, Perkins & pagliuca (1994).

⁹ Voir notamment Heine & Kuteva (2002).

¹⁰ J'avais alors évoqué, pour la récuser, l'analyse de Lyons (1970) qui affirmait ainsi la nature modale des auxiliaires futurs anglais, *will* et *shall* sous prétexte qu'ils étaient à l'origine des auxiliaires modaux, ainsi que l'analyse de N. Davidsen-Nielsen (1988), qui, s'appuyant sur le cas du subjonctif plus-que-parfait latin *venisset* qui a donné en roumain la forme *venise*, considérée aujourd'hui comme une forme de l'indicatif plus-que-parfait, affirmait que les constatations diachroniques ne pouvaient pas déterminer le statut synchronique des formes du futur.

¹¹ Baeyens (2012), Bourova & Dendale (2006), Dendale (2014), Patard & De Mulder (2012), Patard, Grabar & De Mulder (2015)

¹² Discutant des différents arguments avancés par les deux approches, Camussi-Ni écrit (2006 : 47) :

« En définitive, conserver le R dans la base, c'est créer un allomorphisme de la base qui ne se trouverait que dans deux variantes : le conditionnel présent et le futur. C'est affirmer la corrélation entre cet allomorphe +R et ces

cause l'essentiel de ce qui suit, j'essaierai de focaliser mes propos sur l'argumentation morphologique investie dans de nombreux travaux en faveur de l'hypothèse du traitement du conditionnel comme temps de l'indicatif.

Du point de vue de sa constitution morphologique, note ainsi Imbs (1960 : 41), le futur français est apparenté au 'conditionnel'. Tous les deux sont des formes en *-r-*. A cet *-r-* s'ajoutent des terminaisons de présent (*-ai*, *-as*, etc.) et l'on obtient le futur, ou des terminaisons d'imparfait (*-ais*, *-ais*, etc.) et l'on obtient le 'conditionnel'.

Touratier écrira quant à lui (1996 : 38) :

Il est difficile de ne pas retrouver dans le conditionnel d'une part la marque de l'imparfait et d'autre part la marque /R/ du futur, comme le montrent les deux paires minimales :
/nu ʃát-əR-i-ʃ/ ~ / nu ʃát-əR-ʃ / ; nous chant-er-i-ons ~ nous chant-er-ons
/ nu ʃát-əR-i-ʃ / ~ / nu ʃát-i-ʃ / ; nous chant-er-i-ons ~ nous chant-i-ons
Ceci veut dire qu'au point de vue morphologique, le conditionnel dit présent a tout l'air d'être un futur imparfait et donc d'appartenir aux temps de l'indicatif. Car si le futur est un temps de l'indicatif et l'imparfait un autre temps de l'indicatif, on ne voit pas comment la combinaison de ces deux temps de l'indicatif pourrait ne pas appartenir aussi au mode indicatif.

Or, à y voir de plus près, cette analyse, aussi répandue soit-elle, n'est pas exempte de critiques susceptibles de remettre en cause la validité de l'argumentation morphologique.

Celle-ci, en plus d'être paradoxale, pourrait en effet paraître défectueuse, si elle s'accommodait de la vision habituelle du système de l'indicatif. Paradoxale, parce que par le simple fait de voir dans la marque du conditionnel la combinaison de deux morphèmes temporels, cette vision l'exclut du paradigme disjonctif de l'indicatif, dans lequel sont censées s'opposer directement les 5 formes temporelles simples (et leurs correspondants complexes), i.e. présent (PR), futur (FUT), passé simple (PS), imparfait (IMP)... et conditionnel (COND). Autrement dit, l'hypothèse de la compositionnalité du conditionnel, si elle devait être prise au sérieux, aurait pour conséquence paradoxale de ne plus admettre cette forme dans le même paradigme temporel que l'imparfait et le futur, et, donc, de porter un coup à l'analyse du conditionnel comme un temps de l'indicatif.

Ce type d'argumentation pourrait même paraître défectueux, et interroger la notion elle-même de morphème. Si le morphème est la plus petite unité linguistique porteuse de sens, on ne devrait pas considérer la marque du conditionnel comme un morphème tout en y voyant la

deux temps, ce qui équivaut à dire que le matériau R participe de la dénotation du futur et du conditionnel présent. Puisqu'il est pertinent d'isoler un suffixe dans l'infinitif, puisque le suffixe R du conditionnel est porteur d'un sème spécifique en français moderne comme le montre l'opposition du conditionnel et de l'imparfait, on ne voit pas l'intérêt qu'on peut avoir à éviter de le découper. On court plutôt le risque alors d'opacifier l'analyse du futur et du conditionnel.

C'est pourquoi nous postulerons le découpage du R qui peut permettre de poser un suffixe commun au futur, au conditionnel et à l'infinitif. »

combinaison de deux autres morphèmes, ceux de l'IMP et du Fut. Car cela reviendrait à attribuer à cette forme deux structures internes concurrentes que l'on peut représenter ici sous la forme simplifiée suivante (faisant notamment abstraction du Mode et de l'Aspect) :

(1) A. Radical + R + ε + PN	B. Radical + temps + PN ¹³
	Ø PR
	e PS
	R FUT
	ε IMP
	R ε COND

L'abandon de l'une ou l'autre des deux lectures A/B m'a ainsi semblé incontournable. Généralement, c'est le traitement A et le(s) schéma(s) morphologique(s) qu'il implique qui se trouvent, sinon abandonnés, du moins largement sous-exploités : mise en avant quand il est question de la genèse du conditionnel ou de son statut catégoriel, la compositionnalité de cette forme n'est que rarement réinvestie au niveau notionnel.

Les raisons de cette sous-exploitation sont nombreuses. D'abord, elle implique un bouleversement non seulement de la manière de percevoir le système temporel de l'indicatif, mais aussi de la catégorie habituelle du temps. Nous y reviendrons. Ensuite, elle donne aux morphèmes zéros un rôle central dans l'organisation des formes temporelles. Or, un principe tacite existe, qui consiste à se méfier du zéro. Et même quand on est obligé d'admettre le zéro, c'est souvent à titre provisoire, en adoptant l'attitude, décrite par Gleason (1969), qui consiste à éliminer de la description finale les zéros introduits provisoirement au cours de l'analyse¹⁴. Camussi-Ni, discutant les travaux classiques en morphologie structurale, écrit (2006 : 61) :

C'est parce que la marque est assimilée à du matériau et non pas à un rapport entre matériaux qu'il paraît difficilement concevable que le matériau zéro soit une marque. Or, précisément le zéro n'est pas une marque, pas plus que ne l'est le matériau auquel il s'oppose, c'est l'opposition significative entre la présence et l'absence de matériau qui constitue la marque¹⁵.

Ce sont des principes généralement admis, notamment quand il est question de décrire des langues « exotiques ». Leur adoption rencontre une certaine résistance lorsqu'on cherche à les appliquer à des langues à forte tradition écrite.

¹³ Personne-Nombre

¹⁴ Le conseil est suivi à la lettre par Pinchon & Couté (1981) qui font littéralement disparaître le zéro, qu'ils avaient admis dans leur description du futur et du conditionnel, pour le remplacer dans les tableaux par des cases grisées, et en généralisant le système de la sorte (1981 : 41) : « Entre la base et la désinence peut s'insérer un élément ou deux éléments ».

¹⁵ Voir également A. Lemaréchal (1990 & 2001).

En bref, l'argumentation morphologique m'a semblé dans une impasse : elle devait soit être abandonnée¹⁶, soit prise au sérieux à condition que toutes ses conséquences en soient tirées. A quelques exceptions notables près, notamment Damourette & Pichon et Coseriu, le prix, une remise en cause totale de l'organisation globale du système de l'indicatif, était jugé trop exorbitant par la plupart des grammairiens et linguistes, qui se sont globalement contentés d'utiliser l'argumentation morphologique sans se poser la question des conséquences théoriques et pratiques d'une telle utilisation.

Mais l'adoption du schéma morphologique (A) permettrait-elle de sauver l'argumentation morphologique ?

Il m'a semblé assez rapidement que l'adoption du schéma (A), si elle était nécessaire, n'était pas suffisante pour asseoir sur des bases morphologiques solides le traitement du conditionnel comme un temps de l'indicatif.

A supposer que soient résolus tous les problèmes, nombreux et complexes – nous y reviendrons –, qu'un tel schéma implique, son adoption n'écarte pas longtemps les traitements concurrents.

Concrètement, le schéma (A-1) peut être généralisé comme sous (2) :

(2) Radical + X + Y + PN

Notre lecture, qui sera développée plus loin¹⁷, plaide en faveur de la fission de la catégorie traditionnelle *temps* en deux catégories distinctes, X et Y. Si elles sont conçues comme des catégories temporelles, le maintien du conditionnel au sein de l'indicatif peut être envisagé. Mais cette interprétation n'est pas la seule possible. On pourrait aussi penser que X est une catégorie de Mode, et non de Temps. Dès lors, rien ne nous empêche de considérer¹⁸, comme

¹⁶ L'abandon de l'hypothèse de la compositionnalité du conditionnel ne remet pas automatiquement en cause son appartenance à l'indicatif, si on parvient par ailleurs à la justifier. On pourrait en effet appréhender la forme en *-rait* comme une marque indécomposable, support formel d'un seul et unique morphème, faisant partie d'un paradigme où il s'oppose aux divers autres temps de l'indicatif, y compris le futur et l'imparfait. Comme le souligne Touratier (1996 : 61), « en vertu du caractère arbitraire du signe linguistique, la langue peut fort bien utiliser la même unité morphologique comme signifiant de deux morphèmes différents, que l'on qualifie alors d'homonymes ». Cela signifie concrètement que la marque *-r-* qui existe dans le morphème du conditionnel pourrait être considérée comme distincte du morphème *-r-* du futur...

¹⁷ § 2.2.1, p. 32 sq.

¹⁸ D'autant que les deux formes entretiennent des relations sémantiques indéniables, ainsi que le rappellent Wagner et Pinchon (1962 : 303) : « Il est vrai que le futur et le conditionnel évoquant l'avenir se chargent parfois de plus de *valeurs affectives* que les temps relatifs au présent ou au passé. C'est pour cette raison que les grammairiens qui considèrent les modes comme des *formes de la sensibilité* ont été conduits à traiter du conditionnel comme d'un mode. Pour être conséquents, ces grammairiens devraient aussi considérer le futur comme un mode. »

le fait par exemple H. Yvon (1951)¹⁹, la marque *-R-* sous X, commune au futur et au conditionnel, comme le morphème d'un Mode qu'on peut appeler *Suppositif*. Le système se déclinerait alors de la manière abrégée et simplifiée suivante :

(3)

	Mode		Temps	
Radical	indicatif	∅	∅	PR
			- ε -	IMP
	suppositif	-R ²⁰ -	∅	FUT
			- ε -	COND

Même si elle ne sera pas adoptée ici, cette vision me paraît tout à fait envisageable, d'autant que le statut modal ou temporel du futur lui-même a toujours fait débat...

Ce rapide tour d'horizon aura montré, je l'espère, que l'argumentation morphologique, aussi répandue soit-elle, ne résiste pas à un examen détaillé des faits et ne permet pas de montrer d'une manière indiscutable l'appartenance du conditionnel à l'indicatif.

1.1.1.3. Arguments analogiques

Si de nombreuses formes verbales ont connu une discussion de leur statut catégoriel (le futur²¹, l'indicatif²², l'impératif²³...), aucune discussion n'a été aussi longue et animée que celle qu'a connue le conditionnel. Outre la complexité des faits diachroniques et morphologiques, et la richesse sémantique de la forme sur laquelle je reviendrai, la raison principale de cette situation s'explique par l'ambiguïté de la question elle-même qui interroge jusqu'aux notions de Mode et de Temps.

Que voudrait-on en effet déterminer lorsqu'on s'interroge sur le statut temporel et / ou modal du conditionnel ? L'examen des différentes réponses montre que la question n'a pas été invariablement interprétée relativement au statut catégoriel du conditionnel, parce que si tel avait été le cas, on aurait obtenu deux types de traitements linguistiques concurrents dont l'un donnerait à la forme en *-rait* le statut de Mode et l'autre le statut de Temps. Or, en plus de ces

¹⁹ L'auteur reconnaît (1951 : 271) que l'idée d'opposer le 'mode' futur à l'indicatif remonte au moins à Destutt de Tracy (1803) qui opposait les formes de l'existence positive aux formes de l'existence éventuelle (futur, conditionnel, simples et composés).

²⁰ Se reposera alors la question de savoir si ce *-R-* est différent de celui de l'infinitif...

²¹ Voir par exemple Veters (1996).

²² Voir par exemple Lyons (1970).

²³ Voir par exemple Guillaume (1929).

deux types de réponses, et même en faisant abstraction de quelques autres²⁴, on constate l'existence d'au moins deux autres approches assez répandues. La première, que j'ai qualifiée de cumulative, défendue, entre autres, par Wartburg & Zumthor (1958), Cornelis de Boer (1954), Korzen & Nölke (1990), reconnaît l'existence d'un 'conditionnel-temps' et d'un 'conditionnel-mode', sans que soit toujours précisé le sens donné à la notion de Mode (catégorie ou emploi ?)²⁵. La seconde approche, que j'ai appelée *disjonctive*, considère le conditionnel comme *tantôt* mode *tantôt* temps. Tout aussi répandue que la première (défendue entre autres par P. Imbs (1960), Chevalier *et al.* (1964), Arrivé, Gadet & Galmiche (1986)), elle instaure une catégorie unifiée en langue, laissant au contexte la tâche de préciser le statut d'occurrence du conditionnel²⁶.

Au cœur du questionnement se trouvent ainsi les emplois.

Dire, à l'instar de R. Martin (1983 : 133) que le conditionnel est un temps de l'indicatif, ne se distinguant des autres temps de l'indicatif que par « le nombre important de ses emplois modaux » n'aura pas suffi à convaincre tout le monde. De même que l'on attend d'un temps qu'il localise les procès dans le Temps, on attend d'un verbe qui exprime de la modalité qu'il soit un mode. Sans doute trouvera-t-on exagéré le poids accordé ici à cette justification, surtout qu'elle n'est jamais explicitement revendiquée, et qu'elle est même quelquefois clairement rejetée, sous la forme de mises en garde, pour moi significatives, rencontrées ici ou là, sur la nécessité de distinguer Mode et Modalité, temps linguistique et Temps extra-linguistique. Il n'empêche que l'on peut débusquer le raisonnement dans un nombre considérable de travaux sur les temps²⁷ et sur les modes²⁸.

²⁴ Comme par exemple celle de Mauger qui écrira (1968 : 231, n.1) : « En fait, il nous semble que le conditionnel n'est ni un mode à part ni un temps particulier, ou plutôt qu'il est à la fois, et selon les cas, mode et temps ».

²⁵ L'hésitation entre Mode en tant que catégorie et valeur modale est par exemple particulièrement visible chez K. Nyrop (1930 : 300) qui considère que le conditionnel « présente en même temps une valeur temporelle et une valeur modale ; il appartient ainsi non seulement à l'indicatif, mais aussi au subjonctif dans les cas où il exprime la possibilité d'une action conditionnelle ou l'incertitude d'une action. »

²⁶ C'est, par exemple, la position de P. Imbs quand il écrit (op. cit. : 199-200) : « Au niveau de la langue, le morphème *-r-*, qui termine le radical des formes verbales en question, oriente celle-ci vers une valeur d'éventualité ; à partir de ce radical les formes bifurquent. Une première bifurcation est celle qui mène soit à des formes en *-ai* soit à des formes en *-ais*. À l'intérieur des formes en *-ais* une seconde bifurcation, relative à la valeur, seulement, conduit soit à l'emploi temporel, soit à l'emploi modal : c'est le contexte de la phrase qui indique le choix. En relation avec un processus au passé avec lequel elle est mise en perspective, la forme en *-rais* devient spécifiquement temporelle ; en relation directe avec un *si*, elle devient modale ; sans relation aucune, elle suit la même pente, qui est celle de sa valeur congénitale d'éventuel, laquelle reste alors la sienne jusqu'au bout de son actualisation. »

²⁷ Voir Weinrich (1973) et sa critique de ce qu'il appelle les preuves *e nomine*.

²⁸ A-t-on, par exemple, oublié que J. Lyons (1970 : 235), face à un indicatif dépourvu de « forte valeur modale », ne lui attribue le statut de mode que « par simple courtoisie » ?!

En bref, le conditionnel, parce qu'il est susceptible d'avoir un nombre important d'emplois modaux, en types et en fréquence, devrait être considéré comme un Mode distinct, et non pas comme une forme de l'indicatif.

A quoi, utilisant une argumentation *analogique*, on pourrait rétorquer qu'il suffirait d'observer les divers emplois des autres formes verbales, et notamment le futur et l'imparfait, pour s'apercevoir que les valeurs dites modales sont loin de concerner uniquement le conditionnel. Pourquoi, se fondant par exemple sur les emplois hypothétique, hypocoristique et atténuatif, aucun grammairien ni aucun linguiste n'a, me semble-t-il, proposé pour l'imparfait le statut de Mode ? Le traitement du conditionnel comme mode sur la base de la multitude de ses emplois modaux, s'appuyant sur un raisonnement *ad hoc*, paraît ainsi fragilisé.

Toutefois, faut-il le préciser, l'argumentation analogique ne justifie pas le traitement du conditionnel comme un temps de l'indicatif. Elle ne fait que neutraliser l'un des arguments adverses. La question de l'appartenance catégorielle du conditionnel reste donc posée.

1.1.2. Arguments syntaxiques

Ayant décidé de travailler sur corpus²⁹, j'ai entrepris, lors de l'une des premières phases de ma recherche doctorale, d'examiner exhaustivement, en les classant syntaxiquement, les phrases où apparaît un conditionnel.

Au-delà des problèmes inhérents à toute classification – s'agissant en l'occurrence d'une démarche inductive, au moins dans ses premières phases, il était nécessaire d'enrichir continuellement la classification en classes et sous-classes nouvelles, et, donc de restructurer sans cesse l'ensemble –, la tâche, en plus d'être fastidieuse, m'avait d'abord paru sans intérêt. Jusqu'à l'émergence d'une hypothèse qui allait donner un sens et un but à une classification initialement strictement descriptive. Et si, en l'absence d'arguments diachroniques, morphologiques et analogiques probants, la syntaxe pouvait aider à identifier le statut catégoriel du conditionnel ?

La réponse à cette question, en plus d'inscrire avec certitude le conditionnel dans le paradigme de l'indicatif, donnera naissance à l'une des hypothèses les plus fortes de mon

²⁹ Le corpus utilisé était composé de 2200 occurrences, dont une centaine d'exemples oraux, et 1898 exemples journalistiques puisés dans le Corpus de Stockholm (le *COSTO*). « Avec un corpus global contenant plus de 2200 exemples appartenant à des sources aussi diversifiées, avais-je alors écrit (1997a : 288), nous espérons que notre description des structures syntaxiques du conditionnel sera la plus exhaustive possible. » J'y reviendrai longuement tout au long de ce travail, et notamment dans sa 3^e partie, sur la question centrale des données...

travail de thèse, même si elle était à la marge de son objet d'étude : l'hypothèse de la localisation syntaxique des modes, qui sera présentée et développée dans nombre de mes publications ultérieures (1998a, 1998b, 1999, 2002 et 2010). L'idée principale était que chaque mode verbal possède en propre au moins une structure syntaxique.

Or, si chaque mode verbal est reconnaissable par au moins une structure syntaxique exclusive, il devient possible, lors de l'inventaire exhaustif des phrases au conditionnel, de se prononcer sur le statut catégoriel de cette forme, en procédant, parallèlement à l'inventaire, à un test systématique de commutation dont l'objectif est de savoir si chacune des occurrences observées peut ou non être remplacée par une autre forme verbale et si oui de quel type. Puisque la question du statut catégoriel du conditionnel concerne en réalité son appartenance, ou non, au mode indicatif, l'objet du test était de savoir si les occurrences au conditionnel étaient syntaxiquement obligatoires ou si elles pouvaient systématiquement être remplacées par d'autres formes de l'indicatif dans le même environnement syntaxique. Autrement dit, il était question de savoir si les structures syntaxiques du conditionnel formaient un sous-ensemble des structures de l'indicatif. Concrètement, j'ai essayé de remplacer chacune de mes occurrences au conditionnel par une forme de l'indicatif et une forme du subjonctif.

Notons au passage un point qui sera développé dans la troisième partie de cette synthèse, concernant le rapport aux données : en combinant inventaire et classification en fonction d'une hypothèse, en admettant à côté de données réelles fournies par le corpus, des données manipulées, je combinais le travail sur corpus à un travail sur le possible de langue.

Je ne reviendrai pas dans ce cadre sur les critères de ma classification syntaxique, présentée dans le chapitre 1 de la thèse (pp. 10-85), et justifiée dans le chapitre 4 (notamment pp. 288-304), ni sur l'inventaire lui-même, disponible dans la thèse et dans Abouda (1997b). Je me contenterai ici, après une brève remarque méthodologique sur la classification, de présenter rapidement ses résultats, qui seront mis en perspective et problématisés dans la section suivante.

L'une des difficultés rencontrées lors de la classification a consisté à identifier la bonne granularité. Il fallait en effet d'une part garantir le plus haut degré de généralité possible, sans laquelle l'apparition du conditionnel dans une structure et surtout son remplacement par une autre forme temporelle ou modale perdraient toute pertinence au niveau syntaxique, et, de l'autre, être sûr que la classification ne laisse de côté aucun sous-type syntaxique pouvant avoir un rôle dans la sélection de la forme morphologique du verbe. En bref, la pertinence

d'une classification se mesurant exclusivement par rapport à l'objectif qu'elle se fixe, la granularité pertinente était celle qui en l'occurrence faisait appel aux paramètres intervenant dans le choix des modes, et, si possible³⁰, uniquement ceux-là.

La question est naturellement délicate, et interroge ce qu'on entend par syntaxe, et par paramètre syntaxique. Partant d'une conception strictement architecturale de la syntaxe, l'examen des données m'aura progressivement poussé à la lexicaliser (ou à l'enrichir en traits lexicaux). Jusqu'à quel point ces traits n'étaient pas syntaxiques ? Ce sera la question essentielle qui sera abordée dans la suite de cette section.

Les résultats bruts de l'inventaire des structures syntaxiques pouvaient paraître de prime abord décevants, puisque le test de remplacement du conditionnel par une forme verbale de l'indicatif et / ou du subjonctif a permis l'identification des quatre cas théoriques possibles :

Cas 1 : le conditionnel ne peut être remplacé que par une forme du subjonctif ;

Cas 2 : le conditionnel peut indifféremment être remplacé par l'indicatif ou par le subjonctif ;

Cas 3 : le conditionnel est obligatoire et ne peut être remplacé ni par un indicatif ni par un subjonctif ;

Cas 4 : le conditionnel ne peut être remplacé que par une forme de l'indicatif.

L'examen précis de chacun de ces cas montre toutefois une très nette tendance : le conditionnel partage très généralement les mêmes structures syntaxiques que l'indicatif.

En effet, le cas 1, qui aurait sans doute fourni les contre-exemples les plus difficiles à admettre dans le cadre de nos hypothèses, doit être aujourd'hui rejeté : il n'a pas pu être obtenu par manipulation directe de nos propres données, mais par une reconstruction, que l'on peut juger aujourd'hui discutable, faisant le rapprochement syntaxique entre certaines occurrences de notre corpus, comme dans :

(4) (a) J'aurais les moyens de m'installer à Paris que je ne le ferais pas. (Costo)

(b) *J'avais les moyens de m'installer à Paris que je ne le ferais pas.

³⁰ Plusieurs sous-types syntaxiques ont toutefois été identifiés dans Abouda (1997a et b), bien qu'ils ne jouent pas de rôle dans la sélection des modes verbaux, pour des raisons qualifiées de « pédagogiques »...

et des exemples, empruntés à la grammaire, et qu'elle juge elle-même archaïques et /ou appartenant à un registre littéraire soutenu, comme les énoncés suivants que j'emprunte à M. Cohen (1965 : 64) :

(5) (a) Leur père eût été malade (que) ces mauvais enfants n'eussent pas appelé le médecin.

(b) Y eussions-nous consenti que les Anglais auraient pris notre place en invoquant des nécessités stratégiques (De Gaulle).

L'argumentation sous-jacente suivie lors de la rédaction de la thèse consistait à affirmer que, malgré l'absence d'occurrences de ce type dans notre corpus, si, à la place du conditionnel simple dans la protase³¹ de la phrase (4)a on avait un conditionnel 'passé', il aurait été difficile de soutenir qu'il n'était pas substituable par le fameux 'conditionnel passé deuxième forme', le plus-que-parfait du subjonctif (*j'eusse eu les moyens...*). Pour plusieurs raisons méthodologiques, qui s'éclaireront progressivement au cours de cette synthèse, on fera aujourd'hui abstraction de ces cas.

Le cas 2, qui concerne un nombre nettement plus important d'exemples, me semble plus intéressant. Si on fait abstraction, de quelques survivances de tours archaïques, comme dans :

(6) (a) Je ne sache pas...

(b) Qui l'eût cru ?

les cas où le conditionnel peut céder la place à la fois à une forme du subjonctif et à un (autre) temps de l'indicatif se rencontrent essentiellement à la fois dans les relatives³², comme dans :

(7) (a) Klaus Barbie a découragé la pitié de ceux de ses compatriotes qui auraient pu avoir des faiblesses pour un vieillard jeté en pâture à la justice vengeresse des vainqueurs. (Costo)

(b) Klaus Barbie a découragé la pitié de ceux de ses compatriotes qui pouvaient avoir des faiblesses pour un vieillard ...

(c) Klaus Barbie a découragé la pitié de ceux de ses compatriotes qui aient pu avoir des faiblesses pour un vieillard ...

et dans certaines complétives (notamment dans la dépendance d'un verbe d'opinion employé négativement ou interrogativement³³) :

(8) (a) Je ne pensais pas qu'on pourrait encore l'utiliser. (MO4)

³¹ On emploiera ici les termes protase/apodose, qui ont le mérite d'une certaine stabilité et clarté.

³² Cas extrêmement commenté dans la littérature grammaticale et linguistique.

³³ Voir Huot (1986, 1987 et 1991) et Abouda (1998b et 2002).

(b) Je ne pensais pas qu'on pouvait encore l'utiliser.

(c) Je ne pensais pas qu'on puisse / pût encore l'utiliser.

Ces emplois, présentant des cas de *variation modale*³⁴, loin de remettre en cause les hypothèses avancées, montrent au contraire, que c'est parce qu'il est une forme de l'indicatif que le conditionnel peut se trouver dans les structures où on rencontre l'indicatif, fussent-elles partagées avec le subjonctif.

Le cas 3, où le conditionnel est obligatoire et n'est remplaçable ni par un subjonctif ni par un indicatif, me semble de loin le plus intéressant et le plus stimulant, à la fois par ses implications théoriques et par la richesse sémantico-syntaxique des occurrences qu'il met en exergue, dont je ne donnerai ici que quelques aperçus.

Un cas de figure intéressant se présente avec le verbe *dire*, gouvernant un GN, comme dans :

(9) (a) On dirait un papillon (costo)

(b) *On dira / *on a dit / *on disait un papillon³⁵.

Lorsque la position objet de *dire* est occupée par une complétive (cf. *on dirait qu'il a terminé son travail*), la substitution au conditionnel d'une autre forme de l'indicatif est possible. De même, lorsqu'on remplace le GN *un papillon* par un GN pouvant être sélectionné par *dire* au sens de *proférer* (*dire une chose, un mot, un mensonge...*), la substitution est possible. Il n'est pas sûr toutefois que ces manipulations préservent la structure initiale, parce qu'en modifiant l'objet, comme lors de la dernière manipulation, on modifie la structure argumentale du verbe, qui aura à nouveau glissé vers le sens de *proférer*. Se posera ainsi la question de la nature, syntaxique ou non, du caractère obligatoire du conditionnel dans ce genre d'environnements. Cela engage la conception elle-même que l'on se fait de la syntaxe. Nous y reviendrons.

Un autre cas sur lequel on peut revenir se présente avec le verbe *aimer* gouvernant une infinitive, comme dans les exemples suivants :

(10) (a) A deux reprises, deux policiers, déguisés en mécaniciens, montent à bord de l'avion. Ils aimeraient maîtriser le pirate, mais l'occasion ne se présente pas. (Costo)

³⁴ Voir Abouda (1998b). L'acceptation de l'indicatif dans cet environnement syntaxique – de nombreux tests d'acceptabilité menés depuis quelques années auprès des étudiants le montrent – ne va pas de soi. Un travail sur des corpus différents (oral et écrit, appartenant à différents genres) pourrait préciser la situation actuelle. Une étudiante de master2 est pressentie pour mener une telle étude dans le cadre de ses futures recherches doctorales.

³⁵ On aura noté qu'employé en mention, ce GN autorise le remplacement du conditionnel par une autre forme de l'indicatif. Mais on n'aura pas ainsi préservé la structure initiale, invalidant le test tel qu'il est annoncé (la substitution ne vaut que si elle est opérée dans un environnement syntaxique identique).

- (a') J'aimerais vous faire part de quelques remarques à ce propos.
(Costo)

où le remplacement du conditionnel par un autre temps de l'indicatif est impossible. Dans le cadre d'une certaine conception de la syntaxe, on peut douter de la nature syntaxique de cette impossibilité puisqu'en remplaçant, dans un environnement similaire, *aimer* par *vouloir*, verbe sémantiquement proche, on rend possible la substitution au conditionnel d'autres temps de l'indicatif. Cette impossibilité de substitution semble d'autant moins dictée par des considérations syntaxiques qu'on peut rencontrer aisément des cas où, dans un environnement syntaxiquement identique, le remplacement du conditionnel par un autre temps de l'indicatif ne rencontre pas la moindre résistance:

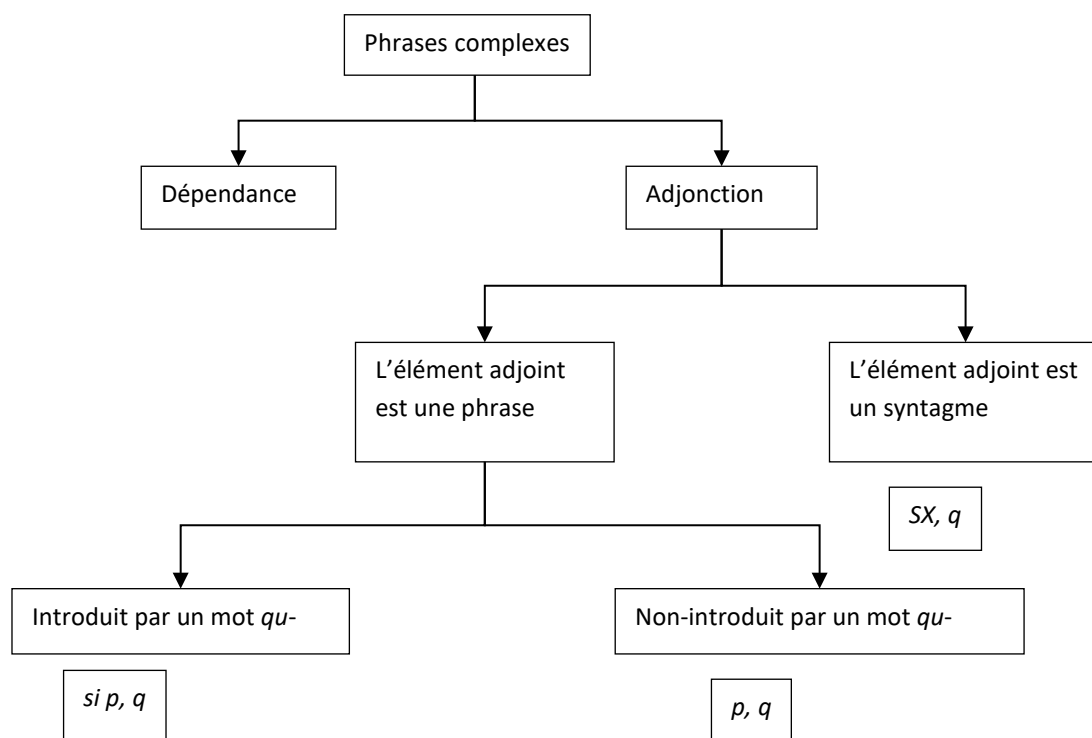
- (11) (a) J'aimerais marcher au bord de la Loire.

- (b) J'aimais / aime marcher au bord de la Loire.

Si, aujourd'hui, il me paraît possible que la résistance constatée soit en rapport avec la combinaison des aspects lexicaux du verbe *aimer* et de l'infinitif qu'il gouverne, c'est l'hypothèse de la performativité qui avait d'abord été envisagée³⁶. Ces traitements écartent ainsi toute considération syntaxique susceptible d'expliquer la non-substituabilité par un indicatif du conditionnel dans ce contexte. Ces observations seront le point de départ d'un examen, initié au chapitre 6 de la thèse, des verbes désidératifs et de leur fonctionnement dans les énoncés atténuatifs, et qui donneront une série de publications à la fois sur le conditionnel (Abouda 2001) et sur l'imparfait (Abouda 2004a et b).

Le dernier exemple, nettement plus redoutable, de contextes où le conditionnel est obligatoire, qu'on aimerait évoquer se présente dans certains types de phrases complexes liées par un lien syntaxique lâche (on parle dans les modèles syntaxiques formels de rapport de modification, d'adjonction ou d'ajout, selon l'école, qu'on oppose à une stricte relation de dépendance). Plusieurs classes et sous-classes doivent être distinguées. Afin de clarifier le débat, on peut schématiser ces sous-classifications de la manière suivante :

³⁶ Les deux paradigmes explicatifs ne sont sans doute pas sans lien. Si la différence entre *Je voudrais me promener au bord de la Loire* et *Je veux me promener au bord de la Loire* peut être décrite par un trait [$\pm$ atténuation], c'est l'opposition [performatif/constatif] qui décrit le mieux la différence entre *J'aimerais me promener au bord de la Loire* et *J'aime me promener au bord de la Loire*. Or il semblerait que le verbe *aimer* constatif au présent se combine plus naturellement avec les verbes infinitifs atéliques, même si la combinaison avec un verbe télique peut être sauvée par des déformations aspectuelles : *J'aime partir* peut ainsi recevoir une lecture itérative ou un glissement sémantique de *partir*. Sans doute le trait [$\pm$ ponctuel] joue-t-il également un rôle.



La première distinction pertinente opposera deux classes principales selon que l'élément adjoint (modifieur) est un simple syntagme ou une phrase conjuguée. La première classe peut être illustrée par les deux séries suivantes d'exemples, extraits du Costo³⁷ :

- (12) (a) Sans Gérard, j'aurais peut-être une gentille petite chaîne hôtelière.
 (b) Or, dans l'hypothèse d'une provocation délibérée, M. Reagan pourrait difficilement ne pas réagir avec fermeté [...].
 (c) Ah ! Evidemment, chez ta belle-mère, ça n'arriverait pas !
 (d) Une chance de vivre en l'an 2000 ? A leur âge, j'aurais été plus pessimiste car la révolution me paraissait inéluctable [...].
 (e) L'affaire Waldheim a pris un mauvais tour, et mérite d'être examinée d'un point de vue plus large. Autrichien, je n'aurais pas voté pour lui.
- (13) (a) En Grande Bretagne, nous n'aurions pas à justifier sa survie.
 (b) A la place de Gomez, j'aurais fait le même choix.

Si, dans la première série, le remplacement du conditionnel par une forme de l'indicatif me paraît possible, il est plus difficile dans la deuxième série. Nous y reviendrons.

Dans la seconde sous-classe où l'élément adjoint est une phrase conjuguée, il convient d'opérer une nouvelle classification syntaxique qui opposera cette fois les cas où la phrase

³⁷ Il convient de préciser que les exemples choisis n'illustrent pas tous les sous-cas distingués dans Abouda (1997a), mais seulement ceux qui éclairent la classification en relation avec le point discuté ici.

adjointe est introduite par un mot *qu-* de type « subordonnant libre »³⁸, et les cas où la phrase adjointe n'est pas introduite par un subordonnant. La première sous-classe peut être illustrée par les extraits suivants :

- (14) (a) « Si les musulmans nous avaient donné leurs filles en mariage, le Maghreb serait devenu un autre Brésil ». (Costo)
- (b) Si ces images avaient été tournées par la télévision libyenne, vous les auriez diffusées tout de même ? (Costo)
- (c) On pourrait être autre chose que ce qu'on est si les parents nous avaient laissé un peu. (Eslo1)

Le remplacement ou non du conditionnel par une autre forme de l'indicatif dans ces structures obéit-il en l'occurrence à des contraintes syntaxiques ? Quelquefois possible comme dans (14)c et même (14)b, me semble-t-il, impossible comme dans (14)a, la substitution obéit à une contrainte qui ne peut pas être décrite en termes strictement syntaxiques. Non seulement la structure syntaxique de type *Si p, q* admet toutes les combinaisons possibles des formes temporelles de l'indicatif, au prix quelquefois, il est vrai, de la disparition de la lecture hypothétique, mais, même dans les occurrences qui manifestent le plus de résistance à la substitution, comme l'exemple (14)a ci-dessus, il est possible de remplacer le conditionnel par un autre temps de l'indicatif, dès lors qu'on accepte de considérer que la substitution de la forme morphologique du verbe de la protase n'altère pas la structure syntaxique. Cela revient à dire que la relation d'accord morphologique à distance qui lie ce couple mythique de la grammaire, plus-que-parfait+conditionnel passé, n'est pas syntaxique. On aura l'occasion ci-dessous d'établir une distinction importante entre les chemins syntaxiques qu'empruntent les traits d'héritage de mode et les chemins qu'empruntent les traits de percolation de temps. Ces derniers, nous le verrons, n'ont aucune des propriétés syntaxiques habituelles, mais s'apparentent aux contraintes logico-sémantiques comme celles qui rendent quelquefois impossibles certaines combinaisons entre formes morphologiques du verbe et certains adverbiaux temporels.

Lorsque la phrase adjointe n'est pas introduite par un mot *qu-*, cas que l'on peut illustrer par les exemples suivants du Costo :

- (15) (a) « Serais-je demain pape, j'aurais des rapports avec le diable », disait Malraux.

³⁸ En gros les conjonctions de subordination et les fameuses locutions conjonctives. Voir à ce propos Milner (1989). Ici, ce cas sera illustré uniquement par les structures de type *si p, q*, les seules pertinentes pour notre propos. Voir à ce propos Anscombe (2004) et Haillet (2007) et (2009).

(b) J'aurais les moyens de m'installer à Paris que je ne le ferais pas.

(c) N'oubliez pas de le citer, vous lui feriez de la peine.

le conditionnel de l'apodose ne peut pas toujours être facilement remplacé par une autre forme de l'indicatif. Ces phrases, que la tradition grammaticale identifie comme relevant de la subordination implicite, et que je range dans les structures en adjonction, se rapprochent des phrases de type *si p, q*, ne s'en distinguant que par deux propriétés à notre avis liées : l'absence de subordonnant et l'absence de mobilité³⁹. L'analyse proposée pour les phrases en *si p, q* s'appliquera ici puisqu'aucune contrainte syntaxique ne rend impossible une phrase de type :

(16) Il vient, je quitte immédiatement la pièce !

Revenons maintenant aux exemples (12) et (13) vus ci-dessus, que je reproduis ci-dessous par commodité sous (17) et (18) :

(17) (a) Sans Gérard, j'aurais peut-être une gentille petite chaîne hôtelière.

(b) Or, dans l'hypothèse d'une provocation délibérée, M. Reagan pourrait difficilement ne pas réagir avec fermeté [...].

(c) Ah ! Evidemment, chez ta belle-mère, ça n'arriverait pas !

(d) Une chance de vivre en l'an 2000 ? A leur âge, j'aurais été plus pessimiste car la révolution me paraissait inéluctable [...].

(e) L'affaire Waldheim a pris un mauvais tour, et mérite d'être examinée d'un point de vue plus large. Autrichien, je n'aurais pas voté pour lui.

Si remplacement, jugé possible, du conditionnel par une forme de l'indicatif ne présente pas la moindre difficulté et n'altère pas la nature hypothétique de la phrase dans les exemples (17)a (le conditionnel peut être remplacé par un imparfait) et (17)b et c (où il peut être remplacé par un futur), une telle substitution, si elle ne touche pas à la structure syntaxique de la phrase, provoque dans les exemples (17)d-e une altération du type sémantique de ces énoncés qui perdent avec cette substitution leur sens hypothétique. Le remplacement du conditionnel par un plus-que-parfait transforme en effet le syntagme prépositionnel *à leur âge* en un simple circonstanciel temporel et l'adjectif *autrichien* en une expression causale⁴⁰.

³⁹ Or, comme l'écrit Milner (1989 : 513), la : « subordination est fondée sur les positions ; la parataxe est fondée sur les places. »

⁴⁰ Le même phénomène s'observe dans les phrases suivantes, où il semble que c'est le conditionnel qui pousse vers une interprétation hypothétique du circonstanciel, simplement temporel dans (b) :

(a) Il y a vingt ans, j'aurais été moins raisonnable qu'eux : par exemple, j'aurais dit que je voulais travailler moins et gagner plus. (Costo)

(b) Il y a vingt ans, j'étais moins raisonnable qu'eux...

Ce paradigme explicatif, qu'il serait long à développer ici mais que l'on peut facilement déduire de ce qui précède, permet, à notre avis, de comprendre la difficulté de remplacer le conditionnel dans (18) :

(18) (a) En Grande Bretagne, nous n'aurions pas à justifier sa survie.

(b) A la place de Gomez, j'aurais fait le même choix.

Le cas 4, d'un conditionnel pouvant aisément être remplacé par un indicatif et uniquement par un indicatif, est, et de loin – il ne faudrait pas que la longue discussion du cas 3 nous fasse croire le contraire –, le cas le plus fréquent, en occurrences et en types de structures syntaxiques : hormis quelques cas de figement ou d'archaïsmes rencontrés sporadiquement ici ou là, et si on fait abstraction des emplois discutés dans les cas 2 et 3, le remplacement du conditionnel par un indicatif est possible dans la quasi-totalité des classes syntaxiques identifiées. La conclusion s'imposera d'elle-même : le conditionnel, n'ayant aucune structure *syntactique* exclusive, n'est pas un mode ; pouvant dans chaque contexte *syntactique* être remplacé par (au moins) une forme de l'indicatif, il sera considéré comme un temps de l'indicatif. En effet, cette identité des structures constatée serait bien singulière, voire surprenante, si l'on maintenait l'hypothèse selon laquelle le conditionnel serait un mode distinct de l'indicatif.

1.2. La visibilité syntaxique des modes

Dans ce qui précède, pour appuyer l'hypothèse de l'appartenance du conditionnel à la catégorie du temps, j'ai affirmé qu'il n'avait aucune structure *syntaxique* exclusive. Pour la plupart des cas étudiés, cette affirmation, ne faisant qu'enregistrer sur un plan descriptif des tendances claires, paraît assez consensuelle et peut être adoptée quel que soit le courant théorique. Pour certaines autres cas étudiés – notamment certains des emplois regroupés dans le cas 3 où le conditionnel est obligatoire – l'adoption d'une telle hypothèse engage une vision particulière de la syntaxe et requiert le choix d'un modèle théorique précis.

1.2.1. Le modèle théorique

S'il était légitime pour un doctorant qui, dans un champ syntaxique conflictuel, allait s'attaquer à un marché du travail extrêmement concurrentiel, d'opter pour une classification syntaxique descriptive et prudente, j'avais conscience que l'inventaire « aussi descriptif soit-il, se prononce fatalement sur un nombre considérable de problèmes syntaxiques » (1997a : 303), et qu'il ne pouvait de toute façon pas s'opérer en dehors de tout modèle théorique : « Pour rendre visibles les données pertinentes et les présenter d'une façon claire et économique, il est nécessaire de disposer d'un modèle syntaxique précis » (ibid., 12).

Mon choix se portera sur une version de la grammaire générative, développée par Chomsky durant les années 1980 et connu sous le nom de Gouvernement & Liage. Ce choix a été justifié à la fois par sa familiarité – il est vrai que j'étais doctorant à Paris 7 – et par son hégémonie épistémologique soulignée par Milner (1989), dans laquelle je ne voyais pas uniquement des inconvénients, mais aussi l'avantage d'offrir au doctorant que j'étais une quantité de travaux disponibles et qui, à défaut de cohérence, pouvaient permettre à celui qui acceptait de faire le tri le développement d'une vue globale tout en fournissant des traitements descriptifs de la quasi-totalité des phénomènes syntaxiques rencontrés *en français*. Même si, face à la grande diversité des travaux et à la richesse des débats théoriques, j'avais tenu à développer une vision personnelle du modèle, largement influencée par Milner, et à laquelle j'ai consacré la totalité du premier chapitre de la thèse (pp. 10-85), ce choix a été moins coûteux que celui de l'un des formalismes grammaticaux qui arrivaient en France⁴¹. D'une part, parce qu'ils m'étaient moins familiers, malgré des enseignements suivis pendant mon année de DEA à Paris 7 et Paris 8, et, de l'autre, parce que leur adoption m'aurait contraint à

⁴¹ Essentiellement les grammaires d'unification, et notamment HPSG et TAG. Voir Pollard & Sag (1994) et Abeillé (1993).

me prononcer, sans l'aide de travaux disponibles, sur un nombre incalculable de phénomènes syntaxiques du français, alors que la syntaxe n'était pas au cœur de mon objet d'étude.

Il me paraît aujourd'hui possible, avec un recul d'une vingtaine d'années, d'admettre que c'est l'adoption de ce modèle qui a permis l'émergence de l'hypothèse de la visibilité syntaxique des modes. Celle-ci présuppose en effet une syntaxe positionnelle que Milner (1989 : 456 & 644 sq.) qualifiera à juste titre de « quantitativement pauvre » et « qualitativement grossière » (*ibid* 645).

Depuis, ce choix a quelquefois été critiqué à la fois par des générativistes, qui trouvaient que mes vues ne correspondaient pas toujours à des versions plus actuelles du modèle, et ils n'ont pas toujours tort, et surtout par les non-générativistes, pour ne pas dire anti-générativistes, particulièrement nombreux au sein des écoles sémantiques françaises, pour des raisons historiques et épistémologiques non sans rapport avec le processus de légitimation de la sémantique en tant que domaine plus récent, sans doute mal traité au sein du courant générativiste depuis ses débuts...

Je reviendrai, dans la 3^e partie de cette synthèse, sur quelques considérations théoriques et méthodologiques relatives à l'adoption d'un tel modèle syntaxique formel et son degré de compatibilité avec une démarche sur corpus.

1.2.2. De la visibilité des modes, et de l'invisibilité des temps en syntaxe

« Pourquoi, demanda Milner (1989 : 664), ne pas supposer que la syntaxe, prise en elle-même, constitue un mode particulier de perception, avec ses propriétés caractéristiques ? [...] On dira que la syntaxe « voit » ou « ne voit pas » tel ou tel élément, dans telles ou telles conditions. Les propriétés générales que nous avons proposées deviennent alors des lois de la perception syntaxique. »

L'hypothèse de la visibilité syntaxique des modes, s'inspirant théoriquement de Milner, part d'un constat simple, qui se mesure mieux dans la comparaison avec les phénomènes temporels : contrairement aux formes temporelles, qui semblent constituer des objets logico-sémantiques, que la syntaxe ne voit pas, les modes verbaux, obéissant à des principes syntaxiques, ont une visibilité de structures qui les rend à la fois descriptibles et prédictibles en syntaxe.

L'hypothèse d'un lien entre syntaxe et mode verbal, si elle n'est pas totalement originale⁴², n'a à notre connaissance jamais été systématisée et appelée à jouer un rôle équivalent à celui qu'on lui attribue ici.

Contrairement à ce qu'on pouvait souvent lire, et depuis longtemps, sur le caractère obligatoire de certaines structures au subjonctif⁴³, l'étude, menée dans le chapitre 2 de ma thèse (pp. 86-174), des modes infinitif, impératif, subjonctif et indicatif montre clairement que le caractère syntaxiquement obligatoire⁴⁴ constitue, non pas une spécificité du subjonctif, mais une propriété générale de tous les modes verbaux, dont chacun possède des structures syntaxiques exclusives.

Je ne reviendrai pas en détail sur cette macro-hypothèse et les analyses qu'elle a nécessitées, parce qu'elles ont été développées dans la thèse dans nombre de publications ultérieures (Abouda 1998a, 1998b, 1999, 2002 et 2010), préférant me concentrer ici, d'une part, sur le point le plus problématique, à savoir la nature non-syntaxique des éléments qui imposent l'emploi du conditionnel dans certains cas, et, d'autre part, les conséquences de cette vision sur la conception de l'objet *temps*.

La question qui me paraît fondamentale est d'identifier la nature de l'élément qui rend le conditionnel obligatoire dans un certain nombre de cas. Face à la complexité des données et l'enchevêtrement des paramètres, la nature syntaxique ou sémantique des contraintes discutées est largement tributaire de la vision que l'on se fait de la syntaxe. L'une des pistes argumentaires, suivie dans Abouda (2010), consiste à comparer, point par point, les contraintes qui régissent la sélection modale et celles qui interviennent dans le choix du temps verbal. C'est ce parcours, parce qu'il est économique et pédagogique, que je me propose de réemprunter ici brièvement.

⁴² Lors de la soutenance de ma thèse, G. Kleiber a indiqué que cette hypothèse avait déjà été avancée dans un texte de 1972 de B. Combettes, que je n'avais jamais réussi à retrouver depuis. Je fais ici surtout référence aux très nombreuses analyses, partielles, avancées dans différents cadres à propos des liens entre tel ou tel mode et telle ou telle structure syntaxique (les constructions du subjonctif chez les grammairiens classiques, les infinitives dans le cadre générativiste...).

⁴³ Observant le lien qui lie le subjonctif à la notion syntaxique de dépendance (lat. *subjungere* : mettre sous la dépendance de...), beaucoup de grammairiens en conclurent qu'il s'agit d'une « servitude grammaticale »...

⁴⁴ On se rappelle de débats qui, de Beauzée (1767) à Lyons (1970), en passant par Brunot (1922), Bally (1932), ou Togeby (1982), auront mobilisé beaucoup d'énergie en s'appuyant sur un principe rarement discuté : ce qui est requis par la forme, n'a pas de sens... Même Damourette & Pichon, parce qu'ils n'avaient pas remis en cause la validité de ce raisonnement, pour défendre la spécificité sémantique du subjonctif (« non-jugement »), s'acharneront à montrer que son emploi n'est pas imposé par la construction. Ils écriront (1911-1927, t. V : § 1869) : « Quant à la syntaxe, elle ne décèlerait une décadence du subjonctif que si l'emploi de ce mœuf était devenu si mécanique, si automatiquement attiré par la forme de la phrase qu'il n'y eût plus aucun intérêt à conserver une forme spéciale, puisque cette forme n'aurait plus aucune signification propre ».

En comparant les différents contextes isolés ci-dessus où le conditionnel n'est pas substituable (cas 3) aux paramètres qui rendent l'emploi de tel ou tel mode obligatoire, l'on se rend aisément compte qu'ils ont des caractéristiques totalement opposées.

1.2.2.1. *Domaine d'influence*

La première différence importante concerne le domaine où s'exerce l'influence des éléments qui interviennent dans la sélection des modes et ceux qui interviennent dans la sélection des temps.

Si l'on fait abstraction de quelques cas, considérés comme peu significatifs (notamment des tours archaïques et/ou figés), à chaque fois que le conditionnel est obligatoire, nous l'avons vu, l'élément qui impose son emploi occupe une position syntaxique adjointe (adverbiaux circonstanciels, protases de phrases hypothétiques...), jamais une position appartenant à strictement parler à la « proposition »⁴⁵. Le mode verbal en revanche n'est jamais obligatoire qu'à cause d'un certain nombre d'éléments qui sont soit *internes* à la proposition elle-même, soit externes mais entretenant un rapport de *domination stricte* avec la proposition contenant la forme verbale en question. Les données étant connues et bien documentées, on trouvera aisément dans la littérature grammaticale et linguistique des exemples qui appuient cette analyse. On se contentera d'évoquer, en guise d'illustration, le cas, développé dans Abouda (2002), d'alternance modale indicatif/subjonctif dans les complétives dépendantes d'un verbe d'opinion employé négativement ou interrogativement (*crois-tu/je ne crois pas...*). Il y a été démontré non seulement que les traits sémantiques responsables de la sélection modale doivent obligatoirement être assignés à des *têtes* syntaxiques, mais en plus que l'héritage ou la transmission de ces traits doit obligatoirement suivre des chemins syntaxiques précis, i.e. qui relèvent de la complémentation (*domination stricte*).

Dans le même ordre d'idées, a-t-on besoin de le rappeler ici, le phénomène de la sélection modale s'exerce dans la limite syntaxique de la phrase et ne la dépasse jamais, c'est-à-dire qu'aucune phrase n'impose le mode verbal d'une phrase voisine. Or, il a été constaté depuis longtemps, que dans le domaine de la temporalité, il arrive⁴⁶ que le choix d'une forme

⁴⁵ Terme souvent et à juste titre critiqué en syntaxe mais que j'emploie par commodité, parce qu'il allie en l'occurrence économie et clarté.

⁴⁶ Cela ne signifie ni qu'il s'agit là d'une condition nécessaire, ni que le choix d'une forme temporelle n'est jamais déterminé à l'intérieur du domaine de la phrase, ni même que cette détermination ne pourrait jamais emprunter des schémas syntaxiques qui relèvent de la dépendance (par exemple n'importe quelle forme aspectuo-temporelle ne peut pas dépendre de *pendant que...*). Notre propos consiste simplement à affirmer que, contrairement à la sélection modale, qui s'exerce exclusivement dans le cadre de relations syntaxiques de dépendance stricte, le choix des temps s'opère bien au-delà de ces limites...

temporelle dans une phrase donnée s'explique par le co-texte et les phrases qui la précèdent. Comment parvenir sinon à expliquer le plus-que-parfait dans l'exemple suivant, commenté par Kleiber (1993) ?

(19) Le voleur s'échappa. Il avait été prévenu par un indicateur.

C'est également le co-texte antérieur qui explique l'emploi, obligatoire, du conditionnel dans la série suivante⁴⁷ :

(20) Chirac ne peut pas me remplacer. Par qui le *ferait-il* ? Séguin ? Ce *serait* la guerre avec l'Allemagne. Balladur ? ça *ferait* bien rire, en tout cas ce *serait* bien trop tôt. Giscard ? Ce *serait* du plus haut comique. Je ne vois que Barre. Et ce *serait* un vrai désastre.

Si, dans le segment textuel en question, aucune occurrence du conditionnel ne peut céder la place à une autre forme de l'indicatif, la substitution devient possible dès que l'on extrait chacun de ces énoncés de son co-texte :

(21) Par qui le fera-t-il ? / ... l'a-t-il fait ? ...

Du reste, le recours à des notions comme l'anaphore temporelle⁴⁸ ou la cohérence textuelle⁴⁹, quel que soit par ailleurs leur degré de pertinence, montre que l'étude de la référence temporelle ne peut pas souvent faire abstraction de la dimension textuelle ou discursive, et doit donc dépasser le cadre de la phrase⁵⁰, qui constitue le domaine de la syntaxe.

Je pourrais multiplier les exemples d'études linguistiques⁵¹ faisant appel pour la description des temps à des outils qui, débordant le domaine de la phrase, relèvent clairement du champ textuel⁵². En bref, ainsi que le note Gosselin (1996: 35) :

La prise en compte de la dimension interphrastique (ou textuelle) apparaît comme une nécessité pour l'étude des phénomènes aspectuo-temporels. La plupart des travaux récents sur le domaine traitent, en effet, de l'enchaînement des temps et des procès dans le texte dans la perspective de l'**anaphore temporelle**.

⁴⁷ *Le Canard enchaîné*, 80^e année, n° 3910, 4 octobre 1995.

⁴⁸ Il serait impossible d'évoquer tous les travaux qui y recourent : voir, entre bien d'autres, Molendijk (1993 et 1996), Kleiber (1993)

⁴⁹ Voir par exemple Moeschler, Reboul, Luscher & Jayez (1994).

⁵⁰ Ainsi que le rappellent Benninger, Carlier & Lagae dans leur introduction de l'ouvrage intitulé *Temps et Texte* (2003 : 7), temps et texte constituent deux notions « intimement liées, dans la mesure où les temps verbaux contribuent à créer la trame narrative du récit. Aussi le fonctionnement des temps verbaux ne saurait-il être étudié dans le cadre trop restreint de la phrase, mais doit-il être analysé au niveau du texte. »

⁵¹ Voir également Battistelli (2009) et la notion de « boom textuel » dans le domaine de la temporalité en TAL.

⁵² Ou, plus généralement, à l'environnement discursif, qui, dans la perspective de Haillet (2007) demeure le domaine pertinent dans lequel s'opère l'agencement des représentations pour expliquer des phénomènes temporels ou proches de la temporalité (atténuation, ironie...).

Etant donné que le domaine de la syntaxe est la phrase et que la sélection du mode s'opère à l'intérieur de ce domaine, contrairement à ce qui peut se passer avec le temps, on peut voir dans ces faits un argument supplémentaire en faveur de l'hypothèse de la visibilité syntaxique des modes, et à l'invisibilité des temps en syntaxe. On verra ci-dessous (notamment 2.2.2.3 et les exemples oraux qui y sont examinés) que la différence est encore plus nette : l'explication d'un emploi de l'imparfait en interaction authentique ne peut réellement se faire que si on remonte bien loin dans l'échange...

1.2.2.2. *L'impossible modal vs l'inacceptable temporel*

La question que je me poserai concerne le type de rejet qui, lors d'un jugement différentiel, peut affecter un énoncé à cause de la forme morphologique du verbe-noyau. On peut y distinguer deux types, selon que l'inadéquation morphologique du verbe concerne le mode ou concerne le temps.

En effet, hormis quelques cas de flottement, identifiés⁵³, les jugements d'agrammaticalité dans le champ des modes verbaux ne présentent généralement pas de difficultés particulières. Notamment parce que ces jugements restent valables quel que soit le contexte d'emploi :

- (22) (a) *Je veux que tu viendrais
(b) *Marie parte
(c) *Je partir

De ce point de vue, ce type de rejet grammatical se rapproche de ce qu'on peut entendre par *impossible de langue*⁵⁴. L'impossibilité en question est similaire à celle qu'on peut reconstruire, en syntaxe, par des exemples fabriqués extrêmement simples, où loin des cas problématiques dont le statut purement syntaxique a pu par la suite être discuté, on inverse l'ordre requis des constituants (*une de Nirvana chanson à la radio joue) ou on supprime un constituant appelé par le verbe (*Le yéti offre à ses visiteurs), ou on tente d'injecter un constituant en trop (*je marche une pomme).

Lorsqu'on compare ces faits aux jugements qu'on rencontre dans le domaine des temps, on constate d'abord que dans, notre activité de linguiste, pour illustrer l'inacceptabilité

⁵³ Par exemple, quel mode verbal employer après *après que* ou après *le fait que* ? Le flottement s'explique largement par des contradictions entre les prescriptions grammaticales et la variété des usages...

⁵⁴ Voir notamment Milner (1989) et Auroux (1998) et Cori & David (2006).

temporelle d'un énoncé, on ne peut plus se contenter d'exemples aussi simples que ci-dessus : on est automatiquement amenés à les « enrichir », soit par l'adjonction de modificateurs :

- (23) (a) #Je viendrai hier
(b) #Il neigea demain
(c) #Hier, il a terminé son travail depuis deux heures

soit par l'apport de précisions contextuelles ou situationnelles⁵⁵. C'est ce que font, par exemple, Tasmowski-De Ryck et Vetters (1996), en expliquant le rejet de (39b) ci-dessous

- (24) (a) - Tu dormais?
(b) - # Tu as dormi?

par la précision contextuelle suivante :

- (25) Paul est à la maison. Il entre dans la chambre de son fils qui doit étudier et qui est installé à son bureau. Paul le voit se redresser comme un ressort et s'indigne

On aura remarqué que, dans les exemples ci-dessus, pour noter l'inacceptabilité dans le domaine de la référence temporelle, on a opté pour le signe dièse (#), proposé par Tasmowski-De Ryck et Vetters (1996) et utilisé dans Abouda (1997a), pour l'opposer à l'agrammaticalité : celle-ci, notée comme usuellement grâce à l'astérisque (*), tente de rendre compte de l'impossibilité de certaines phrases, quel qu'en soit le contexte, tandis que la première cherche à enregistrer une inadéquation, une incompatibilité *logico-sémantique* entre marqueurs, qui peuvent coexister ou non au sein du même énoncé. Or, il est évident que dans les énoncés (23) et (24), le rejet est dû à un phénomène de répulsion, de contradiction directe, entre des entités référentiellement incompatibles entre elles ou entre des marqueurs et une situation contextuelle précise.

Que l'on supprime les indications scéniques (25) et toutes les formes temporelles de l'indicatif deviennent possibles dans ce même environnement syntaxique :

- (26) Tu as dormi ? / Tu dormais ? Tu dormirais ?...

On peut trouver dans le modèle de Gosselin (1996 & 2005) une confirmation de ces vues, en rappelant brièvement le traitement qu'il réserve aux cas de conflit, qui peuvent surgir, dans

⁵⁵ Il est vrai que si ces précisions situationnelles ne permettent pas toujours de justifier un rejet (que l'on songe, dans la même situation, à la forme *tu t'es endormi* ?), elles ne permettent pas non plus de toujours sauver un énoncé. Haillet (2007 : 21), raconte que pour justifier leur acceptation de l'énoncé *Les bicyclettes sont en train d'avoir deux roues*, un groupe d'étudiants a été amené à imaginer toute une scène où ils pouvaient l'admettre. Ils n'ont été ramenés à la raison que lorsque l'enseignant leur a posé la question *Le diriez-vous spontanément ?*

le domaine temporel, soit entre les instructions sémantiques des marqueurs d'un énoncé, soit entre les instructions aspectuo-temporelles de ces marqueurs et des contraintes situationnelles.

Pour un énoncé donné, écrit-il (1996: 36), les instructions peuvent converger et donner directement lieu à une représentation cohérente, ou diverger, entrer en conflit, et contraindre le lecteur-auditeur à élaborer des stratégies interprétatives de construction de la signification visant à résoudre ces conflits. L'hypothèse qui fonde notre entreprise est que ces stratégies interprétatives peuvent être modélisées de façon prédictive ; autrement dit, qu'il est possible de décrire des modes de résolution de conflit ainsi que leurs conditions de déclenchement.

Ainsi, si on observe les deux exemples suivants :

(27) (a) Pierre a mangé en cinq minutes

(b) Pierre mange en cinq minutes

on constate que le second n'est acceptable que s'il reçoit soit une lecture itérative, comme en (28)a, soit, mais dans des contextes plus contraints, une lecture future comme en (28)b :

(28) (a) Généralement/ habituellement, Pierre mange en cinq minutes

(b) Pierre va manger en cinq minutes (et ensuite il va faire ceci ou cela)

Ces interprétations sont le résultat d'intervention de deux mécanismes de résolution qui donnent une issue à un conflit qui surgit « entre le circonstanciel de durée, qui implique l'accessibilité aux bornes du procès, et l'aspect inaccompli, qui l'exclut. » (1996 : 88). Dans la lecture itérative, l'aspect inaccompli ne porte plus uniquement sur le procès mais sur la série itérative dans son ensemble, alors que dans la lecture future, il marque un glissement vers l'état préparatoire du procès...

1.2.2.3. *Alternance, concordance, liberté*

Parallèlement au cas, majoritaire, dans lequel la structure syntaxique n'admet qu'un seul mode verbal, il arrive qu'une même construction admette plus d'un mode. Malgré sa relative marginalité, ce cas, qui relève de ce qu'on a appelé *alternance modale*, est intéressant à observer en comparaison avec la situation qui se présente dans le domaine de la temporalité.

Si on fait abstraction de certains emplois, que l'on peut facilement juger marginaux, comme dans :

(29) (a) Le Ciel fasse que vous ayez raison !

(b) Dieu te bénisse !

où l'apparition du subjonctif, dans une structure typique de l'indicatif, est extrêmement contrainte⁵⁶, on peut penser que le phénomène de l'alternance modale est relativement marginal. Non seulement, il ne concerne que le couple indicatif/subjonctif⁵⁷, mais en plus, il ne concerne qu'un nombre réduit de structures, notamment certaines relatives (très contraintes) et certaines complétives (également), illustrées respectivement par les exemples qui suivent :

- (30) (a) Je cherche un livre qui rend/ rende compte de ce phénomène
 (b) Crois-tu que la planète est/ soit en danger?
 (c) Je ne crois pas que la planète est/ soit en danger

S'il ne m'est pas possible, aujourd'hui, de maintenir l'intégralité des analyses publiées⁵⁸, il me semble que l'existence de ces cas n'invalides pas l'hypothèse de la visibilité syntaxique des modes, telle qu'elle a été présentée ici. Celle-ci gagnerait toutefois à mesurer le phénomène de l'alternance modale, en s'appuyant sur des études sur corpus, qui manquent cruellement aujourd'hui⁵⁹...

Dans le domaine temporel, j'ai suggéré que, du moment où l'indicatif est admis dans une structure *syntaxique*, toutes les formes temporelles de ce mode sont *a priori* admises, *modulo*, naturellement, le respect des règles habituelles d'ajustement référentiel et de cohérence, dont il a été question ci-dessus et qui ne relèvent pas de la syntaxe. Descriptibles et prédictibles (voir Gosselin 1996), ces règles offrent tout de même au locuteur une latitude et une liberté dont il ne peut que très rarement jouir sur le plan modal, ainsi que l'atteste le long débat historique de la notion de « servitude grammaticale ». C'est dans ce sens qu'on propose ici d'opposer alternance modale à liberté temporelle. L'existence de mécanismes de résolution de conflits, proposés par Gosselin, montre dans un sens que l'étendue de l'acceptable a peu de limites, puisque l'interlocuteur et le linguiste procèdent à des calculs, qui peuvent être de plus en plus complexes, pour sauver l'attesté⁶⁰...

⁵⁶ elle n'est possible qu'avec « un sujet religieux », écrit Togeby (1982: 68).

⁵⁷ Cela implique que l'on ait pris acte des descriptions en termes syntaxiques que propose par exemple l'école générativiste à propos de l'alternance entre infinitives et « phrases à temps fini ».

⁵⁸ Notamment certaines notions techniques ainsi que la nature des traits propagés.

⁵⁹ Le choix du corpus sera crucial : il devra à la fois montrer la fréquence réelle de ces cas en français moderne, à l'écrit et à l'oral, et rendre compte des différents « registres », tant il nous semble vrai que ce type d'alternance est particulièrement sensible aux paramètres variationnistes... Voir aussi la note 34.

⁶⁰ C'est du reste ce que semble affirmer Gosselin quand il écrit (2005 : 122) : « Il est apparu qu'en fait très peu de combinaisons de marqueurs sont effectivement impossibles, mais que, bien souvent, elles déclenchent des conflits entre instructions qui sont résolus par des procédures régulières de déformation. »

Il reste le phénomène dit de la concordance des temps, qui semble *a priori* contredire l'hypothèse de leur invisibilité en syntaxe.

S'inscrivant à l'encontre des puristes qui allaient jusqu'à considérer comme des défections les cas où la concordance des temps n'est pas respectée, Brunot affirmait (1922: 782) : « Ce n'est pas le temps principal qui amène le temps de la subordonnée, c'est le sens. Le chapitre de la concordance des temps se résume en une ligne : Il n'y en a pas ».

Je ne reprendrai pas le détail des analyses et les exemples donnés dans (Abouda 2010) – là aussi, une étude sur des corpus variés serait aujourd'hui la bienvenue –, mais rapidement l'hypothèse principale qui y est mentionnée. Trop mécaniques, alors qu'elles semblent souvent transgressées, présentées comme opérant au niveau de la phrase alors qu'elles peuvent être débusquées au-delà, les règles de la concordance des temps ne seront pas maintenues, mais remplacées par une notion plus sémantique, et du coup plus appropriée, i.e. l'ancrage référentiel⁶¹.

L'idée principale consiste d'abord à inverser le sens de la relation : ce n'est pas le verbe principal qui transmet des traits temporels au verbe subordonné (sinon on ne comprendrait pas toutes les transgressions), c'est le verbe subordonné qui, parce qu'il a besoin d'un repère référentiel, peut aller le chercher dans la principale. Et si tous les temps de la subordonnée ne sont pas concernés, c'est parce que seuls le font ceux qui sont indéterminés référentiellement (en gros, les temps qu'on a pu appeler *relatifs*⁶², qui ne se calculent pas par rapport à t_0). J'y reviendrai dans la 2^e partie de cette synthèse.

L'autre avantage de cette hypothèse, c'est qu'elle cesse de voir ces relations comme un phénomène qui s'exerce en syntaxe, permettant ainsi son traitement au niveau textuel, comme dans le cas du discours indirect libre. Dans l'exemple suivant de *Madame Bovary* :

- (31) Et il se mit à faire des préparatifs intérieurs, il arrangea d'avance ses occupations. Il se meubla, dans sa tête, un appartement. Il y mènerait une vie d'artiste. Il y prendrait des leçons de guitare! Il aurait une robe de chambre, un béret basque, des pantoufles de velours bleu !

l'on peut comprendre naturellement que les verbes au conditionnel s'attachent au point référentiel fourni par le verbe *meubler* au PS. Il s'agit d'ailleurs non seulement d'une attache temporelle, mais également énonciative.

⁶¹ L'hypothèse ne contredit pas, dans son esprit, le traitement de Brunot, qui, toutefois, après avoir rejeté le principe de concordance, en y voyant un phénomène purement sémantique, reconnaît l'existence de qu'il appelle une « certaine attraction de formes », où « la pensée n'est pour rien » (1936 : 787)...

⁶² Voir Vetters (1993) et (1996) pour une bonne présentation historique et une discussion de cette notion.

Tout cela concourt à montrer que c'est à la cohérence sémantique (entre marqueurs linguistiques et avec les contraintes situationnelles), et non à la syntaxe, qu'il convient d'attribuer l'obligation pour le locuteur d'employer le conditionnel dans certains énoncés (vs phrases)⁶³. C'est cette adéquation logico-sémantique qui veut que dans une boulangerie par exemple un francophone soit tenu d'énoncer 43(a), et non pas 43(b) et encore moins 43(c), qui sont pourtant toutes trois des phrases tout à fait grammaticales et interprétables (et même possibles) :

- (32) (a) Je voudrais une baguette
 (b) Je veux une baguette !
 (c) J'ai voulu une baguette

En bref, le conditionnel, comme tous les temps de l'indicatif, est un objet sémantique.

⁶³ Cette distinction n'est pas sans rappeler la distinction opérée par Ducrot (1984). « Ce que j'appelle "phrase", note-t-il (1984 : 174), c'est un objet théorique, entendant par là qu'il n'appartient pas, pour le linguiste, au domaine de l'observable, mais constitue une invention de cette science particulière qu'est la grammaire. Ce que le linguiste peut prendre pour observable, c'est l'énoncé, considéré comme la manifestation particulière, comme l'occurrence *hic et nunc* d'une phrase. Supposons que deux personnes différentes disent « Il fait beau », ou qu'une même personne le dise à deux moment différents : on se retrouve en présence de deux énoncés différents, de deux observables différents, observables que la plupart des linguistes expliquent en décidant qu'il s'agit de deux occurrences de la même phrase française, définie comme une structure lexicale et syntaxique, et dont on suppose qu'elle leur est sous-jacente. »

2. Une vision de l'indicatif : propositions

La relative discrétion, rappelée au début de la section précédente, avec laquelle le conditionnel a changé de statut catégoriel dans les textes officiels interpelle. Ce n'est pas qu'on attende en pareille occasion des cérémonies officielles, ni même qu'on s'étonne de tout le temps écoulé entre un pseudo-accord chez les classiques et sa répercussion dans les circulaires de l'éducation nationale⁶⁴. Mais on pouvait espérer restituer le moment qui dans le champ linguistique pouvait justifier ce revirement.

Sauf qu'il me semble plus exact de parler d'évolution lente, peut-être même laborieuse, que de revirement. A l'image du *Bon Usage* qui, édition après édition, envisageait de plus en plus sérieusement l'idée, l'institution académique a fini par noter que, hormis la manifestation sporadique de quelques voix discordantes⁶⁵, une nette majorité s'était dessinée chez les grammairiens et les linguistes en faveur du traitement du conditionnel comme une forme de l'indicatif. Aussi était-il légitime et même naturel de s'appuyer sur ces travaux, et de ne pas reposer, ou pas si lourdement, la question du statut catégoriel du conditionnel. Après tout, beaucoup de recherches sur cette forme ou sur un ou plusieurs de ses emplois, datant des années 1990 et 2000, n'abordaient pas la question, ou se contentaient, si elle s'avérait nécessaire à leurs analyses, d'en donner une réponse le plus souvent courte, en s'appuyant sur cet accord, ou sur l'une des références classiques. On pouvait donc légitimement estimer qu'après une longue période de débat l'affaire était close : la question, qui avait été longuement débattue au sein de la linguistique structurale, intéressait peu une linguistique énonciative émergente, portée sur la parole et le discours, et privilégiant par conséquent des recherches sur les *emplois* du conditionnel. Cette nouvelle tendance s'est d'autant plus aisément imposée que certains traitements structuralistes en morphologie et en sémantique rencontraient des difficultés, dont nous verrons quelques aspects ci-dessous.

De revirement, il n'y a point eu, mais un désintérêt, certainement.

Ce constat historique peut paraître court car, au-delà de la fragilité, déjà signalée en 1.1, des arguments classiques, il semble que le terme *accord* n'est sans doute pas le plus approprié. Tant il est vrai que, derrière l'étiquette *temps* que chacune des écoles attribuait au conditionnel ne se cachaient pas toujours des notions similaires. Entre la conception guillaumienne et la conception de Damourette & Pichon, pour ne citer que les écoles les plus influentes dans le champ de la temporalité, il n'y avait pas en effet beaucoup de points

⁶⁴ On était visiblement dans le rythme habituel : Fuchs et Léonard (1979) rapportent que la première circulaire qui a introduit explicitement la notion de l'aspect dans l'enseignement du français pour le second degré date du 22 juillet 1975.

⁶⁵ Notamment Feuillet (1988) et Sabaneeva (1993).

communs. Non seulement leurs conceptions respectives des temps et des modes différaient sensiblement, mais les raisons qui avaient poussé l'une et l'autre école à admettre le conditionnel au sein de l'indicatif étaient assez éloignées. On parlerait là plus spontanément de trêve que d'accord.

En bref, il me semble aujourd'hui que la réouverture de ce dossier, même coûteuse – parce qu'il est indéniable que la question a pris une place certaine dans ma thèse, et peut-être même au détriment d'autres questions –, était épistémologiquement justifiée et, en lien avec l'hypothèse de la visibilité syntaxique des modes dans le cadre d'une syntaxe pauvre, à nouveau redevenue possible.

2.1. Préliminaires

2.1.1. Paradigmes de paradigmes

M'inscrivant résolument dans une conception saussurienne du signe, il m'a paru primordial de situer le conditionnel dans le bon paradigme. Si, ainsi que l'écrira Saussure (1916 : 162) dans une formule qui deviendra célèbre, la « plus exacte caractéristique [d'un signe] est d'être ce que les autres ne sont pas », il était nécessaire d'identifier le sous-système, le sous-espace conceptuel dirait Gosselin (2010), dans lequel le conditionnel s'opposerait à tous les autres signes de même nature.

Se prononcer sur le statut catégoriel du conditionnel revient ainsi à identifier le réseau d'oppositions dans lequel se déterminent différenciellement les propriétés de ce marqueur. Celles-ci doivent d'une part l'inscrire dans un paradigme, qui s'oppose lui-même à d'autres paradigmes, et, d'autre part, le différencier de tous les autres marqueurs du même paradigme.

Concrètement, cette exigence revient à essayer d'attribuer au conditionnel une valeur qui soit suffisamment discriminante pour l'opposer aux autres formes temporelles de l'indicatif, mais pas au point de le sortir du paradigme indicatif, quand viendra le temps de s'occuper de l'opposition indicatif-subjonctif. Le défi n'est pas mince, et on comprend aisément, pourquoi, dans une sorte de vide idéologique post-structuraliste, il était nettement plus payant de s'occuper des emplois...

Le défi est d'autant moins mince que l'indicatif ne constitue pas lui-même un paradigme homogène mais présenterait des « sous-systèmes d'opposition locaux » (Lemaréchal, 2001). C'est la thèse qui sera défendue ci-dessous quand il sera question de distinguer entre deux groupes de formes temporelles.

Ayant conduit au terme de cette organisation paradigmatique hiérarchisée, le chemin ne s'arrêtera pas là et devra affronter un obstacle connu, celui des emplois, qui, dans un domaine grammatical hautement polysémique, constituent le défi majeur à toute proposition d'organisation en systèmes.

Autrement dit, il convient non seulement de proposer pour le conditionnel une valeur qui, tout en l'opposant aux autres temps du système (et du sous-système qui sera proposé ci-dessous), ne le sorte pas du paradigme indicatif (ni du sous-paradigme qui lui sera proposé), mais aussi que la dite valeur rende compte de tous les emplois empiriquement identifiables. La problématique posée ainsi s'inscrit au cœur d'un débat théorique qui voit s'affronter actuellement les écoles sémantiques, et dont il est impossible de rendre compte exhaustivement dans ce cadre. Je me contenterai donc d'évoquer quelques étapes clés du chemin que j'emprunte, en renvoyant à Gosselin (2010⁶⁶) pour une présentation précise de ces questions dans le champ de la temporalité.

2.1.2. Mode vs temps VS Modalité vs temporalité

Ainsi qu'il a été rappelé au début de cette synthèse⁶⁷, s'interroger sur la nature temporelle ou modale du conditionnel constitue une question ambiguë, qui, même lorsqu'elle est ramenée comme ici à la question de l'appartenance ou non de cette forme au mode indicatif, n'écarte pas définitivement tout risque de malentendu. En cause les liens abusifs, qualifiés de mnémotechniques par Weinrich (1973), entre temps linguistique et Temps extra-linguistique, et leurs corollaires dans le champ de la modalité, mode verbal et modalités⁶⁸. Une tendance, souvent implicite mais qu'on peut régulièrement débusquer dans les raisonnements les plus courants, consiste à attendre d'un temps linguistique qu'il serve exclusivement à situer chronologiquement les procès sur l'axe temporel, et d'un mode verbal qu'il exprime de la modalité.

Si je suis d'accord avec Gosselin pour attribuer la cause de cette dichotomie exclusive, selon laquelle « ce qui est temporel ne saurait être modal, et réciproquement » (2010 : 73), à la conception atomiste des relations entre sémantique et morphologie, particulièrement visible dans le paradigme structuraliste, il me semble possible d'y voir une nouvelle⁶⁹ manifestation

⁶⁶ Notamment le 1^{er} chapitre de la seconde partie.

⁶⁷ § 1.1.1.3, p. 8 sq.

⁶⁸ Voir notamment Confais (1990) et Gosselin (2010).

⁶⁹ Mais cette tendance remonte à loin, et influencera le champ de la temporalité d'une manière décisive. On peut par exemple se rendre compte, à la lecture de Fournier (1986), que c'est le traitement de l'IMP comme une

de ce qu'on pourrait appeler le *substantialisme* qui, refusant aux morphèmes zéros toute possibilité d'existence, ne consent à reconnaître l'existence d'une valeur sémantique (donc en fait le statut de morphème) « que dans la mesure où l'on est capable d'identifier un marqueur spécialement consacré à son expression » (Gosselin, op. cit. : 76). La notion de *marqueur* n'est d'ailleurs sans doute pas exempte de tout reproche, puisque, renvoyant au concept phonologique de *marque*, sort la *valeur* de son rapport différentiel pour la tirer vers la substance. En bref, ainsi que l'affirme Gabriel Bergounioux dans une communication personnelle, la *marque*

est un concept clé de la régression jakobsonienne en phonologie (un peu par rapport à Troubetzkoy, beaucoup par rapport à Saussure) qui a trouvé son prolongement dans l'ensemble des niveaux de l'analyse linguistique où il a été transposé. Avec la notion de *marque*, l'idée est de prendre pour repère un élément de substance qui sert d'identifiant au lieu de s'en tenir au seul principe de la différence, i.e. de l'opposition (et non de la *distinction* comme le reformulera Jakobson dans ce projet constant d'imposer un substantialisme).

Dans le champ de la temporalité, ce substantialisme sera à l'origine de nombreuses analyses qui, par exemple, déniaient au présent de l'indicatif (à cause de son zéro) toute valeur sémantique particulière à tel point qu'elles ne parviennent même plus à rendre compte de ses emplois les plus courants et les plus basiques, comme celle par exemple d'exprimer le présent⁷⁰.

La dichotomie exclusive décrite ci-dessus, en plus d'être critiquable sur le plan théorique et méthodologique, est clairement contredite sur le plan empirique, où il est difficile, dès que l'on affronte les données, de ne pas identifier dans chaque forme verbale « une formule qui combine une variable modale avec une variable temporelle et une variable aspectuelle », ainsi que l'écrira Wilmet (1997 : 329).

Pour ne donner qu'un seul exemple, je peux évoquer la recherche que nous menons actuellement avec Marie Skrovec sur la distribution dans un corpus oral des formes analytique et synthétique du futur en français⁷¹. Refusant, lors de l'annotation du corpus, conçue comme une phase strictement descriptive, de partir d'une taxinomie déjà établie, nous avons été amenés très vite, à identifier, parallèlement aux classes dichotomiques habituelles d'emplois

forme de l'indicatif qui finira, dès le XVII^e siècle, par privilégier son emploi comme temps du passé au détriment de ses emplois dits modaux, qui relevaient du domaine du subjonctif latin... Voir § 2.2.2.2 ci-dessous.

⁷⁰ Gosselin rappelle cette critique de Le Goffic (éd., 2001), par P. Attal (2002 : 203, cité dans Gosselin, op. cit.) : « si je dis à quelqu'un « Je suis malade », il ne lui viendrait pas à l'idée de me demander « Hier ou demain ? » ».

⁷¹ Abouda & Skrovec (2015, à paraître).

temporels (f) et d'emplois modaux (m), une troisième classe, mixte, d'emplois (étiquetés "fm") :

La reconnaissance de cette troisième classe d'emplois est l'une des implications directes de notre méthodologie de travail, i.e. une approche guidée par les données. Celle-ci nous conduit à étiqueter toutes les occurrences, y compris celles qu'on aurait sans doute écartées dans le cadre d'une analyse simplement basée sur corpus, parce qu'elles n'auraient pas correspondu aux exemples typiques qu'on aurait cherchés pour illustrer des classes d'emplois préétablies. Or il nous a paru très vite impossible d'inclure ces emplois mixtes dans l'une des deux classes dichotomiques habituelles sans sacrifier l'une des deux faces sémantiques exprimées par les occurrences en question. (Abouda & Skrovec, 2015 : 8).

Toutes ces considérations me permettent d'affirmer que ce n'est pas parce que le conditionnel appartient à l'indicatif qu'il cessera d'avoir (éventuellement) des emplois modaux. Il s'agit là d'une question empirique dont le traitement requiert l'examen des données.

2.1.3. Traitement de la polysémie temporelle

Il est aujourd'hui couramment admis que le phénomène de la polysémie constitue le défi majeur à tous les modèles linguistiques et qu'elle peut même être considérée à l'origine de bouleversements majeurs, qu'ont notamment connus au cours des quatre dernières décennies les champs de la sémantique et du TAL.

Si la polysémie est générale et peut concerner tous les signes linguistiques, elle « caractérise au premier chef les marqueurs grammaticaux, et tout particulièrement les marqueurs de temps, d'aspect et de modalité », ainsi que l'écrit Gosselin (2010 : 105). Le conditionnel, à l'image de toutes les formes temporelles, a ainsi une multitude de significations identifiables en contextes, ayant fait l'objet de diverses propositions de regroupement au sein de classes d'emplois, dont la plus classique distingue des emplois modaux face à un emploi temporel, et les plus actuelles identifient trois classes d'emplois, i.e. temporel, hypothétique et de non-prise en charge⁷².

La reconnaissance, inévitable, de la polysémie peut s'expliquer par le principe de contextualité, qu'on attribue à Frege⁷³, et que Gosselin (2010 : 107) se propose de formuler ainsi : « la signification d'une expression est au moins partiellement déterminée par le

⁷² Cette tripartition, qui semble aujourd'hui largement admise, a d'abord été, me semble-t-il, proposée par Haillet (1992) et (1995), qui, remettant en cause explicitement la bipartition classique en emplois temporel et modaux, s'attachera (1995 & 2002) à proposer des critères explicites et opératoires permettant de distinguer les trois classes d'emplois. Ce que j'appelle « non-prise en charge » est une famille d'emplois qui a reçu dans la littérature grammaticale et linguistique une multitude d'appellations (journalistique, d'altérité énonciative, évidentiel...).

⁷³ Voir toutefois Milner (1989 : 316, note 1).

contexte dans lequel elle apparaît ». Dès lors se crée, comme le note Benmakhlouf (2002 : 141) une « tension entre le principe de contextualité [...] et le principe de compréhension des énoncés nouveaux qui nous oblige à construire le sens de l'énoncé à partir de blocs constitutifs ». Comme le rappelle Gosselin (*ibid.*), pour résoudre la contradiction entre les deux principes de compositionnalité et de contextualité, et rendre ainsi possible une prédictibilité des modèles sans renier les emplois, toute une tradition sémantique européenne allant de Saussure à Culioli a opté pour une distinction fondamentale entre une valeur en langue, stable et abstraite, et des significations en contexte (ou effet de sens, ou désignés contextuels, etc.)⁷⁴. Si la valeur en langue s'associe aux signes et se détermine au niveau du système, la détermination des valeurs d'emplois s'opère empiriquement en contexte.

Ne restera alors plus qu'à les relier !

Or c'est bien là que réside, selon Gosselin (1996 : 187-188)

l'un des problèmes épistémologiques fondamentaux de la sémantique, qui prend la forme d'un paradoxe : seule la valeur en langue appartient en propre au marqueur (elle constitue l'identité différentielle du signe, selon Saussure), mais elle n'est pas directement observable (hors contexte et hors situation de discours) ; alors que les effets de sens, qui seuls paraissent observables, en particulier par le jeu des relations de synonymie, dépendent autant du contexte que du marqueur lui-même, au point qu'il arrive souvent qu'on ne sache déterminer précisément quelle part revient à l'une ou l'autre de ces entités dans la production d'un effet de sens. Et il y a une bonne part d'arbitraire et d'illusion dans les grammaires qui dressent la liste des effets de sens d'un marqueur donné, comme s'il en constituait le support unique.

Je reviendrai longuement sur ces considérations épistémologiques et méthodologiques dans la troisième partie de cette synthèse, où il sera question de la gestion des données. La polysémie sera à nouveau abordée auparavant en 2.3.2, p. 82 sq..

En attendant, il me paraît incontournable, dans le cadre de la vision systémique que je tente de développer ici, d'opter pour une démarche hypothético-déductive. Elle consiste à proposer pour chaque forme verbale de l'indicatif une valeur en langue, déterminée à titre d'hypothèse et dans son rapport différentiel avec d'autres valeurs. Les valeurs proposées devront par la suite affronter l'épreuve de l'observable, le prédire et l'expliquer, sous peine d'être invalidées.

⁷⁴ J'attire l'attention sur la très grande profusion terminologique que la plupart des écoles sémantiques s'évertuent à entretenir, allant jusqu'à proposer des acceptions diamétralement opposées à des couples de termes utilisés par ailleurs dans une autre école : locuteur/énonciateur, sens/signification...

2.2. Deux groupes de formes temporelles

2.2.1. Morphologie de l'indicatif⁷⁵

Si les considérations strictement morphologiques étaient jugées ci-dessus insuffisantes pour déterminer le statut catégoriel du conditionnel, l'hypothèse de son appartenance à l'indicatif, adoptée sur la base de données syntaxiques, a des conséquences morphologiques sur tout le système temporel, qu'il est légitime, et même nécessaire, d'examiner si on tient compte des exigences méthodologiques rappelées ci-dessus. C'est ce que je proposerai dans cette section : tirer les conséquences morphologiques de ce traitement sur tout le système de l'indicatif. Ce chemin, s'il n'est pas habituel – les traitements généralistes classiques commencent systématiquement par le présent pour finir par le conditionnel – me semble à la fois plus riche et, sans doute, paradoxalement plus simple.

Ainsi qu'il a été dit ci-dessus⁷⁶, les deux structures concurrentes A et B du schéma (1), ne peuvent pas être simultanément attribuées au conditionnel. M'appuyant sur un faisceau d'indices et d'arguments présents dans de nombreux travaux qui soulignent les liens du conditionnel avec l'imparfait et le futur, je me propose d'explorer dans ce cadre la lecture A, que je reproduis ici par commodité :

$$(33) \quad \text{Radical} + R + \varepsilon + \text{PN}$$

On peut naturellement penser que la structure morphologique (et sémantique) du conditionnel est singulière : de très nombreux travaux, à la fois en morphologie⁷⁷ et en sémantique⁷⁸, lui accordent un traitement spécifique.

⁷⁵ La question morphologique sera ici considérée exclusivement dans son interface avec la sémantique, laissant de côté le volet morpho-phonologique. Dans une communication orale (Abouda & Moukrim, 2010), nous avons adopté le modèle de la Morphologie Distribuée (désormais MD), développé par Halle & Marantz (1993 & 1994). Née dans le paradigme générativiste, et plus précisément dans le cadre du Programme Minimaliste (Chomsky, 1995), la MD offre l'avantage indéniable d'être compatible avec n'importe quel modèle grammatical, grâce à l'hypothèse de l'Insertion Tardive (Marantz 1995). Le morphème *y* est vu comme une matrice de traits morphosyntaxiques dépourvu de contenu phonologique. Le composant Morphologique, autonome, intervient sur l'output de la syntaxe pour redistribuer les traits morphosyntaxiques de ces éléments par le biais des règles morphologiques (fusion, fission, appauvrissement...). Ce n'est qu'après les règles morphologiques que le contenu phonologique est associé aux nœuds terminaux. Le processus réalisationnel (*spell-out*) au cours duquel le morphème reçoit son contenu phonologique se présente comme un mécanisme de computation post-syntaxique, qui intervient en tout dernier. En plus d'être compatible avec nos vues épistémologiques (notamment en ce qu'il adopte une phonologie abstraite), le modèle me semble particulièrement bien placé pour traiter les données flexionnelles extrêmement complexes du français. Voir Lampitelli (2015) pour une description morpho-phonologique de l'imparfait dans ce modèle.

⁷⁶ 1.1.1.2 (p. 4 sq.).

⁷⁷ Voir Pinchon & Couté (1981), et la note 14.

⁷⁸ C'est la seule forme temporelle *simple* de l'indicatif qui a besoin, dans le système de Gosselin par exemple, de deux intervalles de référence distincts...

Mais il me semble également possible de penser que cette singularité est douteuse, et que la structure du conditionnel n'a de spécifique que le double marquage phonologique de deux morphèmes qui coexistent dans tout le paradigme de l'indicatif. Cela revient à voir dans la catégorie traditionnelle « temps » une catégorie générique, amalgamant systématiquement deux catégories distinctes, selon le schéma morphologique (simplifié, où l'on ne tient compte ni du Mode ni de l'Aspect) :

$$(34) \quad \text{Radical} + X + Y + \text{PN}$$

A partir de ce schéma global, on peut organiser les formes temporelles de l'indicatif de différentes manières⁷⁹. Je me propose de décliner ce schéma morphologique simplifié de la manière suivante :

(35)

Radical	X	Y	Temps
	Ø	Ø	PR
	Ø	ε	IMP
	e	Ø	PS
	R	Ø	FUT
	R	ε	COND

où on note la présence de deux paradigmes parallèles, appelés provisoirement X et Y, nés de la fission de la catégorie habituelle « temps ».

Le paradigme X oppose trois morphèmes :

Ø : morphème de concomitance : PR ou IMP, selon la valeur de Y

e : morphème d'antécédence : PS (qui n'apparaît que si Y = Ø)

R : morphème de prospection : FUT ou COND, selon la valeur de Y.

Si, servant à situer chronologiquement les procès par rapport à un repère référentiel, la catégorie X se présente comme une catégorie strictement temporelle, la détermination de la catégorie Y peut en revanche s'avérer plus délicate.

⁷⁹ Voir par exemple le traitement de Camussi-Ni (2006), très différent de celui qui sera proposé ici, malgré des choix méthodologiques initiaux comparables.

Opposant un morphème $\emptyset$ au morphème ε , la catégorie Y permet de répartir les temps de l'indicatif en deux sous-groupes distincts, comme on peut le visualiser dans le tableau suivant :

(36)

Radical	X	Y	Temps
	$\emptyset$	$\emptyset$	PR
	e		PS
	R		FUT
	$\emptyset$	ε	IMP
	R		COND

- (i) le premier groupe ($Y = \emptyset$) a pour centre le présent, et contient en plus de lui, selon la valeur de X, le futur et le PS (et leurs correspondants composés).
- (ii) Le second groupe ($Y = \varepsilon$) comporte l'IMP et le COND (et leurs formes temporelles composées correspondantes).

L'identification de la catégorie Y exige d'identifier au préalable ce qui oppose les deux groupes de formes temporelles, et d'abord leurs deux repères-origines respectifs, correspondant au PR et à l'IMP.

L'opposition PR/IMP, et plus généralement le point référentiel à partir duquel s'envisage l'IMP, est l'une des questions les plus traitées dans le champ de la temporalité. Si c'est loin d'être la seule piste explorée, l'opposition PR-IMP a souvent été décrite en termes temporels. Or, si on envisage sérieusement le schéma (34), voir dans Y une seconde catégorie temporelle reviendrait de fait à la faire disparaître : les deux paradigmes parallèles finiraient par fusionner à nouveau et les zéros n'auraient été qu'un outil avancé provisoirement au cours de l'analyse avant de disparaître... Bref, on reviendrait au schéma habituel, où il n'y aurait qu'une seule catégorie temporelle, et où on peut voir chacune des formes simples de l'indicatif comme une forme homogène, sauf le conditionnel, matériau oblige, dans lequel on va continuer à voir tout de même une forme composite et singulière. Ce traitement singulier va, à mon avis, à l'encontre de l'intégration pleine du conditionnel au sein de l'indicatif.

2.2.2. Arguments pour une bipartition

Je n'ai pas eu l'occasion de consacrer un article spécifique à l'hypothèse de la bipartition des formes temporelles, telle qu'elle est présentée ici, même si j'ai pu la défendre dans un certain nombre de communications orales⁸⁰, et si quelques indications dispersées dans différents textes publiés depuis la thèse⁸¹ témoignent de son émergence. C'est qu'il me semble, entre autres raisons, que le format habituel d'un article linguistique ne s'y prête pas : une taille relativement réduite peu compatible avec les exigences scientifiques du genre, où chaque argument mériterait un développement important et l'examen d'une littérature antérieure conséquente, sans parler de la nécessaire description d'un grand nombre d'emplois temporels et notamment ceux qui semblent entrer en contradiction avec les prédictions défendues...

2.2.2.1. Une piste ancienne mais actuelle

L'hypothèse de l'existence à l'indicatif de deux groupes de formes temporelles s'appuie ici sur des considérations morphologiques : c'est l'une des lectures rendues possibles par les jeux paradigmatiques. Mais l'idée elle-même de cette bipartition (ou de bipartitions proches) n'est pas nouvelle. On peut même penser que, remontant à plus de trois siècles, elle est bien ancrée dans la tradition linguistique, dès lors qu'on accepte de la détecter sous des appellations et des formes variées : temps simples/temps composés (dans le sens) chez Arnauld & Lancelot (1660)⁸², temps absolus/temps relatifs (Girard (1747), Brunot (1922)...); temps noncaux / temps toncaux (Damourette & Pichon (1911-1940); temps actuels/temps inactuels (Coseriu (1980)...); temps déictiques/temps anaphoriques (Maingueneau, Tasmowski, Molendijk, Kleiber...), premier sous-système/second sous-système (Wilmet)...⁸³

Naturellement, toutes ces bipartitions, avancées dans les cadres théoriques les plus divers, ne sont pas équivalentes, et comportent même des différences majeures. Par exemple, si certains auteurs classent des tiroirs, certains autres classent des emplois, en s'appuyant le plus souvent sur des exemples fabriqués. L'extension elle-même de chaque groupe, s'appuyant sur des définitions différentes plus ou moins explicites de la nature de l'opposition, peut varier

⁸⁰ Dans la rubrique « Communications » du CV : Abouda (2004a) et (2004b), et Abouda & Moukrim (2010).

⁸¹ Cf. Chapitre VI, pp. 513-532.

⁸² Il s'agit non pas d'une opposition entre les *formes* verbales simples et complexes, telle qu'on se la figure actuellement, mais bien d'une opposition notionnelle équivalente à celle qui m'occupe ici, puisque les auteurs distinguent nettement entre les temps « considérés simplement dans leur nature de présent, de prétérit et de futur » et les « temps composés dans le sens » (1660 : 76) parmi lesquels figure l'imparfait.

⁸³ Voir aussi encore récemment Vet (2005). Tout cela sans parler du modèle de Benveniste et celui de Weinrich, dont la bipartition en plans ne correspond que partiellement au domaine temporel.

d'une bipartition à une autre⁸⁴. Enfin, ces bipartitions ont souvent laissé place à des tripartitions, comme on peut le voir avec l'apparition de la classe des anadéictiques ou encore la classe de la temporalité absolue-relative proposée par Comrie (1985) dans une perspective de linguistique générale⁸⁵.

Il n'empêche que l'idée qu'il existe au sein de l'indicatif deux sous-systèmes parallèles dont l'un est organisé autour du présent et l'autre autour de l'imparfait semble être une constante dans un champ grammatical et linguistique, connu par ailleurs pour sa grande diversité. Ce constat historique et épistémologique a déjà été souligné par Wilmet (1996 : 200) :

Les grammairiens ont longtemps vu dans l'imparfait français un "présent du passé" (Maurice Grevisse jusqu'à la septième édition du *Bon usage*, en 1961) ou un "deuxième présent" (Weber 1954), ou un "passé du présent" (la majorité des francistes espagnols et italiens). Ces formulations maladroites traquent une ressemblance intuitivement perçue de l'imparfait et du présent.

Je ne résiste pas à la tentation de citer l'Abbé Girard (1647 : t.2, p. 25) qui présente sa vision de la bipartition de la manière suivante :

On pourroit en faire la distinction par les simples épithètes numérales de *premier* & *second* : mais pour les mieux caractériser je les distingue par celles d'*absolu* & *relatif*. Lorsqu'ils représentent le temps de l'événement par la seule comparaison avec celui où l'on parle ils sont TEMPS ABSOLUS ; lorsqu'ils le représentent par une double comparaison, non seulement avec le temps de la parole, mais avec celui de quelque autre événement, ils sont TEMPS RELATIFS.

Sans aller jusqu'à dire, à l'instar de Vetters (1996 : 45), que cette définition « n'a rien perdu de son actualité et que les définitions actuelles reviennent exactement au même », il me semble en effet que, malgré une évolution lente, une traque collective s'est engagée depuis longtemps pour identifier ce différentiel insaisissable entre deux paradigmes temporels, et partant entre leurs deux centres respectifs, PR vs IMP.

2.2.2.2. De l'opposition PR/IMP

Comme je l'ai indiqué ci-dessus, l'opposition entre le présent et l'imparfait a fait l'objet d'un nombre considérable d'études linguistiques, que le cadre du présent travail ne permet pas de discuter. Il serait en effet irréaliste de penser que je pourrais ici, au détour d'une argumentation, aborder exhaustivement un domaine très largement débattu et très complexe.

⁸⁴Voir Vetters (1996), pour une présentation historique assez détaillée mais qui, curieusement, ignore Damourette & Pichon.

⁸⁵ Une linguistique générale, mais pas dans une perspective typologique... C'est pour rendre compte, dans le système de Reichenbach, de certaines formes temporelles de l'anglais, i.e. le « pluperfect » et le « future perfect », qu'il a été amené à reconnaître la classe du « combined absolute-relative tense » (1985 : 125 sq.).

Je me contenterai donc de faire un certain nombre d'observations, en privilégiant la présentation des analyses qui, quel que soit leur cadre théorique, semblent appuyer la vision que je défends.

D'abord, il me semble important de rappeler, suite à Fournier (1986)⁸⁶, que c'est bien l'héritage latin qui explique le traitement de l'imparfait comme un temps du passé. « Jusqu'à Port-Royal, écrit-elle (1986 : 15), la problématique de l'imparfait frappe par sa stabilité, évidemment due à l'héritage latin : pour les grammairiens, *j'aimais*, c'est *amabam*, et ils se félicitent de cette équivalence ». Centrées sur la morphologie et l'étude des distributions dans une visée essentiellement didactique, les « petites grammaires d'usage », dont le représentant le plus illustre sera Maupas, auraient pu, grâce à l'attention qu'elles portaient aux emplois, y compris « non-canoniques », et aux propriétés aspectuelles de la forme, nuancer plus fermement la caractéristique exclusivement chronologique attribuée à l'imparfait. Mais Port-Royal, s'attachant à proposer une systématique temporelle « logique », et même dans le cadre de la bipartition temps absolus/temps relatifs, écartera toute considération aspectuelle et négligera les emplois, avec les conséquences que l'on sait⁸⁷. On ne peut ne pas faire le lien entre la domination de la lecture chronologique de l'imparfait et le triomphe institutionnel et symbolique de Port-Royal. Cela me paraît d'autant plus vrai que j'ai en tête des exemples d'études, quelquefois de très grande qualité, où la nature temporelle de l'imparfait, posée par le modèle à titre axiomatique, ne joue quasiment aucun rôle explicatif et n'est maintenue, me semble-t-il, qu'artificiellement.

S'il est vrai aujourd'hui que la plupart des approches proposent pour l'imparfait un traitement aspectuo-temporel, il a toujours existé, avancés dans les cadres les plus divers, des traitements qui refusent de voir en l'IMP un temps (exclusivement) du passé.

Ainsi, pour Damourette & Pichon, ce temps, qui « est loin d'exprimer forcément un passé » (§ 1709)⁸⁸, signifie toujours « l'abandon du repérage par rapport au moi-ici-maintenant » (§

⁸⁶ Voir aussi n. 69, p. 33.

⁸⁷ « La conclusion que nous tirerons de cet aperçu de l'indicatif, écrit Fournier (1986 : 19), est que l'imparfait, de par son appartenance sans discussion à ce mode, est voué à l'expression du certain et de l'assertion pleine, ce qui ne laisse pas place a priori aux emplois dits modaux de ce tiroir qui sont renvoyés au subjonctif latin, sans autre commentaire ». Le glissement *Certain* > *Réel* > *Passé* fera le reste. Voir Gosselin (2010 : 73 sq.).

⁸⁸ Plus précisément, pour ces auteurs, la lecture temporelle de l'IMP n'est que l'une de ses lectures possibles qui ne semble émerger que compositionnellement : « il ne semble pas douteux que le rôle du *saviez* comme agent d'expression d'un passé n'est qu'un cas particulier de son sens toncal » (§ 1707).

1730 : 226)⁸⁹. Ils attribuent ainsi à l'IMP une valeur sémantique négative et vague de type non-présent, ou non-actuel, une sorte de non-[Moi-Ici-Maintenant].

On trouvera chez Coseriu (1980) une analyse similaire qu'il défendra dans une approche romanistique. Distinguant deux dimensions temporelles, i.e. le plan et la perspective, il définit la première (1980 : 19) comme

la dimension au moyen de laquelle certaines actions verbales sont placées directement par rapport à l'acte de la parole sur une ligne de réalité non-diminuée, tandis que d'autres actions sont placées comme sur une arrière-ligne (conditions, circonstances, etc. des actions "actuelles") et présentées comme plus ou moins diminuées dans leur effectivité.

Cette distinction, qui constitue d'après Coseriu le système de base de toutes les langues romanes, permet d'opposer deux sous-systèmes temporels : l'actuel, centré sur le PR ; et l'inactuel, ayant pour centre l'imparfait, qui, selon l'auteur (ibid.) « n'est pas à proprement parler un temps du passé ».

Solidement ancrée dans la tradition linguistique, cette hypothèse reste défendue dans de très nombreuses études linguistiques actuelles, où elle joue souvent un rôle explicatif central. Leur peu de visibilité et d'impact, en tout cas en dehors du champ de la temporalité, s'explique à nos yeux à la fois par la résistance, soulignée ci-dessus, de l'institution grammaticale, et, surtout, par leur dispersion dans des cadres théoriques étanches qui, se livrant à des polémiques théoriques sans fin, ne permettaient pas toujours aisément de voir la similitude entre les traitements.

Pour ne donner qu'un exemple illustrant l'actualité de cette hypothèse⁹⁰, on peut évoquer les études directement inspirées du modèle théorique de Culioli et de l'analyse esquissée dans son article de 1980. Afin de rendre compte des propriétés de l'imparfait, l'auteur y propose l'établissement d'un repère *translaté*, défini par sa relation de différenciation par rapport au repère d'origine. Reprenant le même cadre théorique, Lebaud (1993) attribuera à l'imparfait, outre la propriété négative de ne pas permettre de singulariser une classe d'instant (seuls les temps calculés par rapport au moment de l'énonciation permettent de singulariser une classe d'instant⁹¹), la propriété positive de marquer un changement de repère, c'est-à-dire la

⁸⁹ Ou encore, l'IMP décrit « un monde de phénomènes centrés autour de quelque chose d'autre que le moi-ici-maintenant » (§ 1740 : 234).

⁹⁰ On peut aussi se reporter par exemple à Le Goffic (1986), et plus généralement, à tous les articles rassemblés dans son recueil, ou encore, plus récemment, à Monville-Burston & Burston (2005)... On aurait pu aussi évoquer l'analyse de De Mulder (2012) qui n'écarter pas la possibilité que la distinction opérée par Langacker (2009) entre réalité immédiate et réalité non-immédiate puisse rendre compte de la plupart des emplois de l'imparfait... sauf peut-être, paradoxalement, l'emploi temporel. J'y reviendrai ci-dessous.

⁹¹ Sur ce point l'analyse évoluera : voir n. 99, p. 46.

construction d'un nouveau repère à partir d'un repère-origine. De nombreuses études développeront ces vues au cours des années 1990 et 2000.

En second lieu, on remarquera que si l'imparfait est, et de loin, la forme temporelle la plus polysémique⁹², le processus, vu ci-dessus, qui a consisté à neutraliser l'extrême variation dont il est porteur, a débuté dès le XVII^e siècle pour le faire entrer dans la seule case que le système logique classique, sous l'influence de l'héritage gréco-latin, pouvait lui laisser, celle d'un temps du passé. Il se trouve que les seuls auteurs qui, comme Damourette & Pichon, remettaient en cause le statut exclusivement temporel de l'imparfait étaient les mêmes qui acceptaient l'extrême richesse des emplois et la grande variété des usages. En face, la décision de traiter l'imparfait comme (d'abord) un temps du passé n'était compatible avec la prise en compte des emplois « déviants » ou « non-canoniques » qu'au prix de leur marginalisation. Ces derniers étaient alors vus comme secondaires, issus de cette valeur première par des opérations de dérivation dont Confais dira (1990 : 53) qu'elles « se font au coup par coup au moyen d'analyses *ad hoc* qui nécessitent un certain talent de la part du linguiste, car elles s'avèrent parfois très acrobatiques ».

A l'appui du traitement de l'IMP comme temps de passé, il n'y a pas uniquement l'héritage d'un champ, mais des arguments plus linguistiques, comme celui, présenté souvent comme décisif, qui consiste à affirmer qu'un énoncé à l'IMP, hors contexte, situe forcément le procès dans le passé, comme dans l'exemple suivant :

(37) Il pleuvait

La validité d'un tel argument me semble douteuse. Objet sémantique, ainsi que je me suis efforcé de le montrer, le temps ne s'observe pas au niveau de la phrase, car l'exemple (37) représente bien une phrase, et non un énoncé⁹³. Il est en effet admis dans la quasi-totalité des courants théoriques que c'est même le contexte qui différencie l'énoncé, seul observable⁹⁴, de la phrase, objet construit. En bref, un énoncé hors contexte n'existe pas et l'objet temps ne peut pas être appréhendé comme on appréhenderait un objet syntaxique.

⁹² Monville-Burston & Burston notent (2005 : 135) à juste titre : « Ce qui singularise l'imparfait [...] c'est la multiplicité de ses usages. Pour avoir une idée de cette diversité, il suffit de se reporter aux différentes rubriques sous lesquelles l'imparfait est décrit dans le classique *Bon Usage* : quinze, comparées à seulement quatre pour le passé composé et tout juste deux pour le passé simple ; de tous les autres temps, c'est le présent qui arrive bon second avec les six usages que M. Grevisse lui reconnaît ».

⁹³ Je renvoie aux excellentes observations de C. Fuchs (1986, notamment 51 *sq.*), qui, discutant du traitement psychomécanique de l'imparfait, souligne une certaine hésitation du modèle vis-à-vis du rôle attribué au contexte. Il semblerait qu'il faille lui reconnaître un plus grand rôle puisqu'il semble en grande partie responsable des coupes dans les saisies... Ce qui invaliderait le raisonnement discuté ici.

⁹⁴ Voir n. 63, p. 29, où je rappelle la distinction chez Ducrot.

Au-delà de ces considérations qui relèvent de l'opposition entre syntaxe et sémantique, il n'est peut-être pas totalement illégitime de se demander si le traitement sémantique qui s'appuie sur un raisonnement « énoncé hors-contexte » (en fait, un exemple fabriqué, court, dépourvu de tout modifieur ou adverbial, et de toute indication scénique), s'appuyant inconsciemment sur une tradition solidement ancrée, n'est pas teinté d'une certaine naïveté, chez le linguiste qui se fait livrer sur commande les exemples qui viendraient appuyer ses vues. J'y reviendrai plus longuement dans la 3^e partie de cette synthèse.

Un autre argument possible (mais quelquefois sous-jacent, ne serait-ce que par l'épithète *prototypique* qu'on lui attribue) consisterait à voir dans l'IMP un temps du passé parce que ses emplois temporels seraient les plus fréquents. Même si elle n'établit pas de lien entre ses données quantitatives et son traitement de l'IMP comme un temps du passé, Patard (2007) signale que dans son corpus (assez diversifié et quantitativement important) l'emploi temporel (auquel elle attache toutefois le contrefactuel) représente 80% du total des occurrences. Les données quantitatives peuvent-elles nous renseigner sur le sens premier de l'imparfait ? Ne sont-elles pas trop tributaires du genre, du corpus, des situations de communications, etc. pour qu'on lui donne sans précaution supplémentaire un rôle aussi déterminant ? D'autant qu'il est difficile d'attribuer avec certitude une telle proportion à l'imparfait et non aux contextes de sa production : que calcule-t-on au juste et comment faire la part des choses ?

Enfin, il convient de rappeler que s'il est vrai, ainsi que nous le verrons ci-dessous, que les modèles qui voient dans l'imparfait une forme verbale de décrochage par rapport à la situation d'énonciation butent paradoxalement sur l'emploi chronologique, il existe toute une série d'emplois qui « paraissent battre en brèche les approches aspecto-temporelles, offrant à ceux qui ne placent pas la référence temporelle au centre de leurs analyses des exemples de choix pour défendre leurs thèses » (De Saussure & Sthioul (2005 : 103). On peut, pour fixer les idées, illustrer ces classes d'emplois par les exemples empruntés suivants, qui sont loin d'être représentatifs de tous les cas identifiables ou identifiés, et de tous les sous-cas que l'on peut rencontrer dans une classe donnée d'emplois⁹⁵ :

(38) Imparfait 'hypothétique'

(a) Il songeait : - Si je *pouvais* tuer tous ceux qu'elle a aimés ! (A. France cité par Damourette & Pichon)⁹⁶

⁹⁵ Comme on peut s'y attendre, les listes d'emplois peuvent varier d'un auteur à un autre... Voir, en guise d'illustration, les notes qui suivent.

⁹⁶ Identifié quelquefois comme un emploi à part : imparfait exprimant un souhait ou un désir...

(b) Et dire que sans vous, j'*étais* tranquillement dans la cuisine en ce moment, ou à me chauffer au soleil dans le jardin en attendant la cloche du déjeuner. (M. Maeterlinck, cité par Damourette & Pichon)

(c) Mon paquet ! eh, dis donc, les bonnes femmes elles *dansaient* à poil ce soir, elles n'avaient pas de maillot ! (cité par Damourette & Pichon)⁹⁷

(39) Imparfait 'hypocoristique'

Dans un phylactère de la bande dessinée Drame de la jalousie (Brétécher, *Les Frustrés*, 5), un homme, qui garde les enfants de son amie en son absence, tente de consoler le petit garçon d'un gros chagrin en le prenant dans ses bras et dit :

– Ch'est un pauvre bébé cha ... il *avait* le cœur brijé ... (Barceló & Bres, 2006 : 62)

(40) Imparfait de 'décrochage subjectif' (Lebaud) /*ground* décalé (De Mulder)

Zut ! Dimanche prochain, il y *avait* un petit marathon sympa à Reims ! (Lebaud)

(41) Imparfait 'atténuatif', 'de politesse' ou 'de discrétion'

(a) Ecoute, Caroline, avant que tu me quittes, je *voulais* te demander quelque chose. (F. Mauriac, cité par Damourette & Pichon)

(b) Je *venais* te dire bonsoir. (A. Gide).

(42) Imparfait 'préludique' :

Après avoir choisi de jouer à l'« école des grands » (A est la maîtresse et B l'élève), A explique la situation de jeu à B.

A. mais tu prends pas le bébé parce que t- t- en fait i i faut i faut pas que ta maman vienne en fait et puis et puis ton papa ni et be et be i faut y aller tout seul d'accord ?

B. mais j'comprends plus rien moi hein !

A. ba au début tu *savais* pas //

B. au début je *savais* pas qu'est-ce que je *savais* pas au début ?

A. au début tu *savais* que / queum qu'en fait que moi j'*étais* // euh // le fait que moi je suis la maîtresse et toi tu es la grande sœur / mais tu apportes pas le bébé et puis moi je // d'accord mais tu apportes rien (Patard, 2007)

⁹⁷ Les emplois (a) et (b) sont souvent rangés dans les études les plus récentes dans une classe distincte sous le chapeau du *contrefactuel*. Mais cette même classe est rangée par Patard (2007) dans la classe temporelle...

Sans minimiser la pertinence et l'apport des études adverses, il me semble que la plupart des modèles non-temporalistes parviennent mieux et à moindre coût⁹⁸ à traiter les emplois dits modaux⁹⁹.

Enfin, reconnaissons tout de même que c'est face à l'emploi chronologique de l'imparfait, notamment en combinaison avec le PS, que les approches non-temporalistes rencontrent le plus de difficultés¹⁰⁰. Examinons, pour fixer les idées, l'exemple suivant de Zola, cité par De Mulder (2012), qu'il commentera à la suite :

(43) Il lui empoigna la tête, il la cogna contre un pied de la table. Elle se débattait, il la tira par les cheveux, au travers de la pièce, bousculant les chaises. Chaque fois qu'elle faisait un effort pour se redresser, il la rejetait sur le carreau d'un coup de poing. [...] La table poussée, faillit renverser le poêle. Des cheveux et du sang restèrent à un angle du buffet. Quand ils reprirent haleine, hébétés, gonflés de cette horreur, las de frapper et d'être frappée, ils étaient revenus près du lit... (*La Bête humaine*)

De Mulder affirme (2012 : 105) à juste titre qu'il n'est pas possible de soutenir que l'imparfait dans cet extrait exprime un déplacement du point de vue, comme c'est le cas des exemples suivants, qu'il emprunte respectivement à Arrivé, Gadet & Galmiche (1986 : 309) et Wilmet (1996 : 206) :

(44) (a) Dommage, il était beau ce vase.

(À propos d'un vase que l'on décide finalement de ne pas acheter)

(b) C'est bien vous qui parliez lors de la prochaine réunion ?

En effet, il paraît indiscutable que les procès à l'IMP dans la séquence (43) soient vus du même point de vue que ceux au PS. Les traitements inspirés du modèle de Langacker selon lesquels l'IMP désignerait une *réalité non immédiate* se trouveraient ainsi en difficulté. Liées au sein d'une même scène narrative, les deux formes temporelles exigeraient un traitement similaire que le modèle, sous peine d'invalidation, ne peut offrir en attribuant à l'IMP une propriété qu'il refuserait au PS.

En réalité, cette critique peut souvent être émise à l'encontre de la plupart des approches non-temporalistes dans ce type de contexte narratif, où la trame narrative fonctionne grâce à

⁹⁸ Les mécanismes descriptifs et explicatifs peuvent être (et quelquefois sont) investis pour traiter des phénomènes qui ne relèvent pas de la temporalité.

⁹⁹ Ainsi par exemple dans un cadre culiolien, le procès à l'imparfait est repéré grâce à l'opérateur ω qui signale une opération de rupture par rapport à l'origine énonciative. Dans ces emplois non-temporels, les procès qui seront posés comme validables verront leurs propriétés aspectuo-temporelles repérées par rapport au repère fictif (*) dans une opération où la valeur de rupture désigne une mise à distance par rapport au plan de l'énonciation...

¹⁰⁰Le constat est fait par Berthonneau & Kleiber (2006).

une répartition de rôles au sein du couple IMP-PS. Des solutions existent mais risquent d'être invalidées dans d'autres contextes. Ainsi, par exemple, si les descriptions en termes d'opposition de plans à la Benveniste (1959), ou textuelles à la Weinrich (1973), parviennent à traiter ce type de cas¹⁰¹, elles butent sur d'autres types d'emplois et d'autres difficultés... Je reviendrai encore une fois sur l'emploi temporel de l'imparfait en 2.2.2.3.

En bref, la complexité des données rend les traitements unitaires de l'imparfait difficiles, voire impossibles, ce qui pousse certains auteurs à revendiquer explicitement le rejet de toute démarche systématique, à l'image de Berthonneau & Kleiber¹⁰² qui écrivent (2006 : 7) :

il nous semble plus bénéfique, en ces temps où, dans la sphère temporelle à la mode, « Jean tombe [beaucoup] parce que Max le pousse », de procéder à l'inverse des approches systématiques habituelles. Au lieu de partir avec une définition bien formatée en tête et de prouver ensuite sa pertinence sur les différents types d'emplois relevés, il est, à notre avis, plus utile et plus rentable de s'attaquer directement à ces emplois pour en décrire le plus finement possible le fonctionnement.

Je partage avec Berthonneau & Kleiber le souci d'une prise en compte et d'un traitement scrupuleux des données. Plus encore, je pense, même si concernant l'imparfait ce n'est pour l'instant que programmatique, que c'est dans le retour aux données (leur nature, leur gestion, leur traitement, etc.) que peut désormais résider le progrès descriptif, et donc théorique, dans le champ de la temporalité. Je renvoie à la 3^e partie de cette synthèse où ces considérations seront longuement développées. Néanmoins, je suivrai Gosselin (1996 et 2005) pour affirmer qu'une pure démarche inductive ne réglerait pas toutes les difficultés, et qu'une démarche hypothético-déductive, pour moi complémentaire, est non seulement légitime, mais nécessaire. Le retour aux données, que cette démarche garantit, permet de valider ou d'invalidier les hypothèses, et donc les capacités descriptives et explicatives du modèle avancé. C'est dans ce cadre que s'inscrit notre vision de l'indicatif, ainsi que la proposition schématique qui suit.

2.2.2.3. *Portrait-robot d'une proposition*

Je considère l'imparfait comme une forme verbale **non saturée**, par analogie avec la métaphore chimique telle qu'elle a été empruntée, dans une syntaxe plus dynamique que

¹⁰¹ Et encore : voir les excellentes observations de Patard (2007 : 283 sq.).

¹⁰² On peut tout de même se demander si les auteurs y rejettent la démarche unitaire, comme ils l'affirment, ou la désinvolture avec laquelle certains travaux traitent les données. La première interprétation semble être contredite par les auteurs eux-mêmes qui revendiqueront un peu plus loin dans le texte une démarche que l'on pourrait qualifier d'unitaire dans le temps : les auteurs prennent successivement chacun des emplois de l'imparfait pour le traiter dans un long article à part (voir la bibliographie pour un petit aperçu de leurs travaux). On reviendra dans la 3^e partie de ce travail sur la seconde interprétation possible...

structurale, par Tesnière (1957) pour parler de la *valence*, terme également emprunté à la chimie.

Cette non-saturation, qui n'est pas sans rappeler le traitement anaphorique ni la notion d'incomplétude (Le Goffic, 1995), provient d'une autre caractéristique, soulignée dans beaucoup de travaux vus ci-dessus, qui voit dans l'imparfait un non-présent. C'est parce qu'elle a perdu son attache par rapport à la situation énonciative (t_0 , actualité, *nunc*, repère-origine... peu importe la terminologie ici), que cette forme cherchera à se fixer au premier repère référentiel disponible, quelle qu'en soit la nature. Et dans les énoncés réels, il y a toujours un repère disponible, même s'il faut quelquefois aller le chercher bien loin, comme on le verra ci-dessous.

Si cette approche, écartant ce que Confais (1990) appelle la thèse du signifié fondamental¹⁰³, évite les procédures de dérivation (ou de conversion, re-catégorisation, coercion, etc.¹⁰⁴), elle reste sous la menace d'une seconde objection que Gosselin (2010 : 114) adresse à l'égard de la démarche dérivationnelle mais qui ne perd pas totalement sa pertinence ici, celle de

vider presque intégralement de leur signification [les marqueurs], ou [de] ne pouvoir décrire cette signification de base que de façon purement négative (par ce qu'elle ne peut exprimer), de sorte qu'il devient extrêmement difficile de définir différentiellement les valeurs des marqueurs aspectuo-temporels, bien que ceux-ci ne soient pas librement interchangeables en contexte. De plus, la pauvreté de la signification en langue qui leur est reconnue rend pratiquement impossible le calcul de leurs effets de sens en contexte.

Nous est-il possible de préciser plus encore cette opposition PR/IMP, décrite vaguement en termes de « différence », de décrochage, de translation, de rupture, d'abandon du repérage par rapport à <Moi, Ici, Maintenant> ?

Se contenter d'une telle instruction sémantique pour l'imparfait, dès lors qu'il s'agit de qualifier la valeur en langue, ne constitue pas une dérobaie : il s'agit au contraire de prendre acte d'une sous-détermination qui existe dans la langue elle-même. En dire plus serait en dire trop et, automatiquement, privilégier, abusivement, un emploi au détriment des autres, qui sont tout aussi « naturels ».

¹⁰³ Qui attribue au marqueur une valeur première en langue et des sens secondaires et dérivés en contexte.

¹⁰⁴ Au-delà de l'inadéquation théorique de ces démarches, Gosselin a montré (2010 : 109 sq.) qu'elles sont inadéquates sur le plan empirique : en neutralisant (par effacement des traits sémantiques) certaines propriétés du marqueur, on perd des éléments nécessaires à l'interprétation globale de l'énoncé. Voir également Confais (1990).

Si cette vision, au même titre que toutes les approches non-temporalistes, ne devrait pas rencontrer de difficultés insurmontables à traiter les emplois dits modaux¹⁰⁵, elle devrait affronter l'emploi temporel, dont j'aimerais dire quelques mots.

Mais au préalable, un petit rappel concernant les propriétés aspectuelles de l'imparfait, en tout cas dans ses emplois chronologiques, et en dehors de cas de conflit (au sens de Gosselin). Il semble y avoir aujourd'hui unanimité chez les linguistes, quel que soit leur choix théorique et terminologique¹⁰⁶, pour considérer que l'imparfait marque l'aspect inaccompli, c'est-à-dire, dans le modèle de Gosselin, que « l'intervalle de référence est inclus dans celui du procès, dont les bornes ne sont pas prises en compte » (2010 : 162)¹⁰⁷.

Pour moi, cette propriété aspectuelle est centrale. La vision interne du procès nécessite le déplacement du repère référentiel qui peut être considéré comme un point de focalisation, une fenêtre, à partir desquels le procès est envisagé. En cela, le modèle de Gosselin dit la même chose¹⁰⁸, mais il se distingue en cela des traitements classiques et courants qui localisent directement le procès par rapport à l'intervalle d'énonciation (ici comme chez Gosselin le procès est calculé par rapport à l'intervalle de référence). Quel autre temps peut offrir cette possibilité, pour moi fascinante, de montrer des procès de l'intérieur ? Uniquement le présent me semble-t-il. Mais avec le présent, hors des cas de distorsion sur lesquels je reviendrai, il n'y a pas de déplacement de repère référentiel (ou d'intervalle de référence).

Cela mérite quelques précisions. Si on examine la grammaire temporelle de Gosselin (1996, tableau 1, p. 29), on constate qu'il attribue la propriété aspectuelle de l'inaccompli à quatre des huit formes temporelles qu'il examine : je donne pour chacune d'elles l'exemple proposé par l'auteur pour illustrer la combinaison aspectuo-temporelle :

¹⁰⁵ Sauf sans doute, et contre toute attente, l'emploi hypocoristique, dont le traitement me semble mieux assuré dans la perspective dialogique proposée par Brès (2003). L'élément central de son analyse consiste à voir dans l'énoncé contenant un IMP hypocoristique un énoncé échoïque comme si le locuteur parlait à la place de son interlocuteur. Mais j'irais plus loin : comme le montrent quelquefois les imitations phonétiques censées restituer la parole de l'enfant (ou de l'animal domestique, mais dans l'esprit du maître c'est équivalent), c'est l'enfant qui parle à travers la voix de l'adulte. Et quand un bébé parle, il parle « mal » : l'imparfait ne serait dans ce cas qu'une forme de transposition du présent (affinités aspectuelles obligent). Quelques observations comparatives avec l'arabe tunisien semblent le confirmer, non pas en termes de distorsions temporelles, mais en termes de déformations syntaxiques, lexicales et phonétiques... La palatalisation qui affecte des sons dans l'exemple (39) (p. 45) est à ce propos significative (et on peut même s'amuser à estimer l'âge des enfants représentés dans ces énoncés selon le type de déformation phonétique qui affecte les sons)...

¹⁰⁶ Vision sécante (Guillaume), vision interne (Martin), aspect sécant (Wilmet), imperfectif, etc. : on parle de la même chose, ou presque.

¹⁰⁷ Voir également Gosselin (1996).

¹⁰⁸ Y compris au niveau de l'arrière-plan conceptuel avec le concept de représentation.

(45) Imparfait (passé + inaccompli)

Luc se promenait.

(46) Présent (présent + inaccompli)

Luc se promène.

(47) Futur simple (futur + inaccompli)

A huit heures, Luc dormira depuis longtemps.

(48) Futur analytique (futur + inaccompli)

A huit heures, il va (être en train de) manger.

Signalons d'abord que le présent et l'imparfait peuvent avoir parallèlement à l'inaccompli une valeur aoristique (imparfait de rupture, présent historique, présent à valeur du futur). Ces cas ne présentent pas de difficultés particulières et peuvent par exemple être traités à l'aide de mécanismes de résolution des conflits. Et, en tout cas, ne jouent pas de rôle dans l'argumentation en cours. Ce qui est en revanche intéressant à signaler c'est qu'aucune instruction aspectuelle n'est codée dans le système de Gosselin pour les deux formes du futur. Autrement dit, l'inaccompli sera considéré par lui comme un effet de sens non-typique que peuvent produire ces formes contextuellement. Ce qui est par ailleurs visible y compris au niveau des exemples donnés : sans l'aide des modifieurs temporels (le futur simple a besoin pour rendre possible l'interprétation inaccomplie à la fois du syntagme prépositionnel *à huit heures* qui fixe le repère, et du syntagme prépositionnel *depuis longtemps* pour faire déborder le procès de part et d'autre du repère : sans ces éléments c'est l'aoristique qui s'impose).

En bref, si on fait abstraction du présent qui, puisque l'intervalle de référence est couplé à l'intervalle de l'énonciation, ne nécessite pas, dans les cas habituels, un déplacement du repère référentiel pour garantir une lecture aspectuelle de l'inaccompli, l'imparfait est la seule forme temporelle qui doit, pour assurer cette même lecture aspectuelle, décrocher l'intervalle de référence qui doit suivre le procès, de l'intervalle de l'énonciation (auquel il est arrimé par défaut).

Mais où doit-il le suivre ?

C'est là où ma vision diffèrera de celle de Gosselin. L'instruction temporelle [+ passé] que son modèle attribue à l'imparfait en tant que marqueur n'émerge, selon moi, que contextuellement. Et il me semble même que ce n'est pas l'imparfait qui en est responsable.

Que dit-on, en effet, au juste quand on affirme que l'imparfait localise le procès dans le passé ? Sert-il par lui-même à situer le procès dans le passé ? Est-il simplement compatible avec une localisation du procès dans le passé ?

L'hypothèse que je défends, à l'image de la plupart des modèles non-temporalistes comme il a été vu ci-dessus, consiste à retenir la seconde possibilité : c'est parce qu'il est indéterminé, décroché de l'énonciation, qu'un procès à l'imparfait, s'accrochera à des marqueurs chronologiques qui, eux, expriment par eux-mêmes le passé.

La valeur d'inaccompli en l'occurrence perturbe plus qu'elle n'éclaire : l'imparfait, ainsi qu'il a été observé dans de nombreux travaux, offre souvent le cadre, l'arrière-plan à des procès au PC (ou au PS). D'où un glissement qui consiste à voir en lui l'élément qui installe le cadre narratif dans le passé, dans lequel se déploient des procès plus ponctuels. On défendra le point de vue contraire : c'est parce qu'il est inaccompli qu'il peut fournir cet arrière-plan, mais ce n'est pas lui le responsable de la localisation dans le passé...

Dans les trois longues séquences suivantes, extraites d'ESLO1¹⁰⁹ (respectivement des modules « entretien », « conférences » et « repas »), j'ai souligné les éléments que je crois responsables de la localisation des procès dans le passé :

(49) ESLO1_entretien 001

- alors monsieur NPERS depuis quand habitez-vous à Orléans ?
- eh bien euh depuis l'âge de trois ans depuis ma plus petite enfance
- oui et alors c'est pas la peine peut-être de vous demander ce qui vous a amené à vivre à Orléans ?
- non ah non puisque c'est mes parents que enfin c'est ma mère mon père a été tué à la guerre de quatorze
- oui
- ma mère se trouvant seule elle est venue à la grande ville vous comprenez pour pouvoir euh y travailler alors que étant je suis natif de Houque-la-Joyeuse je sais pas si vous connaissez
- non
- euh c'est entre Marchenoire et Beaugency
- oui
- et alors là y *avait* pas d'éléments pour pouvoir travailler elle est donc venue puisque j'*étais* fils unique elle est donc venue à Orléans pour pouvoir trouver du travail

(50) ESLO1, CONF 503

l'enseignement des langues pour le professeur que j'ai été euh l'enseignement d'une langue vivante ça a d'abord été euh comme pour vous probablement quelque chose

¹⁰⁹ Nos exemples oraux sont extraits du corpus ESLO, qui sera présenté dans la 3^e partie de cette synthèse. Ces exemples sont disponibles (transcription et son) sur le site <http://eslo.huma-num.fr/>.

qu'on *entreprenait* avec certitude avec la certitude de réussir n'est-ce pas on nous avait donné euh des méthodes on nous avait donné un système on *pratiquait* de de mon temps euh ce qu'on *appelait* la méthode directe hein qui n'*était* pas du tout directe d'ailleurs et où on nous *obligeait* à certaines gesticulations à une petite comédie constante

(51) ESLO1_Repas 272

- mais bon on a toujours pas commencé euh de construire j'aurais dû commencer plus tôt au moins presque un mois à l'avance
- ben oui mais tu payes pas de location ici
- non
- ça c'est bon hein
- ah oui oui
- parce que je sais pas combien ça pourrait valoir ça un truc comme ça mais bon
- oh ben ça c'est pas à l'endroit où c'est <...>
- oh oh plus que ça
- oh pas tellement plus
- avec la la terrasse et tout le bazar dis donc
- ben oui mais tu as pas de commerçants tu as
- ben et alors
- regarde au Clos de Lorette tout au début donc quand on s'était renseigné pour se loger on *savait* pas si on allait emménager y *avait* au Clos de Lorette tu *avais* un F3 ils sont bien là-bas au Clos de Lorette
- oui
- un F3 avec garage alors c'*était* bientôt il y a deux ans remarque presque quarante-cinquante et un mille
- ça a dû augmenter
- ah oui ça a augmenté bien sûr maintenant je suis sûr que c'est ça doit être dans les soixante mille

Il me semble que dans ces exemples les procès à l'imparfait, à cause de leur indétermination référentielle, ne font que s'agglutiner à des contextes passés exprimés quant à eux par des marqueurs spécifiques du passé (principalement le PC et des adverbiaux temporels, ainsi que des éléments qui relèvent de l'environnement discursif et qui sont accessibles à la fois par les métadonnées et un co-texte plus large sur le site internet).

Le cœur du traitement esquissé dans cette synthèse consiste à affirmer que la lecture temporelle de l'imparfait n'est que l'une de ses possibilités interprétatives. Son indétermination référentielle l'oblige à s'attacher aux repères référentiels que le co-texte ou l'environnement discursif lui attribue. Je fais donc l'hypothèse que, dans une situation authentique de communication, il y a toujours un élément co-textuel ou contextuel qui permet l'interprétation chronologique de l'imparfait. Dans le cadre holiste que défend Gosselin, et

que je suis, ce fonctionnement serait naturel et attendu. Il s'agit donc d'une hypothèse empiriquement vérifiable : y a-t-il toujours un marqueur qui permette à l'imparfait d'avoir contextuellement cette lecture chronologique ? Existe-t-il des cas authentiques où aucun marqueur (y compris contextuel) ne vient donner à l'imparfait cette attache temporelle dans le passé dans lequel il parviendrait tout de même à situer le procès ?

Discutant l'approche anaphorique-méronomique de Berthonneau & Kleiber, Patard affirme l'existence d'énoncés où l'imparfait ne se raccroche pas à une situation antécédente, mais à un moment passé fourni par des modificateurs temporels. Mais elle ajoute aussitôt (2007 : 252) : « De plus, on peut observer que dans certains cas (notamment dans les textes littéraires), le procès à l'imparfait ne peut se rattacher à aucun antécédent, temporel ou conceptuel, même de façon rétrospective ». Et elle donne l'exemple suivant tiré de Modiano, *Une jeunesse* :

(52) [Début de chapitre] Elle suivait un couloir et lui parvenaient des bruits de conversation et des sonneries de téléphone. Des gens entraient, sortaient, claquaient des portes. Dans le bureau de Bellune régnait un grand calme et, si l'on était resté quelques instants sur le seuil, on aurait pu penser que personne n'occupait le local. Pas le moindre éclat de voix. Pas même un cliquetis d'une machine à écrire. Bellune, debout devant la fenêtre à guillotine, fumait. Ou bien il était assis sur le bras de l'un des fauteuils de cuir, écoutant des enregistrements de chansons à l'aide d'un magnétophone. Il lui demandait son avis, mais la musique et la voix étaient si basses qu'elle n'entendait presque rien. Un après-midi, elle l'avait même surpris qui regardait pensivement tourner la bande, sans juger nécessaire de mettre le son.

Le choix de l'exemple ne me convainc pas : à moins de ramener le travail linguistique à une activité stylistique, ou de nier toute spécificité créatrice aux œuvres littéraires, il me semble que le choix d'un exemple littéraire n'est admissible méthodologiquement que dans la mesure où il pourrait éclairer, éventuellement avec éclat – il est vrai que certains sont irrésistibles –, un mécanisme de langue par ailleurs parfaitement identifié. Mais le risque est réel, en voyant dans ce qui n'est peut-être qu'un exercice de style¹¹⁰ un fonctionnement authentique de la langue, de gommer les spécificités de l'un et de l'autre.

Si l'hypothèse de l'existence d'un imparfait temporel sans antécédent fourni contextuellement ne peut s'appuyer sur des textes littéraires, peut-on la vérifier sur des données orales authentiques ? Loin d'avoir exhaustivement exploré le corpus-test

¹¹⁰ Il ne viendrait à l'esprit de personne naturellement de fonder son analyse du PS sur l'examen du célèbre extrait suivant d'*exercices de style* de Queneau : « Ce fut midi. Les voyageurs montèrent dans l'autobus. On fut serré. Un jeune monsieur porta sur sa tête un chapeau entouré d'une tresse, non d'un ruban. Il eut un long cou. »

Imparfait¹¹¹, les dizaines de cas observés montrent que si les procès à l'imparfait ont toujours un élément « passé » auquel ils s'attachent, il faut quelquefois remonter bien loin pour le retrouver. Pour s'en rendre compte, on peut examiner l'encart intitulé *Etude d'un cas* (page 70 sq.), qui représente l'analyse d'une interaction (extraite du « Repas » 272, ESLO1), dont j'extrais trois passages (les passages supprimés par économie ne comportent pas d'imparfait)¹¹².

On pourrait trouver l'analyse qui y est développée, qui s'approche en l'occurrence du traitement anaphorique-méronomique de Berthonneau & Kleiber, peu convaincante et coûteuse. Sur le premier point, il convient d'abord de savoir que si j'expose mon analyse de cette occurrence, ce n'est certainement pas pour me livrer à « une certaine gymnastique théorique pour domestiquer l'emploi récalcitrant » (Berthonneau & Kleiber (2006 : 8)). Il serait contre-productif de s'arc-bouter sur un modèle donné : l'idée est de donner un exemple d'analyse d'un cas que je trouve difficile. D'autant que, même si je ne me suis pas livrés à un examen exhaustif d'un large corpus, l'occurrence analysée ci-dessus représente le seul cas « récalcitrant » sur des dizaines de cas examinés aléatoirement. En tirer des conclusions générales sur la validité du modèle ne me semble pas méthodologiquement légitime. Concernant la seconde critique envisageable, il me semble possible d'affirmer que le coût d'un traitement est proportionnel à la complexité des cas. Tous les cas examinés sont loin de nécessiter un traitement aussi lourd. En revanche, s'il est question de traiter des données authentique sans tri préalable (voir la fin de § 2.1.2, p. 49 et surtout la 3^e partie), il faudra sans doute s'habituer à l'existence de ces cas « récalcitrants » qui éprouveront tous les modèles¹¹³.

Plus généralement, il est vrai que la démarche ne peut être validée et la valeur retenue que si on parvient à décrire à grande échelle les différents emplois et prédire leurs sens en contexte. Pour cela, le modèle devrait s'appuyer sur l'étude d'un vaste corpus.

¹¹¹ Pour des raisons qui s'éclaireront ci-dessous, j'ai, dans une phase que je pourrais qualifier de pré-tests, utilisé le corpus que nous avons constitué avec Marie Skrovec pour l'étude du futur (cette étude ainsi que le corpus constitué seront présentés dans la 3^e partie) pour explorer rapidement les données qui relèvent de l'imparfait, avec un triple objectif : mesurer l'adéquation de ce corpus pour l'étude de l'imparfait, avoir une idée du nombre d'occurrences, et essayer d'établir la liste des propriétés à annoter. Ce travail, notamment la dernière phase, ne fait que commencer : après nettoyage, le nombre total d'occurrences à l'imparfait s'élève à 11485 (4463 dans ESLO1 et 7022 dans ESLO2).

¹¹² La transcription intégrale de ce repas, alignée sur le son, est disponible en ligne à l'adresse : <http://eslo.humanum.fr/CorpusEslo/html/ecoute/real?idEnregistrement=213&mode=enregistrement&idTranscription=1826>

¹¹³ Le projet *Temporal* (MSH-VdL 2013) auquel j'ai participé montre à quel point sont difficiles ces questions de l'annotation en référence temporelle des données orales. Voir Lefeuve & al. (2014) et, dans le CV, un descriptif du projet.

Etude d'un cas

Lors d'un repas entre deux couples d' « amis », à l'heure de l'apéritif, où les femmes s'affairaient encore visiblement dans la cuisine, s'engage spontanément un très long échange entre les deux hommes qui semblent passionnés par la mécanique et les belles voitures. Lors de cette interaction, une course souterraine s'engage spontanément entre l'Alfa Roméo – voiture de 272LOC1 – et la Simca 1100 – voiture de DM95. Nous étions le 19 mai 1970.

Il me semble que 272LOC1 voulait simplement, au début, montrer la puissance de sa voiture et la pertinence de son choix, sans chercher de confrontation avec son interlocuteur, dont il ignorait jusqu'à la nouvelle acquisition. Mais cela ne durera pas longtemps, et, ainsi qu'on tentera de le montrer, l'enjeu de l'interaction va se déplacer. Si, même après de très grandes coupures, les passages reproduits ci-dessous restent très longs, c'est pour permettre de mieux saisir l'enjeu caché de cette interaction, qui est le seul à pouvoir expliquer une occurrence de l'imparfait qui semble en apparence contredire nos prédictions.

(1) ESLO1-Repas 272_séquences 1

272LOC1: mais écoute moi en Alfa Roméo à cent cinquante ou à cent soixante j'ai je prends moins de risques que en quatre L à cent dix

DM95: oui oh ben remarque c'est sûr je hm

272LOC1: ça y a pas ça y a pas à discuter là-dessus

DM95: oui oui oui non ce qu'il faut avoir évidemment c'est c'est toujours une bonne réserve de puissance

272LOC1: c'est ça hm

DM95: mais alors quand tu es à fond en R4 à cent dix ben eh

272LOC1: tu es bloqué

DM95: moi je vois à cent dix avec la j'ai une [pi:pronounce:instantaneous] maintenant

272LOC1: ah c'est la tienne qui est là devant

DM95: oui

eh bien ça va y a du y a du répondant

272LOC1: mais quand est-ce que tu t'es payé ça

c'est pas le genre de voiture qui te va pourtant je trouve

DM95: ben avec quoi tu m'aurais vu

272LOC1: ben je sais pas une deux cent quatre euh

DM95: oh ben c'est équivalent

272LOC1: coupé

oui une deux cent quatre coupé ou

DM95: ben c'est la place qu'est-ce que tu veux

DM95 cherchera alors à justifier, au cours de plusieurs tours de parole, son choix, motivé essentiellement par sa situation familiale (il lui fallait suffisamment de place). Avant de tenter une contre-attaque, à l'imparfait :

(2) ESLO1-Repas 272_séquences 2

DM95: bon alors un coupé deux cent quatre et puis c'est marrant la deux cent quatre ne me **disait** rien

272LOC1: elle te dit rien ?

DM95: oui

alors comme dans tous les essais de la le coupé onze cent à la onze cent est est nettement devant

272LOC1: ouais

DM95: hein

et puis j'avais travaillé chez Simca alors c'**était** plus facile qu'on que j'en trouve une

Et la bataille reprend de plus belle. Sur plusieurs tours de parole, 272LOC1 ne lâche rien (tout y passe : la consommation, l'ergonomie, la forme, la boîte de vitesse), mais DM95 ne se rend pas, se défend, sans parvenir à convaincre son interlocuteur :

(3) ESLO1-Repas 272_séquences 3

272LOC1: ben je sais pas enfin j'aurais choisi entre la la la deux cent quatre et la onze cent j'aurais préféré la deux cent quatre quand même comme forme euh

DM95: elle consomme plus hein

272LOC1: oh non

DM95: si

272LOC1: ben penses-tu

[...]

272LOC1: ben tu sais la deux cent quatre tu la conduirais dans les mêmes conditions c'est pareil hein

DM95: je crois pas ben dans les essais si tu veux la la onze cent est donnée entre sept et onze litres et la deux cent quatre entre huit et vingt litres

[...]

272LOC1: la deux cent quatre est plus jolie enfin bon

DM95: oui oui un peu plus oui mais moi je enfin c'est

272LOC1: elle fait plus jeune que la Simca onze cent tu vois euh un peu genre quatre L un peu tu vois voiture euh

[...]

DM95: moi je sais pas elle ne m'a jamais tellement emballé

272LOC1: mouais c'est marrant

DM95: nous on aurait pris Mercedes

après c'est un peu plus gros hein

272LOC1: tu te la joues toi

Devant un tel acharnement, DM95 tente alors l'esquive, qui finira par payer, parce qu'il reconnaît un échec, celui de ne pas être parvenu à vendre sa vieille R4 : et l'interaction se focalisera là-dessus pendant une longue séquence (que je ne reproduis pas ici, puisque les occurrences de l'imparfait en emploi temporel qu'elle contient sont systématiquement attachées à des marqueurs de passé) :

(4) ESLO1-Repas 272_séquences 4

DM95: moi ce qui me fait suer cependant c'est cette pourriture de R4

272LOC1: là là tu me surprends vraiment hein

De ces quatre longs extraits, seule la séquence 2 contient deux occurrences de l'imparfait, dont la dernière ne pose pas de problème dans le cadre de l'hypothèse défendue ici. La première occurrence (*ça me disait rien*) semble en revanche *a priori* n'être attachée à aucun repère temporel passé fourni dans le co-texte antérieur (dans l'extrait coupé qui précède immédiatement la séquence 2, il n'est question que de la Simca 1100).

Que fait cet imparfait ici ? A quoi réfère-t-il ? Et à quoi est-il attaché ?

Il me semble que l'enjeu principal de l'échange concerne le **choix** de DM95, que 272LOC1 trouve **mauvais**, de la Simca 1100. Peu importe la symbolique que l'on peut être tenté d'attribuer à ce bras de fer. D'un point de vue strictement linguistique, de très nombreux indices montrent que l'imparfait ne crée pas, mais pointe vers un moment passé, qui, certes n'est pas spécifiquement marqué linguistiquement en tant que tel, mais que DM95 s'acharne d'une manière souterraine à en montrer l'existence : le moment du choix. A commencer par l'expression verbale elle-même (« la 204 ne me disait rien ») : *ne rien dire* laisse entendre (ce qui sera confirmé dans la suite de l'échange) que le locuteur a envisagé l'acquisition d'une 204 (il pouvait donc se l'offrir) mais qu'il a écarté cette option. Dans le micro-échange en question, il est remarquable que cet imparfait soit repris chez 272LOC1 par un présent (DM95 : la 204 ne me disait rien _ 272LOC1: elle te dit rien ? _ DM95: oui), une manière de demander à l'adversaire : et tu l'assumes toujours, ton choix ?

Il s'agit là d'une tâche que je trouverais intéressant de mener mais dans le cadre d'un projet collectif. Non seulement à cause de son ampleur et de sa complexité, mais aussi parce que, ainsi qu'on le verra plus clairement dans la 3^e partie, la phase, incontournable, d'annotation de corpus exige que l'on soit au moins deux, plus si possible, pour réguler le mieux possible la part subjective dans l'activité interprétative¹¹⁴. A ce titre, l'étude que, depuis bientôt deux ans, nous menons, Marie Skrovec et moi, sur la distribution à l'oral des deux formes analytique et synthétique du futur peut être vue, malgré la grande quantité de travail qu'elle nécessite (à la fois à cause de l'importance quantitative des occurrences et de la complexité des données) comme une première expérience qui, au-delà des résultats spécifiques au futur, nous permet de forger une démarche concrète et précise, transposable, et donc susceptible de s'appliquer à l'étude d'autres formes temporelles dont l'imparfait. Pour moi, un tel projet pourra valoir la peine : l'étude en corpus – oral et situé¹¹⁵ – de l'imparfait¹¹⁶ constitue sans conteste, pour les raisons expliquées ci-dessus, le verrou qui pourrait permettre un progrès dans le champ de la temporalité. C'est à ce prix, me semble-t-il, qu'on parviendra à préciser comment des éléments du co-texte (y compris de l'environnement syntaxique, mais pas uniquement) ou de l'environnement discursif (Haillet, 2007), déclinent cette valeur abstraite de l'imparfait en un sens contextuel *relativement* précis¹¹⁷.

L'identification elle-même des différents types (ou classes) d'emploi de l'imparfait est une question à la fois théorique et empirique. Sur le plan empirique, on sait que le corpus et les données qu'il contient ou ne contient pas dépendent d'un nombre considérable de paramètres (genre, situation de communication, registre, thèmes discursifs, profils des locuteurs, etc.). Certaines données, on le sait, ne sont pas facilement saisissables : comment, par exemple, obtenir suffisamment (en quantité et en représentativité) d'énoncés contenant un imparfait atténuatif alors qu'ils n'apparaissent que dans des situations de communication fugaces, où l'enquêteur n'est pas censé être présent ? Quels problèmes techniques et juridiques devons-

¹¹⁴ On pourrait penser qu'il s'agit là d'un vœu naïf. Il n'empêche que la tâche d'annotation, particulièrement ingrate et fastidieuse, est menacée par une subjectivité tentée, à la longue, de simplifier les données, de minimiser certaines spécificités, et de forcer les occurrences à entrer dans des groupes identifiés...

¹¹⁵ Entre autres exigences. J'y reviendrai.

¹¹⁶ Il existe depuis une quinzaine d'années des études sur l'imparfait qui se fondent sur des corpus assez conséquents et qui contiennent en partie des données orales. Le mérite revient d'abord à l'approche dialogique (voir notamment Brès 2005 et Patard, 2007) qui a été, me semble-t-il, la première à intégrer dans ses analyses des exemples d'interactions orales authentiques. Notre vision sur le corpus est toutefois différente, et sera mieux précisée surtout dans la 3^e partie de cette synthèse.

¹¹⁷ L'expérience que j'ai de l'annotation des occurrences du futur montre qu'il n'est pas toujours possible de leur attribuer un sens contextuel unique. Et ce n'est pas toujours uniquement à cause des taxinomies existantes (en l'occurrence nous les avons toutes *a priori* écartées), mais aussi à cause d'une certaine indétermination, une ambiguïté subtile, qui permet à des sens contextuels parfaitement identifiés par ailleurs de coexister dans certains cas...

nous affronter pour garantir un accès à ce type de données, ou à des données plus privées (l'hypocoristique, sans doute aussi le préludique, etc.)¹¹⁸. Sur le plan théorique, les questions ne sont pas moins complexes. Dans cette nébuleuse sémantique qu'est un énoncé réel, le marqueur, ainsi que le montre la théorie de la Polysémie Contextuelle Généralisée de Gosselin (2005), n'est isolé ni des autres marqueurs (et de leur propre polysémie) ni du contexte et de ses divers effets pragmatiques. Dans ces conditions, comment parvenir à distinguer ce qui relève du contexte, de ce qui relève du sens holiste de l'énoncé, et de ce qu'on laissera au marqueur en guise de sens contextuel spécifique ? Il « est bien arbitraire, avait déjà écrit Ducrot (1984 : 50) de déterminer, à l'intérieur d'un énoncé donné, quel est le sens qu'y possèdent, pris un par un, les mots dont il est constitué ». A *arbitraire*, Gosselin ajoutera *illusion* (1996 : 188) : « Et il y a une bonne part d'arbitraire et d'illusion dans les grammaires qui dressent la liste des effets de sens d'un marqueur donné, comme s'il en constituait le support unique. »

Est-ce à dire qu'une telle tâche est impossible ?

Aussi délicate et ardue soit-elle, cette tâche pourra désormais bénéficier de l'avènement des humanités numériques, qui non seulement peuvent aider à rendre disponibles les données pertinentes auprès de la communauté linguistique, mais aussi peuvent fournir les outils informatiques qui permettent de les traiter d'une manière ergonomique : naviguer aisément dans une masse de données structurables à l'infini, en fonction des hypothèses linguistiques, quel progrès depuis les fiches d'exemples de Damourette & Pichon !¹¹⁹

¹¹⁸ Voir sur ces questions juridiques et éthiques Baude (coord., 2006).

¹¹⁹ L'avènement de cette ère numérique permet un accès aisé à un autre type de ressources, moins souvent signalé mais dont l'importance est capitale : les travaux et recherches linguistiques les plus récents. En plus des archives ouvertes et des plates-formes d'échange entre chercheurs (notamment www.academia.edu et www.researchgate.net), le statut d'UMR obtenu par le LLL nous permet désormais l'accès à l'intégralité des contenus disponibles sur le Portail d'information scientifique des unités CNRS en Sciences humaines et sociales (<http://biblioshs.inist.fr>). Il n'est pas ici nécessaire de remonter à Damourette et Pichon pour mesurer le gain en productivité que cela représente pour le chercheur, vingt ans suffisent : aller à la Bibliothèque de la Sorbonne depuis mon studio dans le 20^e arrondissement était une expédition qui se prépare ! Je me souviens de la file d'attente dans l'escalier menant à la porte de la bibliothèque nichée au 1^{er} étage (la capacité d'accueil étant limitée) ou devant le petit meuble en bois (il y en avait un seul) où étaient rangées les fiches classées par ordre alphabétique, de la mauvaise humeur des agents (il y en avait qui étaient plus sympas que d'autres) à qui il ne fallait pas demander trop d'ouvrages à la fois (mais il n'était pas facile de lutter contre cette fichue tentation de vouloir éviter de refaire la queue), sous peine de voir sa fiche barrée avec un commentaire lapidaire et à peine audible (« pas disponible ») et leurs regards se porter déjà sur la personne suivante (il faut reconnaître qu'ils étaient très forts, et sans le moindre support informatique, ces agents pouvaient en quelques secondes se prononcer sur la disponibilité du document). Je ne parlerai pas de la petite salle où étaient parquées les trois photocopieuses en surchauffe (j'ai oublié de dire que j'avais un talent particulier pour réclamer des ouvrages non-empruntables) ni des petites pressions exercées par des gens qui attendaient derrière et qui voulaient qu'on veuille lâcher une photocopieuse qu'on avait mis trop longtemps à avoir (parce qu'il y avait là deux files d'attente : une pour se procurer la carte de photocopies, et une autre pour accéder aux machines)...

Naturellement, l'aide de l'informatique ne réglerait pas tous les problèmes signalés, mais elle peut y contribuer d'une manière conséquente. J'y reviendrai encore une fois dans la 3^e partie.

Résumons-nous.

Mon hypothèse¹²⁰ consiste à scinder la catégorie traditionnelle « temps » en deux sous-catégories distinctes :

- une première catégorie (Y) qui fixe le repère énonciatif à partir duquel se déploie la temporalité des procès. Cette catégorie *énonciative* permet d'opposer deux groupes de Formes Temporelles, selon le marqueur choisi ($\emptyset$ ou ϵ).

1. *Hors distorsion*¹²¹, le premier groupe (qui a pour centre le présent) est celui de l'assertion dans la sphère <Moi-Ici-Maintenant>. J'y inclus, en plus du présent, le PS, le FS¹²² et leurs formes composées ou périphrastiques correspondantes. Ces formes s'opposent, mais à l'intérieur d'un même paradigme énonciatif, par l'instruction temporelle (et éventuellement aspectuelle) que chacune d'entre elles code. Ici, aussi, on suivra Damourette & Pichon¹²³ qui parlent de localisation

¹²⁰ Il me semble que le modèle, tel qu'il est formulé ici, peut faire l'objet de vérifications psycholinguistiques relatives aux étapes d'acquisition et de développement de la temporalité chez l'enfant. Gosselin (2005 : 24 sq.) présente une traduction en français des 4 étapes d'acquisition distinguées par R. Weist dans le cadre théorique de H. Reichenbach (S : moment de l'énonciation, E : moment de l'événement, R : moment de référence) comme suit :

Etape 1 : « speech time system » (entre 1 an et 1,6 environ) : E = R = S (les trois points coïncident).

Etape 2 : « event time system » (entre 1,8 $\pm$ 2 et 2,8 $\pm$ 2) : R = S, E est libre (i.e. il peut être antérieur, simultané ou postérieur par rapport à S) ; apparition, en français, du passé composé – avec des procès téléliques – et du futur périphrastique : [aller + Vinf].

Etape 3 : « restricted reference time system » (de 3,0 $\pm$ 2 à 4,0 $\pm$ 2) : R devient libre relativement à S, mais lorsqu'il est antérieur ou postérieur à S, il coïncide nécessairement avec E. En d'autres termes, R et E sont libres par rapport à S, mais (R $\neq$ S) $\implies$ (R = E) ; utilisation de circonstanciels de localisation temporelle, à l'exception de ceux qui indiquent une localisation relative du type « avant », « après »).

Etape 4 : « free reference time system » (à partir de 4 ans environ)

Les trois points sont libres ; apparition, en français, du PQP et des prépositions avant, après.

Il est naturellement difficile d'interpréter les résultats des rares (me semble-t-il) études disponibles sur l'acquisition de la temporalité chez les enfants francophones monolingues, et il est de toute façon nécessaire d'avoir un modèle linguistique pour les interpréter, distinguer ce qui relève du marquage morphologique de ce qui relève du marquage adverbial. Voir pourtant Labelle (1994) dont l'étude, adaptant la théorie de Weist dans le cadre de Vet apporte quelques arguments à la vision développée ici. Des vérifications supplémentaires sont naturellement requises.

¹²¹ La distorsion temporelle sera abordée en 2.3.2.

¹²² Je suis entièrement d'accord avec Gosselin (1996) et (2005) pour m'opposer à la conception, courante dans différentes écoles, qui consiste à établir un lien entre l'incertain et le futur. Ce traitement ne tient ni sur le plan conceptuel (ne pas confondre réalité et représentation linguistique), ni sur le plan empirique (on peut présenter un procès futur comme certain ou un procès passé comme incertain...).

¹²³ L'étonnement amusé qui accueillait le renvoi il y a vingt ans à Damourette & Pichon (alors que le renvoi à Guillaume, leur contemporain, était courant) s'estompe : plusieurs linguistes s'y réfèrent désormais explicitement, et même ceux qui ne retiennent pas leurs analyses reconnaissent une œuvre « très moderne » (Patard, 2007 : 219)...

temporelle « sèche » particulièrement visible au PS. On ne développera pas davantage la vision aspectuelle qu'un tel système ne manquera pas d'impliquer : on pense qu'il est possible de se passer de toute caractérisation aspectuelle pour décrire le PS¹²⁴. Il ne serait dans cette vision qu'une forme purement temporelle. En cela je serais sans doute plus guillaumien que les disciples de Guillaume, dont « [b]eaucoup [...], ainsi que le constate Wilmet (1997 : 332), n'ont pas accepté sans états d'âme qu'il rejette en dehors du domaine ne fût-ce que le couple *marchai/marchais* ». L'une des principales motivations qui a poussé à traduire ce qui n'était chez Guillaume qu'une caractérisation négative (le global par opposition au sécant, seul porteur parmi les deux formes d'une valeur aspectuelle) en une caractérisation positive est le maintien au sein du système d'un rapport différentiel au sein de ce couple¹²⁵. Dans la vision défendue ici, il ne s'agirait pas d'un couple, leurs différences se situant ailleurs¹²⁶, ils n'auraient pas besoin de s'opposer aspectuellement pour maintenir la « symétrie » du système, à laquelle beaucoup semblent très attachés.

2. Le second groupe a pour centre l'imparfait et contient également le conditionnel et leurs formes composées (ainsi que la forme périphrastique *aller* à l'IMP + inf). Il s'agit de formes temporelles *a priori* indéterminées parce qu'elles codent des instructions aspectuo-temporelles calculées par rapport à un intervalle de référence lui-même *a priori* indéterminé. Leur repère référentiel se définissant en langue uniquement par sa différence (décrochage) par rapport à l'intervalle d'énonciation, ces formes sont susceptibles de prendre toutes sortes de significations contextuelles (aussi bien énonciatives que chronologiques). On comprendra ainsi pourquoi le phénomène de la polysémie temporelle, s'il est général, s'observe surtout avec ces formes temporelles. On comprendra aussi pourquoi, contrairement à la localisation sèche du PS (où le procès passé est, selon moi, vu du présent),

¹²⁴ Le pas, nous l'avons vu, a été franchi dans le système de Gosselin pour le futur simple qui n'y code aucune instruction aspectuelle en langue...

¹²⁵ On peut s'en rendre aisément compte en lisant Mellet (1981).

¹²⁶ On entrevoit naturellement d'autres conséquences en termes d'évolution linguistique : quand on fait le rapprochement avec ce qui est sans doute en train de se passer au sein d'un autre couple, FP-FS (voir Abouda&Skrovec (2015, à paraître), difficile de ne pas voir dans le déclin, quoi qu'on en dise, du PS face au PC un même phénomène, que Fleischman (1983) tente d'en rendre avec son hypothèse de *present/current relevance*, mais qui semble à nos yeux plus motivé par une pauvreté aspectuelle (je n'ai pas osé parler d'*appauvrissement aspectuel*...). Les deux hypothèses ne s'opposent pas toutefois frontalement.

l'imparfait, lorsqu'il parvient à localiser contextuellement dans le passé, donne cet effet de sens « autre époque¹²⁷ » (*c'était mieux avant, c'était une autre époque*)...

- une seconde catégorie (X), que l'on peut qualifier d'aspectuo-temporelle, et qui donne à la forme temporelle ses caractéristiques temporelles (présent, passé, futur, ou, mieux encore, si on exprime le phénomène en termes de relations temporelles entre le procès et le repère référentiel, coïncidence, antériorité et postériorité) et, éventuellement, aspectuelle (essentiellement accompli, pour les formes composées, inaccompli pour le PR et l'IMP, et prospectif pour le futur et le conditionnel).

Nous retrouvons là les relations temporelles habituelles qui s'opèrent entre les repères référentiels (ou intervalles de référence dans le modèle de Gosselin) et les procès.

Cette vision schématique (car elle laisse ouvertes différentes possibilités) me semble offrir de nombreux avantages, et sans doute présente-t-elle aussi des difficultés qu'on n'entrevoit pas encore. Parmi les avantages visibles, on peut évoquer ici son économie : elle met en place moins de catégories : non seulement, elle abandonne la catégorie de l'aspect global puisque la simple localisation dans le temps suffit à décrire le fonctionnement des formes concernées, mais en plus elle ne propose le recours à un intervalle de référence autre que S (le fameux point référentiel R de Reichenbach) que s'il y a décrochage par rapport à la situation d'énonciation¹²⁸. Autrement dit, pour le premier groupe de formes temporelles (où R=S), je propose de calculer la relation temporelle du procès directement par rapport à S, ce qui est, me semble-t-il, en cohérence avec la vision globale esquissée ci-dessus. En plus de son économie, cette vision semble garantir une cohérence sur un point difficile, nous l'avons déjà vu au niveau morphologique, et auquel la plupart des modèles ont dû apporter une réponse qu'on peut juger *ad hoc* en proposant un repère spécifique pour le conditionnel : ici, le conditionnel se déploie à partir du même repère référentiel que l'imparfait¹²⁹.

¹²⁷ Le *Il était une fois* ne localise pas dans le passé (malgré les formules trompeuses qui peuvent suivre l'expression comme *il y a très très longtemps* ou encore, en arabe, *كان يا مكان في قديم الزمان في سالف العصر والوان* - littéralement « était ce qui était dans l'ancien temps en une époque précédente ») mais ouvre un univers différent, magique, celui du conte... Mais, on le sait, ne peut être raconté que ce qui se donne pour passé.

¹²⁸ Cette caractéristique permet de se prémunir contre la critique la plus forte à laquelle le système de Reichenbach (1947) a dû faire face : proposé dans une démarche onomasiologique, son système paraît trop puissant par rapport à ce qui est généralement observé dans les langues (sans d'ailleurs parvenir à expliquer les différences entre IMP et PS par exemple) : il prévoit 9 possibilités temporelles résultant des différentes combinaisons chronologiques entre les trois points S, R, et E (de *event point*).

¹²⁹ On pourrait sans doute évoquer aussi le PC : le système proposé permet de rendre compte de ses deux valeurs repérées depuis longtemps par les linguistes : accompli du présent (sa valeur d'origine, avec l'effet résultatif qu'on lui reconnaît) et temps du passé où il tend à remplacer le PS...

Ce sont les raisons qui m'ont poussé à qualifier dans (Abouda 1997a) le premier groupe de formes temporelles *monochroniques* (car elles ne s'appuient, en dehors des cas de distorsion, que sur le repère S) par opposition au groupes de formes temporelles *polychroniques*, imparfait et conditionnel (et leurs formes composées correspondantes) qui, se déployant par rapport à un repère R qui se définit lui-même différentiellement par rapport à S, nécessitent une prise en compte des deux repères à la fois. Ce qui complique alors particulièrement la donne, c'est la possibilité qu'ont les formes temporelles monochroniques d'être employés, dans les cas de distorsion, polychroniquement...¹³⁰

2.3. De la polyphonie à la polychronie

Je me rappelle du jour où je me suis mis à lire le chapitre 8 du *Dire et le dit*. Ce devait être au même moment, sans doute, que Haillet faisait la même rencontre. Pour celui qui travaille sur le conditionnel, la théorie polyphonique vient comme une bouée de sauvetage, un moyen possible de mettre de la cohérence dans un univers qui paraît chaotique.

Initiée par Bakhtine dès la fin des années 1920, la théorie polyphonique n'arrivera en France que beaucoup plus tard. Traduits en anglais et en français à partir des années 1970, les ouvrages de Bakhtine auront une influence considérable en France dans le champ de la stylistique et de la critique littéraire¹³¹, avant de parvenir chez les linguistes.

La théorie polyphonique a connu un vif succès dans les années 1990 et surtout dans les années 2000, où de nombreux chercheurs se sont attaqués à différents phénomènes linguistiques, armés d'un cadre théorique esquissé par Ducrot. Les successeurs le préciseront ou le modifieront sur certains points, quelquefois le développeront d'une manière conséquente comme dans le cadre de la Scapoline¹³², projet collectif qui dure depuis plus de vingt ans.

Au même moment va se développer en France une autre école, le dialogisme, qui se revendique explicitement de Bakhtine¹³³. Elle connaîtra également un vif succès, notamment dans le champ de la temporalité.

¹³⁰ Cela nous permet, me semble-t-il, de tenir compte de la remarque suivante de Vetters (1996 : 50) : « On peut donc conclure qu'Yvon a raison quand il remet en question l'appartenance exclusive des tiroirs à l'une des deux catégories. Il se trompe pourtant lorsqu'il avance qu'un tiroir relatif comme le plus-que-parfait peut être employé de façon absolue. Les exemples que je viens de donner suggèrent exactement le contraire : ce sont les tiroirs absolus qui peuvent être employés de façon relative. »

¹³¹ Voir J. Kristeva (1969), et notamment le concept d'intertextualité.

¹³² La théorie SCAndinave de la POlyphonie LINGuistiquE. Avec un nombre conséquent de publications, dont il est ici impossible de rendre compte : voir les références de Nølke dans la bibliographie.

¹³³ Voir Brès dans la bibliographie et les références auxquelles il renvoie.

Si on fait le lien, tellement les similitudes sont frappantes malgré les différences connues dues essentiellement aux arrière-plans théoriques, avec les travaux de Jacqueline Authier-Revuz, on peut bien parler d'une exception française, suivant en cela Brès & Mellet (2009)¹³⁴.

2.3.1. Polyphonie linguistique

Comme l'écrit à juste titre Donaire (2006 : 61), « la polyphonie est devenue elle-même polyphonique ». Il n'est pas question de rendre compte de la diversité des approches que l'on peut enregistrer au sein du courant polyphoniste, ni même de présenter notre propre cheminement au sein de cette école. On se contentera donc de renvoyer aux références présentes dans la bibliographie, et notamment celles de Donaire, Nølke et de Haillet, dont la conception me semble sur de nombreux points proche de la vision défendue ici. L'objectif de cette section est simplement de fournir un kit, que j'espère léger, qui permette de situer le prolongement que j'ai cherché à donner à cette théorie dans le domaine temporel sous le nom de *polychronie* (Abouda 1997a, 2003 et 2004c).

On s'appuiera ici pour cette présentation rapide de la théorie polyphonique essentiellement sur la version de Ducrot (1984), même si sur de nombreux points les recherches polyphoniques actuelles¹³⁵ s'en écartent (notamment sur la question des énonciateurs), parce qu'elle me semble la plus pédagogique et la plus facilement exportable dans le domaine de la temporalité.

Le point de départ de cette théorie est le rejet par Ducrot de ce qu'il appelle *la thèse de l'unicité du sujet parlant*, c'est-à-dire le postulat, très généralement admis plus ou moins explicitement, qui consiste à voir comme une entité unique à la fois la personne réelle en chair et en os qui parle, l'entité discursive qui se désigne au sein de son énoncé par *je* et enfin l'être discursif qui prend en charge le contenu de l'énoncé. L'unicité en question est maintenue grâce à un lien de naturalité qui fait que dans les cas habituels et sans doute les plus fréquents les trois entités coïncident. L'existence de cas, même rares, de non-coïncidence entre ces trois entités suffit en bonne méthode à justifier leur distinction systématique sur le plan théorique.

¹³⁴ Ils notent (2009 : 9) : « L'explication nous semble tenir à un fait bien connu [...] : alors qu'une partie de la linguistique en France s'est construite, à la suite notamment des travaux de Benveniste (faisant eux-mêmes suite aux recherches de Damourette et Pichon, de Bally, de Guillaume), sur l'opposition énoncé/énonciation, il n'en a pas été de même en domaine anglo-saxon, ce que signale par exemple la difficulté à traduire le terme d'*énonciation* (chez certains, *utterance* ; chez d'autres, par emprunt au français et intégration phonologique, *enunciation*). Le retravail des notions de dialogisme ou de polyphonie dans les cadres d'une approche de linguistique énonciative ne s'est donc pas produit. »

¹³⁵ Voir notamment Donaire (2006), Haillet (2007) et Nølke (2009).

Ducrot distingue, se cachant derrière la notion de sujet parlant, trois entités :

- I. le producteur réel du message ; noté ici par P (de Personne) ;
- II. le locuteur, noté L, qui correspond à l'être discursif qui est à l'origine de l'énoncé, et à qui renvoie *je* ;
- III. et, enfin, l'énonciateur, noté E, d'abord défini comme étant l'être discursif qui accomplit les actes illocutoires d'assertion, d'interrogation, d'ordre, etc., et, ensuite pour des raisons internes à la théorie, comme le responsable du contenu de l'énoncé, c'est-à-dire l'entité qui prend en charge ce contenu.

Plusieurs exemples ont été fournis pour illustrer les différents cas de disjonction entre P, L, et E. Pour visualiser la disjonction P/L-E (entre le producteur physique de l'énoncé et le locuteur, sans distinction pour l'instant avec l'énonciateur), on peut, en suivant Ducrot, dresser le tableau suivant (auquel j'ai adjoint la dernière ligne¹³⁶) :

(53)

Personne réelle	Locuteur
Auteur	Narrateur
Acteur	Personnage
Lecteur	Narrataire

S'appuyant sur Genette (1972), Ducrot fait plusieurs observations sur la différence entre auteur et narrateur ayant des répercussions directes dans le domaine de la temporalité¹³⁷. Prisonnier de sa chair, l'auteur ne peut exister que dans une époque donnée, alors que le narrateur, en tant qu'individu fictif ou discursif peut vivre dans le Passé, le Futur, ou le Présent. Dans un roman d'anticipation, l'auteur enverra donc dans le Futur ou le Futur du Futur un narrateur pour que le récit puisse avoir lieu. C'est commode : en plus d'être omniprésent, le narrateur peut continuer à raconter même quand il est mort...¹³⁸

¹³⁶ En effet, il me semble possible, m'appuyant sur Jauss (1978), d'étendre la comparaison au couple lecteur-narrataire, sur lequel on peut faire des observations équivalentes.

¹³⁷ Sur ces questions, voir également Weinrich (1973) et Vuillaume (1990) et (1993).

¹³⁸ Quant à la confusion entre acteur et personnage, Gabriel Garcia Marquez montrera avec éclat à quel point cela peut s'avérer dangereux. « Ils furent indignés par les vivantes images que le riche commerçant [...] projetait dans le théâtre [...] à cause d'un personnage mort et enterré dans certain film, sur le malheur duquel on versa des larmes amères, et qui reparut bien vivant et métamorphosé en arabe dans le film suivant. Le public, qui payait deux centavos pour partager les retours de fortune des personnages, ne put supporter cette inqualifiable moquerie et brisa tous les sièges. Le maire, cédant aux instances de don Bruno Crespi, dut faire expliquer par l'annonceur public que le cinéma n'était qu'une machine à illusions, laquelle ne méritait pas ces débordements passionnels du public. » Gabriel Garcia Marquez, *Cent ans de solitude*, éditions du Seuil, (1968 : 237).

Quant à la distinction entre L et E, considérée par la plupart des auteurs¹³⁹, y compris par Ducrot lui-même, comme la seule qui devrait intéresser les linguistes, elle est à l'origine *stricto sensu* de la théorie polyphonique, dont le principe de base consiste à admettre l'existence du phénomène de l'altérité énonciative, i.e. la possibilité qu'au sein d'un même énoncé coexistent, hiérarchisées par le locuteur, différentes instances énonciatives, assumant chacune un point de vue. L'un des exemples que donne Ducrot pour illustrer sa distinction me paraît particulièrement éclairant de la démarche :

(54) Ah, je suis imbécile ; eh bien, attends un peu !

Il est évident que le locuteur de (54) ne prend pas en charge le contenu du premier énoncé, et en attribue même la responsabilité à l'allocutaire, peu importe que ce dernier ait ou non affirmé l'imbécillité du locuteur. Sur ce point, Ducrot insistera à plusieurs reprises. Il n'est pas nécessaire que l'énoncé soit attribué à une entité précise, ou même qu'il ait été effectivement produit. L'énonciateur est une entité abstraite (un être discursif qui sature un point de vue, traduiront les polyphonistes scandinaves) :

J'appelle « énonciateurs » ces êtres qui sont censés s'exprimer à travers l'énonciation, sans que pour autant on leur attribue des mots précis ; s'ils « parlent », c'est seulement en ce sens que l'énonciation est vue comme exprimant leur point de vue, leur position, leur attitude, mais non pas, au sens matériel du terme, leurs paroles. (Ducrot, 1984 : 204).

C'est ici que réside à notre avis la différence fondamentale entre la polyphonie et le dialogisme, nettement plus attaché, même s'il ne faut pas le confondre avec le dialogal, à décrire l'interaction que tout discours (ou tout texte) ne manque pas de provoquer avec d'autres discours (ou d'autres textes), même s'il peut s'agir, de son propre autre discours (ou de son propre texte). Même si Brès & Mellet (2009) écartent le concept d'interaction dans le sens pris habituellement au sein de l'approche interactionnelle (alternance de tours de paroles, etc.), l'interprétation qu'ils font du dialogisme les pousse à chercher les énoncés d'autrui dans son propre énoncé, car « les rapports dialogiques (y compris ceux du locuteur avec son propre mot) sont un objet de la translinguistique. [...] Dans la langue, objet de la linguistique, n'existe et ne peut exister aucun rapport dialogique »¹⁴⁰. La perspective dialogique telle qu'elle vient d'être dessinée sera clairement écartée par Ducrot (au même titre que le volet psycho-physiologique) :

¹³⁹ Dont Nølke qui écrit (2009 : 85) : « Si intéressante que soit l'étude de cette relation entre le sujet parlant et le locuteur, elle dépasserait cependant le cadre de la ScaPoLine. »

¹⁴⁰ Bakhtine (1929/1963/1970 : 239), *La poétique de Dostoïevski*, Paris : Seuil (coll. « Pierres Vives »), cité par Brès & Mellet (2009 : 5).

Ce n'est pas ce type de problèmes que je considère comme les miens – ce qui n'implique bien sûr aucune dévalorisation de ces problèmes, mais seulement l'hypothèse que les miens peuvent être traités séparément. (1984 : 178)

On reconnaît là une problématique majeure qui a traversé les différents courants et domaines linguistiques, celle de l'autonomie du linguistique. La conséquence initiale de ces choix théoriques opposés est multiple : tourné vers la parole (vs langue), le dialogisme travaille d'abord sur le discours (vs *énoncé*, conçu comme un élément du discours), la polyphonie fixe comme domaine de pertinence l'énoncé (c'est même là la différence principale qu'évoque Ducrot lui-même pour se distinguer de Bakhtine), mais posera rapidement l'existence, parallèlement à la polyphonie discursive, d'une polyphonie en langue.

Il est remarquable de constater que chacune des deux approches, entrées en dialogue essentiellement dans les années 2000, fera un pas vers l'autre : la polyphonie, surtout dans le cadre de la Scapoline¹⁴¹, revendiquera un programme de recherche qui s'intéresse au discours et au texte, tandis que le dialogisme fera le chemin inverse, et finira par étendre son investigation à l'intérieur de l'énoncé. Le pas ultime de cette évolution est la reconnaissance de l'existence de marqueurs dialogiques *grammaticaux*¹⁴² : quand on travaille sur le conditionnel, la polyphonie vient comme une panacée !

2.3.2. De la polyphonie à la polychronie : un modèle de traitement de la polysémie temporelle ?

Si, lorsqu'on travaille sur le conditionnel, l'approche polyphonique (ou dialogique) s'impose comme une évidence¹⁴³, l'extension de cette approche à l'étude d'autres formes temporelles, si elle n'était pas initialement vraiment envisageable, commence depuis quelques années à faire son chemin, notamment pour l'imparfait (Donaire, Patard) et le futur (Azzopardi et Brès)¹⁴⁴.

Je ne reviendrai pas en détail sur cette hypothèse (elle est esquissée dans la thèse, et dans Abouda 2003 et 2004c) mais en présenterai brièvement une version moins forte qui articule la polychronie aux mécanismes de résolution de conflits (Gosselin 1996 et 2005)¹⁴⁵.

¹⁴¹ Voir également Donaire (2006).

¹⁴² Brès (2009) & (2014), Azzopardi (2011).

¹⁴³ Haillet (1995 & 2002), Abouda (1997a), Donaire (1998), Azzopardi (2011), Brès (2014)...

¹⁴⁴ Il me semble que dans ses derniers travaux, Brès développe une vision très proche de celle que je défends. Il écrit, par exemple (2014 : 22) : « Notre hypothèse est que le conditionnel a en langue une structure dialogique, et que le futur peut avoir en discours un fonctionnement dialogique, ce qui leur permet, de façons différentes, de participer à la production de l'indirection de la source ».

¹⁴⁵ Présentée en 2008 lors de la journée Temporalité organisée à Rouen par L. Gosselin, mais non publiée.

2.3.2.1. *Le problème et un diagnostic*

L'objectif initial de l'esquisse d'un modèle général pour le traitement de la polysémie temporelle a été le déséquilibre, d'abord constaté par Confais (1990) me semble-t-il, entre la généralité d'un phénomène et les réponses particulières apportées par les grammairiens et les linguistes, réponses qui « se font au coup par coup au moyen d'analyses *ad hoc* qui nécessitent un certain talent de la part du linguiste, car elles s'avèrent parfois très acrobatiques » (Confais (1990 : 53), déjà cité ci-dessus) ¹⁴⁶. Réponses dispersées, développées pour des marqueurs particuliers : pour le présent, il n'y aurait pas polysémie parce que cette forme, comme l'indiquerait son zéro morphologique, serait dépourvue de valeur temporelle spécifique... Pour le futur, on convoquera l'incertitude (extralinguistique !) liée à l'avenir pour expliquer des emplois modaux, et pour les autres formes temporelles toutes sortes d'hypothèses dérivationnelles¹⁴⁷ seront proposées...

Ces difficultés proviennent en plus des données elles-mêmes, dont la complexité n'est plus à démontrer, de l'ambiguïté du point référentiel S – de *speech act* – proposé par Reichenbach, et ses corollaires dans la plupart des autres systèmes (t_0 , R_0 , etc.). Le diagnostic a été posé par plusieurs linguistes, dont Wilmet qui l'exprime ainsi (1996 : 201) :

S confond le moi-ici-maintenant – ce repère que tout énonciateur pensant, parlant ou écrivant a une propension naturelle à choisir comme base de la division du temps en époques – et l'actualité authentique, dissociable du précédent, qu'il s'agisse d'une invitation expresse à le délaissier (type "reportons-nous en arrière") ou d'une décision arbitraire du locuteur (type "nous sommes en 2040").

Wilmet affirme ainsi qu'il est nécessaire de reconnaître en plus d'un point zéro du Temps extra-linguistique où se situe la personne réelle qui parle (et sur lequel elle n'a aucune prise), un autre point zéro du temps linguistique où se situe une entité qui ressemble fort au locuteur. Mais cette distinction ne parvient, selon moi, à traiter qu'un type de distorsion temporelle.

¹⁴⁶ Voici, par exemple, le traitement que propose Imbs (1968 : 97) de l'imparfait hypocoristique : « Le recul dans le passé symbolise l'écart qui sépare le monde de la personne adulte du monde de l'enfant auquel elle s'adresse : en pénétrant dans l'univers des enfants, l'adulte suggère, par l'imparfait, à la fois un univers réel comme le passé, et un univers éloigné du présent des adultes. »

¹⁴⁷ « Pour confirmer la thèse de la signification première et des significations secondes, écrit Confais (1990 : 53-54), il est nécessaire d'interpréter les anomalies comme des images, des figures de style ou plus généralement des emplois figurés. »

2.3.2.2. Polychronie

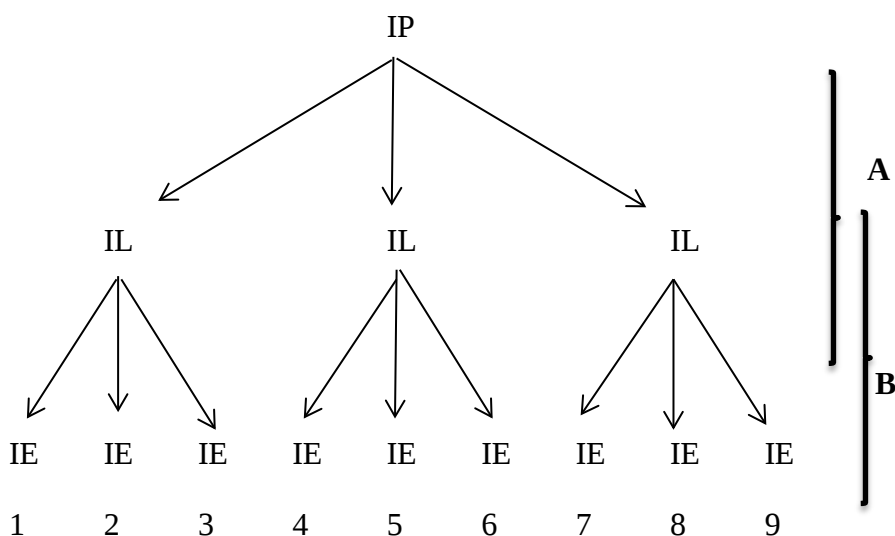
L'idée de base de l'hypothèse polychronique consiste à fournir à chacune des entités identifiées par Ducrot au sein de la notion du sujet parlant un attribut temporel. On aboutit ainsi à une distinction entre trois temporalités (ou trois axes temporels parallèles) :

- I. un Temps extra-linguistique où se situe P, qu'on identifie par le repère référentiel IP, i.e. intervalle référentiel où se place la personne réelle qui parle ou écrit ;
- II. un temps linguistique où se situe le Locuteur, identifié par le repère référentiel IL ;
- III. et un temps linguistique où se situe l'Enonciateur, identifié par l'intervalle référentiel IE.

Or, de même qu'un énoncé quelconque, avant d'être ou non attribué au Locuteur, doit d'abord être attribué à un Enonciateur, être théorique nécessaire à l'existence même de l'énoncé, on proposera par analogie, que le procès avant même d'être calculé par rapport à IL, soit d'abord calculé par rapport à IE. Cela revient à attribuer aux formes temporelles des structures temporelles décrites en termes de rapports avec IE.

D'un point de vue strictement onomasiologique, on aurait *a priori* neuf cas de distorsions possibles (mais il n'est pas sûr qu'ils se réalisent tous), qui sont répertoriés sous 0:

(55)



Le cas (5) de ce schéma constitue la situation 'normale', non-marquée, c'est-à-dire le cas où il n'y a pas de distorsion. Tous les autres cas présentent des distorsions, distinguées en deux groupes : A, et B.

La distorsion de type A, qui s'opère entre IP-IL est une distorsion extralinguistique qui semble être liée à la création d'un univers relativement autonome. C'est ce type de distorsion qui a été pointée par Wilmet cité ci-dessus. Je fais l'hypothèse que ce type de distorsion est coûteux : il ne se rencontre pas dans les énoncés ordinaires, à l'occasion d'un ou deux énoncés, mais semble relié à la création d'un univers narratif particulier, et doit s'appliquer, sinon à une œuvre entière, au moins à une séquence textuelle relativement longue : c'est là, à notre sens, où interviennent les formules explicites de transition, de types *Reportons-nous en arrière*, ou *Nous sommes en l'an 2040*¹⁴⁸ qui ouvrent les séquences narratives en question. Ces formules n'ont rien d'obligatoire : la distorsion de type A peut se donner à titre axiomatique dans un univers donné, et ne jamais être marquée. Si ce type de distorsion n'intéresse pas en particulier le linguiste, son identification lui est nécessaire pour éviter certaines confusions possibles. Dans un passage comme « Nous sommes au 5 mai de l'année 1555. Henri II règne sur la France [...] »¹⁴⁹, il ne me paraît pas judicieux d'analyser le présent du verbe *régner* comme un présent narratif.

La distorsion de type B, en revanche, qui s'opère entre les deux intervalles IL et IE intéresse le linguiste dans la mesure où elle est systématiquement, c'est en tout cas l'hypothèse que je formule, marquée. Elle peut se rencontrer fréquemment dans les énoncés ordinaires (*hier un type m'aborde dans la rue...*) et peut surgir à l'occasion d'un ou deux énoncés.

2.3.2.3. *Articulation avec les mécanismes de résolution des conflits*

Après avoir cru que l'hypothèse polychronique esquissée ci-dessus pouvait permettre le traitement de tous les cas de polysémie temporelle, il m'a semblé il y a quelques années qu'il était plus juste de la réserver au traitement d'un type particulier de distorsions, et qu'elle peut judicieusement s'articuler aux mécanismes de résolutions de conflits tels qu'ils sont conçus par Gosselin (1996).

¹⁴⁸ Voir Vuillaume (1990), qui rappelle cet exemple des *Misérables* : « Retournons en arrière, c'est un droit des narrateurs, et replaçons-nous en l'année 1815 [...] », cité dans Vuillaume, p. 37.

¹⁴⁹ A. Dumas, *Le Page du duc de Savoie*, t7, 7, cité dans Vuillaume (1990 : 83).

Le modèle de Gosselin étant largement connu, on se focalisera sur certains aspects des mécanismes qu'il propose pour résoudre les conflits qui peuvent se présenter entre les différentes représentations linguistiques qui interagissent au sein de l'énoncé.

Pour identifier le domaine où la polychronie pourrait intervenir, je me propose de prendre un exemple simple, celui du présent.

En plus de son emploi typique codé en langue (comme dans *Marie mange un fruit*), le présent de l'indicatif est susceptible d'avoir des effets de sens non typiques (*Cette semaine, Marie est en vacances*), et des effets de sens, qualifiés de dérivés, car obtenus par résolution de conflits.

Gosselin dresse une typologie complète des types de conflits et des mécanismes de leur résolution. Je ne m'attarderai que sur un type de conflit, qu'on peut qualifier de temporel¹⁵⁰, qui émerge soit entre éléments de représentation temporelle qui coexistent au sein de l'énoncé, soit entre représentations aspectuo-temporelles et contraintes pragmatico-référentielles (en l'occurrence conversationnelles). Dans le cas du présent, on peut illustrer ces cas successivement par le présent narratif, et par le présent à valeur de futur :

- (56) (a) Lundi dernier, il arrive comme un fou et me demande un tire-bouchon
(b) (Au cours d'une conversation) J'entre chez lui et je vois une chose extraordinaire
(c) Demain, je rentre à la maison.

Le mécanisme de résolution de conflit qui intervient dans ces cas est celui du déplacement de l'intervalle de l'énonciation qui peut s'opérer soit vers le passé soit vers l'avenir.

Comment peut-on justifier, à l'intérieur de ce modèle, ce phénomène de reduplication de l'intervalle d'énonciation ?

La réponse réside en partie dans le concept de représentation, définie comme une simulation de la présence et de la perception :

Dire que la simulation de la perception se trouve libérée de l'exigence de la présence, c'est dire que le passé comme l'avenir deviennent perceptibles et montrables, et qu'ils se comportent donc comme s'ils étaient également présents [...].

C'est manifestement pour souligner le phénomène, et renforcer ainsi la stratégie de simulation, que les procès passés et futurs sont parfois présentés au moyen de marqueurs

¹⁵⁰ Les déformations (au moins en partie) aspectuelles reçoivent dans ce cadre théorique des traitements parfaitement convaincants.

indiquant normalement le présent [...]. On propose de modéliser ce phénomène en autorisant une opération de duplication de la position du sujet. Cette multiplication constitue un mode de résolution de conflit entre des instructions temporelles normalement incompatibles et correspond à la multiplication (polyphonique) des points de vue temporels dans un même énoncé. (Gosselin, 1996 : 102)

Si tel est le cas, peut-on continuer à définir l'intervalle d'énonciation comme l'intervalle qui indique les limites temporelles de l'acte physique de l'énonciation ?

La notion elle-même du sujet de l'énonciation, qui est donc réduplicable, n'est pas sans ambiguïté. On est là tout proche d'une approche que l'on peut qualifier – le mot a d'ailleurs été utilisé par Gosselin lui-même – de polyphonique.

Sans être la seule possible (des modèles cognitifs proposent actuellement des alternatives intéressantes que Gosselin semble retenir), l'hypothèse polychronique me semble apte à traiter certains types de conflits, ceux précisément qui ne peuvent être résolus que par réduplication de l'intervalle de l'énonciation, réduplication liée, nous l'avons vu, à celle du sujet de l'énonciation, notion très proche du sujet parlant, dont l'unicité a été remise en cause par Ducrot (1984).

Elle permet, sans abandonner les acquis du modèle de Gosselin, de fournir un arrière-plan linguistique à cette notion de réduplication. Elle répond d'autre part à une intuition, largement partagée au-delà de toutes les divergences de traitements, qui fait que le procès est vu de deux points de vue à la fois, de l'intervalle référentiel du locuteur et de celui de l'énonciateur. Avec tout ce que cela implique, non seulement sur le plan chronologique, mais aussi énonciatif.

3. Retour aux données

Une question importante, mais qui, parce je préférais lui consacrer une partie spécifique, était jusqu'ici sous-jacente, a traversé les deux premières parties de cette synthèse, i.e. la question des données. J'ai ainsi voulu figurer, sur le plan de la forme, une évolution importante sur le plan individuel mais qui me semble représentative de l'évolution de tout le champ linguistique. La question des données a en effet été progressivement appelée à jouer un rôle accru dans le travail linguistique. Avec quelque retard par rapport au domaine anglo-saxon, cette tendance, dont la manifestation la plus claire était la constitution de corpus, d'abord écrits ensuite oraux, a fini par s'affirmer en France à partir des années 1990 et surtout à partir des années 2000. Non sans provoquer des réactions d'hostilité, voire de mépris pour une occupation qui serait étrangère au vrai travail du linguiste. En réaction, cette nouvelle approche n'était pas moins critique à l'encontre d'une linguistique qui passait son temps à développer un appareillage extrêmement sophistiqué et complexe pour décrire des données dont l'existence même ne serait pas toujours garantie. Outre l'aspect scientifique, les rivalités de personnes interféraient dans un champ traversé par des enjeux de légitimation et de domination.

Aujourd'hui, avec un recul qui devient chronologiquement significatif, il me semble possible d'établir un premier bilan d'un mouvement collectif, d'après le parcours d'un chercheur qui a vécu intensément cette mutation¹⁵¹ : formé à Paris 7, connu pour être l'un des endroits où l'on pratiquait une linguistique formelle soupçonnée d'ignorer les données, j'ai intégré, suite à une demande de mutation, l'université d'Orléans, où on allait constituer et mettre à disposition le plus vaste corpus oral du français.

Cette position d'intermédiaire, au début extrêmement perturbante, se manifeste également au niveau de l'articulation entre mes activités d'enseignant et de chercheur. Si, ainsi que nous l'avons vu, je m'étais principalement focalisé sur l'étude d'objets sémantiques, c'est à l'enseignement de la syntaxe formelle au sens large que mon service d'enseignement a été essentiellement consacré. Depuis mon recrutement comme Attaché Temporaire d'Enseignement et de Recherche à Paris 7, jusqu'à mon arrivée à Orléans, en passant par les quatre années passées à Grenoble, j'ai en effet systématiquement, le plus souvent exclusivement, enseigné des UE de syntaxe¹⁵². Pour des raisons multiples (profil et projet

¹⁵¹ Des bilans plus généraux existent désormais en nombre en français : voir notamment Habert et *al.* (1997), Williams (dir., 2005), Habert (2005)... Il est à noter que sur de nombreux points, notamment techniques mais également méthodologiques, l'évolution extrêmement rapide du domaine relègue une partie de ces bilans, pourtant relativement récents, au rang de témoignages historiques.

¹⁵² Voir CV, rubrique « enseignement ».

professionnel des étudiants, cohérence du service par rapport aux autres enseignements, ressources disponibles pour les étudiants, etc.), il s'est agi d'une syntaxe formelle intégrant peu de données, surtout peu tirées de l'oral. S'il me paraît indispensable, pour des raisons pédagogiques légitimes, d'enseigner aux étudiants la syntaxe distributionnelle, de les initier à la syntaxe générative et aux grammaires d'unification, une ouverture à des modèles syntaxiques plus tournés vers les données orales¹⁵³ aurait constitué un complément de formation utile et intéressant, une fois que les étudiants auraient acquis des bases solides. Sauf que, à l'image sans doute de la plupart des universités de taille comparable, la structure de l'enseignement des SDL à Orléans et le profil des étudiants, sans évoquer ici le contexte macro-économique et social, empêchent toute accumulation de connaissances au-delà de la licence. En effet, ne s'inscrivent généralement en master que des étudiants qui viennent d'autres filières, avec une formation linguistique quasi inexistante, ou des étudiants de SDL dont le niveau général est si faible qu'ils n'ont pas pu intégrer leurs filières rêvées¹⁵⁴, et à qui il est généralement plus utile d'administrer des piqûres de rappel que des enseignements nouveaux et plus en rapport avec les recherches menées au sein du laboratoire : il s'agit là à mes yeux du défi principal au niveau des masters SDL à Orléans et il est loin de concerner uniquement les enseignements de syntaxe.

En bref, chercheur en sémantique où j'ai été amené progressivement à intégrer les corpus oraux dans mes recherches, ma position d'enseignant de syntaxe, qui par ailleurs est loin de me déplaire, ne m'a pas encore permis d'intégrer suffisamment de données orales ou très peu. Si on pourrait espérer mieux en termes de synergie entre les deux activités, cette position inconfortable aurait, je l'espère, au moins l'avantage d'éviter l'écueil d'un regard caricatural sur le fait linguistique.

¹⁵³ Des travaux dans le domaine français, entrepris dans des cadres théoriques divers, commencent à exister en nombre. Voir dans la bibliographie, entre bien d'autres, les références de Blanche-Benveniste, Deulofeu, Sabio (1995 et 2011), et Skrovec (2014).

¹⁵⁴ Ayant un cours d'introduction à la syntaxe en première année, j'ai pris l'habitude depuis plus de quinze ans de faire remplir aux étudiants des fiches de renseignement personnel qui m'aident à en cerner le profil. D'une année à l'autre, les données varient peu et à la question du projet professionnel, si on écarte ceux qui ne savent pas (généralement ce sont les mêmes qui se sont inscrits en SDL parce qu'ils voulaient éviter les filières classiques qu'ils connaissaient, i.e. lettres modernes, histoire, géo, anglais, etc.), une bonne moitié se destine au professorat des écoles, un tiers voudrait intégrer une école d'orthophonie, et la partie restante voudrait passer un concours de la fonction publique (animation sociale, petite enfance, etc.).

3.1. La linguistique par l'exemple

3.1.1. *Pierre, Marie et la pomme*

Dans mes cours de syntaxe, j'utilise massivement des phrases fabriquées, ou quelquefois glanées mais simplifiées, pour illustrer certains phénomènes ou certains concepts, comme dans le couple suivant d'exemples, où trois éléments jouent des rôles descriptifs et explicatifs importants, *Pierre, Marie et la pomme* :

(57) Pierre rencontre Marie

(58) Pierre mange une pomme

Or, on le sait, le recours à des exemples fabriqués fait l'objet depuis quelque temps de vives critiques : la linguistique, et tout particulièrement la syntaxe, travaille sur des données artificielles, taillées sur mesure pour déduire des principes généraux sur le fonctionnement du langage :

Syntacticians focus on the data that they believe the fewest complications factors, and "clean up" the data to remove what they believe to be remaining complications that obscure simple, general principles of language. (Abney, 1996 : 11).

Pire encore, à force de fabriquer ses propres données épurées, cette « linguistique de bureau » finit par ne travailler que sur des matériaux marginaux, non-représentatifs, voire inexistantes. C'est d'abord la syntaxe générative qui fera l'objet de cette critique :

To go out on a limb for a moment, let me state my view: generative grammar has produced many explanatory hypotheses of considerable depth, but is increasingly failing because its hypotheses are disconnected from verifiable linguistic data. Issues of frequency of usage are by design made external to matters of syntax, and as a result categorical judgements are overused where not appropriate, while a lack of concern for observational adequacy has meant that successive versions have tended to treat a shrinking subset of data increasingly removed from real usage. (Manning, 2003 : 295-296).

Mais elle atteindra rapidement la sémantique formelle, qui, selon Wilks (2011) par exemple, s'intéresserait à des faux problèmes, comme celui des *donkey sentences*¹⁵⁵ qui, si elles font l'objet d'un nombre important d'études,

n'existent pas, à part bien sûr dans les corpus de linguistique. Elles n'existent pas, pas plus qu'on ne dit « John veut épouser une Norvégienne » pour signifier une Norvégienne particulière – on ne parle tout simplement pas comme cela ! Donc ces phrases ne sont pas ambiguës, contrairement à ce que dit la théorie. [...]Le problème n'est pas lié à ceux qui s'intéressent à un nombre limité de données, mais il est lié à toute une tradition qui fabrique

¹⁵⁵ Pour une définition, on peut consulter http://www.semantique-gdr.net/dico/index.php/Donkey_sentence.

et discute d'exemples qui ne correspondent en aucune façon à ce qui pourrait se dire dans la vie réelle¹⁵⁶.

Et de fait, si on cherche sur *Google* « Pierre mange une pomme », on se rendra aisément compte que je n'étais pas, loin de là, le seul linguiste ou grammairien à l'avoir « emprunté ». Aussi longtemps que notre patience tiendra à feuilleter les pages des résultats, l'exemple en question apparaîtra en mention, comme exemple d'une leçon de grammaire. Le seul cas où la phrase apparaît en usage c'est sous la forme d'un titre d'une petite vidéo débile sur YouTube, comme on peut en voir tant sur la Toile, où Pierre ne mange pas la pomme, mais la casse contre sa tête... Notre exemple semble emblématique de la critique de Wilks : il n'existe que dans les corpus des linguistes¹⁵⁷.

Il est pourtant frappant que les étudiants, qui pouvaient légitimement être interloqués par l'artificialité de tels exemples, n'aient jamais manifesté la moindre réticence quant à leur réalité ou la probabilité de leurs emplois, tout au plus manifestaient-ils sporadiquement un agacement vis-à-vis de la pauvreté de ces exemples et... quelquefois, du choix des prénoms. Je m'amusais alors à leur répondre mystérieusement que quand ils seraient en 3^e année de licence ils auraient le droit à Max et Léa, qui font plus de choses que Pierre et Marie... Le rôle de l'exemple, ainsi que le montre le choix plus ou moins conscient des prénoms qui change en fonction du modèle syntaxique, n'est pas de montrer des données linguistiques « brutes », mais des données construites en fonction d'un regard. Et il serait naïf de croire qu'elles le sont dans les corpus oraux : dès qu'elles deviennent un observable, et déjà pour entrer dans le domaine de l'observable, les données sont construites dès les phases qui paraissent *a priori* les plus neutres, comme lors de la transcription, qui constitue toujours un choix, et sans doute même avant cette phase.

Répetons-le : la question de savoir si les exemples sont inventés ou non n'est pas la véritable question. De toute manière, les exemples sont *construits*. Milner (1989 : 113)¹⁵⁸

L'exemple se présente ainsi comme un fait linguistique construit – et il est vrai que cette construction ne doit pas être bâclée¹⁵⁹ – qui s'assume dans une mise en scène différentielle (Marandin, 2001 : 8). L'exemple, ainsi que l'écrit Auroux (1998 : 185), « n'est pas un usage

¹⁵⁶ Wilks (2011) cité et traduit dans Poibeu (2014 : 11).

¹⁵⁷ Cette question avait déjà été longuement discutée par Auroux (1998), et avant lui par Milner (1989).

¹⁵⁸ Attesté ou fabriqué, l'exemple est en effet selon Milner (op. cit. : 113-114) toujours doublement construit. D'une part, par sa qualité de concrétion, l'exemple constitue une sélection qui permet d'exhiber la propriété voulue à l'exclusion des autres, et, d'autre part, il est toujours choisi après une analyse qui permet de l'intégrer ou non à la famille d'emplois considérée.

¹⁵⁹ Les risques d'une telle construction sont connus. Voir Milner (1989 : 113, n. 23) et surtout Auroux (1998 : 192 sq.).

de la langue, mais une mention sui-référentielle d'un de ses éléments : c'est une occurrence qui représente son type. » Il est dans ce cas légitime de le choisir « pour illustrer au mieux un phénomène, et, du même coup, une règle. Par conséquent, il doit être simple et, si possible, ne doit illustrer qu'un phénomène à la fois. Dans ces conditions, il y a interchangeabilité entre l'exemple et la règle » (Auroux 1998 : 191). Ces exigences sont toutefois contrariées par ce que Milner appelle la *concrétion* du langage naturel, qui fait que le moindre élément de la langue combine toujours plusieurs dimensions. Le caractère artificiel constitue ainsi une réponse à la concrétion, et vise une simplicité qui permet « de ne pas encombrer la perception linguistique de propriétés étrangères à la propriété soumise à examen » (Milner, 1989 : 113, n. 23).

En bref, l'exemple se présente comme une coupe, censée attirer d'une manière efficace le regard vers le phénomène expliqué. S'il n'existe pas en tant que fait brut, il est censé exister plus fondamentalement comme un représentant crédible de ce qui est répétable, et qui l'arrache à la singularité de son énonciation. Celle-ci, en lui donnant une réalité, risque en effet par là même de lui ôter sa fonction, par un encombrement perceptif qui le fait dévier vers le fait divers, là où on cherche à capter du général :

Cela va de soi : l'énonciation est ce qu'il y a, dans la profération d'un énoncé, qui ne se répétera pas. Or, la linguistique (comme la grammaire) s'intéresse à ce qui se répète de profération en profération. Comme la grammaire, elle suppose donc qu'il y a, dans le langage, du répétable et l'exemple est ce qui saisit justement cela. Que du même coup, elle néglige des propriétés réelles, on n'en saurait douter, mais elle ne peut faire autrement que de les négliger. » (Milner, *op. cit.* : 110).

Les exemples (57) et (58) sont utiles pour illustrer respectivement le caractère obligatoire du complément du verbe, et la possibilité d'absence d'un complément sous-catégorisé. Même pour décrire des phénomènes aussi basiques, on ne peut pas permuter ces exemples alors qu'ils sont, d'un point de vue strictement syntaxique, identiques. Il ne s'agit pas de cacher pour toujours aux étudiants le phénomène de ce que la tradition grammaticale appelle l'emploi absolu des verbes, mais d'abord de leur présenter une propriété censée être plus générale, à l'aune de laquelle d'autres phénomènes peuvent être présentés. Le chemin des étudiants est ainsi balisé, débarrassé non seulement des données authentiques et de leur charge sémantique qui pourrait les empêcher de voir dans leur pureté théorique les phénomènes expliqués, mais aussi des exemples fabriqués improvisés.

Cette construction des données, si elle semble pédagogiquement et méthodologiquement légitime en syntaxe, garde-t-elle sa pertinence en sémantique ?

Une première précision s'impose : on ne peut pas convoquer ici l'absurdité de certains exemples fabriqués par certains auteurs pour interdire à la sémantique de recourir à la construction d'exemples. A peine caricatural, l'exemple fabriqué par Berthonneau & Kleiber – *Jean tombe [beaucoup] parce que Max le pousse* – pour dénoncer le phénomène n'est malheureusement pas rare. Or, des exemples aussi absurdes se rencontrent en masse dans des ouvrages de syntaxe, quelquefois extrêmement sérieux et publiés dans des universités prestigieuses... La désinvolture avec laquelle on fabrique quelquefois des exemples dans des langues qu'on ne maîtrise pas, sans forcément prendre la peine de vérifier les données auprès de locuteurs natifs, a effectivement de quoi agacer quand on voit quel appareillage théorique sophistiqué est mis en place pour généraliser des fonctionnements présentés comme universels. Cela ne peut que nourrir un certain scepticisme vis-à-vis des exemples et des analyses sur des langues peu connues (et qui n'ont pas de défenseur dans le champ) et par contrecoup à l'encontre du déploiement théorique dans son entier¹⁶⁰.

En bref, le mauvais exemple, aussi bien en syntaxe qu'en sémantique, est un problème qui risque de jeter le discrédit sur toute démarche théorique. On peut donc suivre Auroux (1998 : 193) pour affirmer la nécessité « de protocoles dans la construction des exemples (ce qu'on peut appeler une théorie des observables) et d'une déontologie dans le choix de l'univers de la langue. »

Peut-on en sémantique, à l'image de ce qu'on semble admettre en syntaxe, recourir à la fabrication d'exemples ?

Il est indéniable que si une partie de l'interprétation – ou du sens de l'énoncé selon Ducrot – se détermine dans l'énonciation, qui, par définition, n'est pas répétable ainsi que l'a affirmé Milner, les recherches en sémantique et en pragmatique ont montré qu'il est possible de saisir non seulement la signification de la phrase, qui elle est répétable, mais aussi des effets de sens typiques qui naissent de l'interaction entre marqueurs et contextes¹⁶¹. Il n'y a aucune raison méthodologique qui interdise à la sémantique, dans une démarche hypothético-déductive, de parvenir à dégager des généralités en construisant des données : si des énoncés comme *Ce*

¹⁶⁰ Je pense qu'entre la tradition anglo-saxonne, notamment américaine, et la tradition européenne, notamment française, il y a sur cette question (comme sur d'autres : la politique de publication, le choix des modèles théoriques, etc.) une ligne de démarcation claire qui a peut-être eu tendance à se déplacer ces dernières années.

¹⁶¹ « On peut dire que les instructions posent des contraintes sur le potentiel interprétatif. Qui plus est, elles indiquent une interprétation par défaut. Cela veut dire que les instructions indiquent un contexte par défaut qu'on pourrait caractériser comme « construit » par la forme linguistique. » (Nølke (2009 : 83). Voir également Haillet (2007) sur le concept d'*interprétation par défaut*. La démarche ne me semble pourtant pas sans risque : nous avons vu que certains auteurs attribuent à l'imparfait la valeur [passé] parce que, hors contexte, un énoncé comme *Il pleuvait* situerait l'action dans le passé. Comment valider une telle hypothèse ?

mur n'est pas blanc ou *Luc marche pendant deux heures* ou encore *Max a vécu en banlieue* souffrent des mêmes défauts que ceux formulés à l'encontre des exemples (57) et (58) – on peut le vérifier sur *Google* –, ils ont été conçus pour rendre patent un phénomène sémantique précis...

Il semble que, aussi bien en syntaxe qu'en sémantique, toutes les données, y compris les exemples fabriqués, peuvent avoir leur légitimité si dans leur construction on a pris soin d'expliciter la démarche et de l'interroger. Mais toutes les données ne se valent pas : elles ne peuvent pas être utilisées indifféremment. Cela dépend crucialement de la nature de l'objet, mais aussi du point de vue et des objectifs qu'on s'est fixé. Globalement, il me semble que l'exemple en sémantique doit être moins pauvre et qu'on est plus souvent amené à exemplifier le contexte ou à le contraindre au niveau de l'énoncé afin qu'il soit plus explicite concernant le phénomène observé. Le travail de construction qui en résulte est plus complexe, et donc plus aventureux, ce qui encourage de nombreux sémanticiens à privilégier les données attestées. Or, l'analyse d'une donnée, qu'elle soit attestée ou fabriquée, requiert souvent des interventions de la part du linguiste, amené à manipuler les exemples examinés (commutation d'un élément de l'énoncé avec d'autres éléments, adjonction ou suppression de constituants, identification des enchaînements discursifs que les énoncés autorisent ou non, etc.). J'y reviendrai.

Cela laisse naturellement ouverte la question des données authentiques et leur rôle dans la vérification en conditions réelles des hypothèses avancées.

3.1.2. De Stendhal à Martin

Si, d'un point de vue historique, les grammaires à visée scolaire ont globalement privilégié les exemples inventés, les grammaires à visée savante se sont appuyées sur des « citations » ou exemples attestés. A l'intérieur de ce dernier courant, qu'on trouvera également dans le champ linguistique, on peut encore distinguer entre les travaux qui s'appuient sur des exemples oraux, généralement « ordinaires », et ceux qui font appel à des exemples écrits. D'abord exclusivement littéraires, ces derniers se sont diversifiés pour intégrer d'autres « genres », notamment, pour des raisons pratiques évidentes, les textes journalistiques. Ce découpage, qui n'est pas historique¹⁶², est schématique et ne reflète qu'une tendance. Sans faire mention des travaux qui combinent différents types de données, sur lesquels on

¹⁶² Les études linguistiques actuelles qui intègrent, au moins en partie, des données littéraires sont légion alors que des auteurs comme Damourette et Pichon ont massivement intégré, il y a un siècle, des exemples oraux.

reviendra, on ne peut établir d'équivalence entre oral¹⁶³ et *ordinaire* : parce qu'elle est plus facilement accessible, c'est la parole publique (discours politiques, émissions de radio et de télévision...), généralement formatée, qui a d'abord été étudiée. Il en va de même entre écrit et *standard* : des recherches sont consacrées aux copies d'élèves, aux sms, aux échanges sur les forums, etc.

Le point sur lequel on voudrait insister concerne la présence d'une tendance, souvent implicite, à fonctionner selon le mode de « la preuve par l'exemple » : elle consiste à voir dans l'exemple attesté une protection qui dispense de toute interrogation méthodologique quant à sa validité, son degré de représentativité et le rôle qu'on lui attribue à l'intérieur du modèle. Paradoxalement, alors que la démarche qui fabrique ses propres données, en reconnaissant leur aspect construit, est conduite à une certaine prudence et à une évaluation de sa production, la démarche inverse, ignorant généralement ce qu'il y a de construit dans ses exemples attestés, court le risque de commettre certaines erreurs méthodologiques¹⁶⁴.

Or, il n'est pas sûr, et cela est particulièrement visible en grammaire, que, sur ce point, la méthodologie qui repose sur les citations soit radicalement différente de celle qu'elle est censée critiquer : lorsque pour illustrer une règle ou un fonctionnement linguistique quelconque on fait appel à un exemple attesté, celui-ci n'a de valeur, sur le point précis considéré, que par et dans l'exclusion d'autres réalisations potentielles. Cela revient à dire que le différentiel entre possible et impossible de langue (noté par l'astérisque) ne disparaît pas avec la disparition apparente de l'un des termes opposés. En soi, la démarche est concevable si elle s'accompagne d'un certain nombre de précautions méthodologiques, au premier rang desquelles figure la représentativité¹⁶⁵ qui seule permet de lui attribuer une pondération et donc de déterminer l'importance de son rôle pour le modèle. Or la représentativité ne peut être mesurée qu'en corpus, entendu comme une « collection de données langagières sélectionnées et organisées selon des critères linguistiques et extra-linguistiques explicites pour servir d'échantillon d'emplois déterminés d'une langue » (Habert, 2005 : 148). Voilà pourquoi je considère, suivant en cela Habert, que paradoxalement l'utilisation, désormais particulièrement facile, de ressources électroniques pourrait constituer un danger et même une

¹⁶³ Voir sur ce point Béguelin (1998), Cappeau & Gadet (2007a).

¹⁶⁴ C'est ce que note Milner (op. cit. : 114) : « Il est vrai que ce double caractère de construction active demeure implicite et par là facilement inaperçu dans la tradition grammaticale et dans la plupart des théories linguistiques. Dans la théorie de l'école de Cambridge [...], le caractère construit apparaît à l'évidence, puisqu'on y use d'exemples nécessairement fabriqués. »

¹⁶⁵ Ainsi qu'on le verra ci-dessous, la question de la représentativité est extrêmement complexe. Voir sur ce point Habert (2000). Ici, on ne parle pas forcément d'une représentativité générale au niveau de toute une langue, mais il peut s'agir d'une représentativité plus restreinte à un genre, registre, époque, type d'interactions, etc.

régression par rapport à une époque pas si ancienne où les données, parce qu'elles étaient difficiles à rassembler, étaient mieux conçues¹⁶⁶ :

Le risque de ces données « à portée de main » ou de la souris est celui de la confusion dans la profusion. Pour ce qui regarde les archives électroniques de journaux nationaux comme *Le Monde*, elles ne représentent qu'un certain emploi du français (écrit tenu). Malgré sa taille, [...], *le web n'est représentatif de rien si ce n'est de lui-même*. [...] On trouve certes tout sur le web, mais sans savoir exactement quoi. On risque alors d'obtenir des amas textuels encore plus mal définis que les archives de journaux. C'est un peu comme si, en mélangeant des tranches aléatoires et de taille variable provenant de radios aussi différentes que France Culture, Radio Notre Dame, Skyrock, NRJ, etc., on pensait parvenir à représenter le français oral. (Habert, 2005 : 136)

Si ces remarques sont valides au niveau de la constitution de corpus, où Habert pointe entre autres le panachage aléatoire et non réfléchi de différents types de données, elles le sont plus encore en ce qui concerne la pratique, très courante, qui consiste à glaner au vol des attestations détachées de leurs locuteurs et de leurs contextes de production. Ce papillonnage, auquel, à l'image de beaucoup d'autres, je me suis quelquefois livré tant il est commode, est méthodologiquement indéfendable : non seulement il met sur un pied d'égalité des données hétérogènes sur les plans qualitatif et quantitatif, mais il ne permet ni de mesurer le poids de ce qui est exploité ni de conjecturer ce que la recherche n'a pu capter. Particulièrement visible au niveau des recherches informatiques (utilisation des moteurs de recherche ou de ressources électroniques comme Frantext...), cette façon de procéder est présente dans des méthodes de recherche plus « manuelles ». Berthonneau & Kleiber (2006), par exemple, tout en critiquant, ainsi que nous l'avons vu ci-dessus¹⁶⁷, la désinvolture avec laquelle des auteurs fabriquaient des occurrences d'énoncés, n'ont cherché ni à présenter, ni à justifier leurs propres données¹⁶⁸. Ils livrent une série de dix-huit exemples, certains fabriqués, d'autres attestés (Stendhal, Hugo), d'autres encore empruntés à des linguistes qui les avaient eux-mêmes fabriqués (Wagner et Pinchon, Martin). Que ces exemples soient donnés à titre indicatif pour fixer les idées est une pratique courante et compréhensible, en regard du travail fastidieux que la constitution d'un corpus implique, mais tel ne semble pas être le cas de cet article où un rôle important est accordé aux phrases citées. Ainsi, l'exemple suivant

(59) Deux kilos de moins, je rentrais dans ma robe

¹⁶⁶ Pour une présentation historique située des différents types de ressources linguistiques, cf. Habert *et al.* (1997), Habert (2000) et Habert (2005).

¹⁶⁷ Voir notamment n. 102, p. 47

¹⁶⁸ La pratique est non seulement courante mais elle semble même naturelle dans une bonne partie du champ linguistique. A tel point qu'un relecteur de notre article soumis à *RSP* (Abouda & Skrovec, 2015, à paraître) a trouvé que la présentation des données et de la démarche méthodologique était trop détaillée...

qui me paraît discutable, joue un rôle important dans la démonstration des auteurs. On peut légitimement s'interroger sur la relation qui peut être établie entre des exemples aussi hétérogènes, en dehors des types d'emploi que les linguistes ont décidé de leur attribuer (en l'occurrence l'imparfait dit *contrefactuel*). On ignore s'il s'agit des seuls exemples disponibles, si certains ont été écartés et pourquoi, etc. Ces questions importent d'autant plus que les linguistes diffèrent sur les classes d'emploi à distinguer et l'extension de chacune d'elles.

En bref, le papillonnage ne permet pas de mener à bien une procédure de vérification des données qui constitue l'une des tâches incontournables de toute activité scientifique.

Un autre point à discuter concerne le poids qu'il faudrait accorder à une attestation dans le modèle descriptif et explicatif. Nous avons vu ci-dessus (p. 68) que Patard (2007) s'appuie sur un seul exemple tiré d'*Une jeunesse* de Modiano pour admettre que l'imparfait puisse ne pas avoir d'antécédent. Au-delà de la critique concernant la qualité d'une source littéraire qui peut ne pas être considérée comme représentative, peut-on s'appuyer sur une seule attestation (le corpus de Patard contient 4403 occurrences) pour prendre une décision aussi tranchée¹⁶⁹ ? Sur un point de désaccord entre les différents modèles, l'existence d'une unique attestation pose problème. On sait désormais qu'il y a dans les corpus des « existants impossibles » (ou des énoncés que l'on juge *a priori* impossibles). Habert (2000) les répartit en quatre catégories : (i) le lapsus ou l'erreur ; (ii) le viol intentionné d'une règle (qui est à la source de recherches stylistiques, de jeux de langage, etc.) ; (iii) la variation interne à la langue (Labov) ; (iv) le témoin de l'évolution des règles (qui n'est pas sans lien avec la troisième catégorie). Il n'est pas exclu que l'exemple de Modiano fasse partie de la deuxième catégorie.

Ecore une fois, la ligne de partage la plus pertinente méthodologiquement ne serait pas celle qui oppose les données selon qu'elles sont fabriquées ou attestées, mais la pertinence du fait linguistique mobilisé, par rapport aux données écartées, et la place qu'on lui accorde, en fonction de sa pertinence, dans le dispositif descriptif et explicatif du modèle. Tout cela revient à affirmer la nécessité d'un statut des données, ce qui ne peut s'envisager en dehors de la constitution de corpus. Or, on le verra par après, les données ne sont réellement explicites que si elles sont accessibles (dans le cas d'exemples oraux, l'accessibilité implique entre

¹⁶⁹ Nous verrons ci-dessous (§ 3.2.1.2, p. 97) que ce type de raisonnement présente deux types d'erreurs statistiques, l'incertitude et la déformation.

autres l'accès au son) et situées, c'est-à-dire reliées aux métadonnées qui fournissent les informations sur la situation de communication, le profil du locuteur, etc.

3.2. La linguistique en corpus

La tradition anglo-saxonne de linguistique descriptive, la première à s'appuyer sur des corpus électroniques¹⁷⁰, a su exister et se maintenir, malgré l'hégémonie du paradigme générativiste jusqu'à l'arrivée d'un renfort venu du domaine du Traitement Automatique du Langage. Malgré leur extrême sophistication, les formalismes utilisés manquaient d'efficacité et l'augmentation continue de la puissance de calcul permettait d'intégrer une quantité croissante de ressources (dictionnaires, grammaires) et d'en rendre compte de façon plus satisfaisante, à travers des règles de plus en plus détaillées et proches des données.

Or, ainsi que le note Habert, entre les « corpistes » talistes et les ethnométhodologues, spécialistes d'analyse de discours, sociolinguistes, linguistes de terrain, etc., la convergence n'est qu'apparente, tant sont nombreuses leurs divergences théoriques et pratiques, concernant non seulement l'objectif de la démarche, la méthode utilisée, mais également la nature elle-même du corpus et ses caractéristiques (en taille, en structure, en format). C'est ce qui a poussé Habert et *al.* (1997) à parler *des* linguistiques de corpus. La démocratisation de l'accès à l'informatique aboutira à ce que Habert (2000) appelle *laïcisation* : avec un ordinateur et un environnement logiciel désormais à portée de main, de nombreux linguistes, sans l'aide d'informaticiens¹⁷¹, sont à même de construire leur propre corpus. On aboutira ainsi à une situation où règnent des pratiques très divergentes et plus ou moins abouties...

3.2.1. Les ESLOs

Le projet des Enquêtes Sociolinguistiques à Orléans (désormais ESLO) fut arrêté en 1966, lorsqu'un groupe de linguistes britanniques, cherchant à renouveler les pratiques pédagogiques de l'enseignement du FLE en Angleterre, décida, en conformité avec une pratique didactique solidement ancrée dans le pays, d'intégrer l'enseignement de l'oral. Devant le constat de l'inexistence pour le français, y compris en France, de matériaux pédagogiques cohérents et actualisés, une décision fut prise de constituer un nouveau corpus. L'entreprise, qui durera plus de cinq ans entre la conception et la réalisation, donnera

¹⁷⁰ Voir Habert (2000) et Habert et *al.* (1997).

¹⁷¹ Nécessitant systématiquement l'appui d'informaticiens (parce que par exemple il est souvent nécessaire de procéder à une adaptation logicielle), le travail sur un corpus numérique, qui requiert un haut degré de spécialisation, sans l'aide d'informaticiens n'est souvent pas une situation choisie, mais subie par le linguiste par manque de moyens.

naissance à l'un des corpus les plus vastes de français oral : 350 bandes magnétiques représentant quelques 317 heures d'enregistrements, ce qui correspond à quelques $\pm 4\,500\,000$ mots. Outre sa taille, ce corpus se distingue, par exemple de son aîné, i.e. le Français Fondamental (FF)¹⁷², par sa volonté de capter au mieux la variété des usages. Dans le champ balbutiant de la sociolinguistique, il a tenté de représenter les productions discursives les plus diverses (en plus des entretiens, qui constituaient la partie essentielle du corpus, ESLO contenait des productions plus spontanées : reprises de contacts informelles, repas de famille ou entre amis, enregistrements en micro caché dans les commerces et sur les marchés...), chez des locuteurs aux profils variés en termes d'âge, de sexe et de catégories socioprofessionnelles¹⁷³.

Au début des années 1990, l'intégralité du corpus (constitué des bandes magnétiques originales ainsi que d'un ensemble de documents papiers, dont un catalogue dactylographié détaillant des métadonnées relatives aux enregistrements, les fiches d'identification des locuteurs, quelques centaines de feuillets de transcription manuscrites) fut confié au LLL (ex-CORAL), qui entreprit au milieu des années 2000 de le numériser pour sauver d'une détérioration inéluctable un témoignage unique des années 1960-1970. Cette numérisation, qui s'apparentait à une véritable reconstruction, fédéra une dizaine de chercheurs du LLL, et nous a poussé à entreprendre, à quarante ans de distance, la constitution d'une nouvelle enquête sociolinguistique à Orléans, ESLO2, qui se voulait à la fois comparable à la première enquête (désormais ESLO1), et en même temps adaptée à la situation contemporaine à la fois sur le plan technique et scientifique¹⁷⁴.

Les enregistrements débiteront en 2008, et l'ensemble ESLO1-ESLO2 finira par constituer l'un des plus vastes corpus oraux disponibles du français : il contient actuellement environ 7 millions de mots, et l'alimentation du site¹⁷⁵ continue.

Si ma participation à ce projet collectif a constitué sur le plan individuel un tournant épistémologique et méthodologique majeur, tous ses effets n'ont pas été favorables à mes propres recherches. En cela, je ne peux que souscrire à ce qu'écrivait déjà en 1992 Bergounioux qui, expliquant les raisons qui favorisent dans le marché de la linguistique la recherche abstraite au détriment des pratiques de recension descriptives, écrit (1992 : 18) :

¹⁷² Voir sur ce point Bergounioux & *al.* (1992) et Abouda & Baude (2009).

¹⁷³ Pour une présentation historique détaillée d'ESLO, voir Bergounioux & *al.* (1992).

¹⁷⁴ Voir Abouda & Baude (2007) pour une présentation de ce double projet.

¹⁷⁵ <http://eslo.huma-num.fr/>

Il convient d'ajouter les temps morts, c'est-à-dire invisibles en termes de résultats, de la constitution de corpus, pour les déplacements, la sélection des opérations et des témoins, les contraintes techniques, la saisie, les transcriptions, le catalogage. Dans toute enquête longue, il y a le risque que le linguiste n'en vienne à faire de son moyen une fin jusqu'à parvenir, à force de précautions, à une forme de présentation d'autant plus achevée qu'elle aura renoncé à l'explication.

Personnellement, j'ai fait partie, dès ses débuts (les premières réunions ont commencé en 2004), de ce projet collectif qui m'a beaucoup appris, notamment une initiation à la sociolinguistique, et sur le plan technique, de la conception du questionnaire à la réalisation des entretiens (j'ai personnellement assuré l'animation d'une quinzaine d'entretiens), de la transcription (sur Transcriber) au catalogage, en passant par la gestion technique (réalisation de tableurs pour assurer le suivi des enquêtes, vérification des métadonnées du mini-corpus test de l'omelette...). Particulièrement chronophage¹⁷⁶, ce projet explique en partie¹⁷⁷ mon éloignement du champ de la temporalité comme le trahit le niveau des publications de 2004 à 2012.

Pour être complet, j'ajouterais que le corpus ESLO n'a été réellement disponible, y compris pour les chercheurs du LLL, qu'à la fin 2013 et qu'aucune exploitation exhaustive n'était réellement envisageable auparavant.

3.2.1.1. *De l'écrit à l'oral*

L'appareil Recordon que nous avons choisi présentait pour notre travail des avantages certains. Son poids était peu élevé : un peu plus de 6 kg. [...]. L'interruption dans les conversations que nous imposait le changement des disques toutes les six minutes, au terme de leur durée maxima d'enregistrement, ne présentait pas non plus d'inconvénient. Nous ne nous soucions pas non plus de la conservation des disques. Nous profitions largement des possibilités qu'offrent les disques en papier d'être effacés et de servir ainsi à plusieurs enregistrements successifs. Il aurait été beaucoup trop coûteux de conserver tous les enregistrements comme beaucoup de bons esprits nous le suggéraient.

C'est ainsi que l'on justifiait dans Gougenheim et *al.* (1964 : 63) la décision, que l'on peut trouver aujourd'hui étonnante, d'effacer, au fur et à mesure de leur transcription, les

¹⁷⁶ Aujourd'hui, alors que semble s'ouvrir au niveau institutionnel le chantier d'une reconnaissance du travail de constitution de corpus, il convient de rappeler que cette reconnaissance ne devrait pas se restreindre au premier cercle mais devrait s'étendre à tous les participants à ce type de projets, tant il est vrai que la plupart des « petites mains », sans lesquels un projet collectif d'ampleur ne pourrait voir le jour, travaillent dans l'ombre et ne sont souvent même pas identifiés comme participants à ce type de projets.

¹⁷⁷ Il est extrêmement difficile de mesurer avec exactitude la part du projet ESLO dans mon éloignement du champ de la temporalité : à la même période, les responsabilités administratives (directeur des études pendant un an, puis, pendant 3 ans, directeur d'un département de plus de 400 étudiants de la 1^{ère} année au doctorat) et un service d'enseignement extrêmement chargé (à cause d'un sous-encadrement chronique, j'ai dû, à l'image de la plupart des collègues, assurer une centaine d'heures complémentaires en moyenne depuis mon arrivée à Orléans en 2003 jusqu'en 2012), associés à une situation familiale qui me réclamait de l'énergie et de la présence ont fini par peser lourd et réduire le temps disponible que je pouvais consacrer à mes recherches dans le domaine de la temporalité.

enregistrements audio réalisés dans le cadre du projet du Français Fondamental (FF) initié au début des années 1950. Mais outre les raisons invoquées, il ne serait sans doute pas absurde de voir également dans ce gâchis scientifique et patrimonial un rapport problématique à la matière sonore¹⁷⁸. En effet, si, ainsi que le rappelle Blanche-Benveniste (1997), le magnétophone a été inventé dans les années 1930, les études portant sur du français oral n'ont débuté que bien plus tard. « Pendant toute une époque, écrit-elle (1997 : 1), le magnétophone n'a servi qu'à étudier des langues dépourvues d'écriture, langues « exotiques », dialectes ou patois. Il n'en était pas question pour les langues dites de "grande culture", comme le français ».

A ce fait historique, on ajoutera un second, non moins significatif : les deux premiers corpus oraux du français, le FF et l'ESLO, ont été conçus dans une perspective didactique, et plus particulièrement dans le domaine du FLE, qui n'a pas la réputation d'être l'un des plus ambitieux scientifiquement. Le FF, conçu comme une réplique du *Basic English*, a été créé pour promouvoir le français dans des colonies qui allaient bientôt devenir indépendantes, tandis que l'ESLO a été initiée par des universitaires britanniques pour renouveler la didactique du FLE en Angleterre. Ce dernier corpus, s'il a été exploité par des universitaires anglais, allemands, belges et néerlandais, n'avait connu d'autre utilisation en France qu'une application pédagogique du BELC¹⁷⁹. C'est ce que rappellent Blanche-Benveniste & Jeanjean (1986 : 85) : « Et en France ? Peu de choses, si ce n'est l'exploitation pédagogique qu'en a tirée le BELC... Le grand corpus de français parlé, recueilli en France, a été acheté et exploité uniquement hors de France. Curieuse aliénation du milieu linguistique français ! »

Les raisons de cette situation singulière sont multiples. Outre les difficultés techniques et financières (la constitution d'un corpus est chronophage et coûteuse), c'est à des raisons historiques et épistémologiques qu'on peut imputer une telle résistance face aux corpus oraux. Pendant longtemps, le français parlé était loin de constituer un domaine légitime dans le champ de la linguistique. L'histoire même du français fournit une première explication : en tant que langue nationale qui a dû s'imposer à l'encontre de toutes les langues régionales, le français avait besoin d'une légitimation symbolique qu'il a su trouver dans l'écrit surtout littéraire. Ensuite, ainsi que l'affirme Bergounioux (1992)¹⁸⁰, la linguistique en tant que

¹⁷⁸ Voir Bergounioux (2004).

¹⁷⁹ Voir également Bergounioux et al. (1992).

¹⁸⁰ « Dans la division entre gens de terrain et gens du livre, la logique universitaire ne peut que favoriser ceux qui sont les plus respectueux de sa tradition lettrée. Cette tradition philologique, qui favorise la langue morte - hier le latin et le grec, aujourd'hui la langue littéraire - contre la langue vivante, ne se serait pas perpétuée si facilement

discipline universitaire a incorporé, au sein même de son organisation, la domination symbolique de la forme écrite de la langue en faisant la part belle à la grammaire normative. Enfin, ainsi que le rappellent, après Blanche-Benveniste & Jeanjean (1986), Abouda & Baude (2009 : 138), « la linguistique s'est construite sur une lecture des dichotomies proposées par Ferdinand de Saussure qui a situé, hors de la discipline, la description des variations en général et des formes orales en particulier. L'étude du français parlé s'est alors trouvée exclue du champ d'une science vouée à la recherche d'invariants structurés en système. »¹⁸¹

Ces phénomènes, qui vont des contraintes techniques et économiques à la représentation que l'on se fait en France de la langue et jusqu'à la conception que se fait le linguiste de sa propre activité, expliquent pourquoi il n'était ni aisé ni valorisé de travailler sur le français oral. Si, à l'image de ce qui s'est passé dans tous les pays, ce sont d'abord les corpus écrits qui ont bénéficié des technologies numériques, la France, hormis quelques précurseurs au premier rang desquels il convient de citer l'école aixoise, longtemps marginalisée, a accusé dans le domaine un retard certain notamment par rapport aux pays anglo-saxons. Or, il n'est pas indifférent de choisir en guise de données de l'écrit ou de l'oral. Même s'il convient de ne pas radicaliser l'opposition, les données écrites sont le plus souvent normées, ayant bénéficié de relectures et de corrections, comme c'est le cas des ouvrages et des journaux qui sont les plus sollicités pour la confection de corpus. Jusqu'à quel point ces données « policées » reflètent-elles le fonctionnement réel d'une langue, et la grammaire normative dans ce qu'elle a quelquefois de plus idiosyncrasique ? Difficile de ne pas rappeler ici l'épisode évoqué par Weinrich concernant l'emploi du passé composé et du passé simple par Madame de Sévigné. Racontant dans ses lettres¹⁸² l'épisode où son mari et elle-même avaient pu observer une comète, visible dans le ciel de Paris durant plusieurs jours, elle eut recours aux deux temps : *M. d'Artagnan VEILLA la nuit passée et la VIT fort à son aise*¹⁸³ ; *J'AI VU cette nuit la comète, sa queue est d'une fort belle longueur*¹⁸⁴. Pourquoi le PS dans un cas et le PC dans l'autre ? Weinrich répond (1973 : 294) :

si elle n'avait rencontré les intérêts bien compris de tous ceux dont la formation s'est accomplie sans terrain et qui, par l'abstraction, désavouent aussi les vérifications pratiques. L'origine sociale des gens de terrain confirmerait leur domination à l'entrée dans l'université et la manière particulière de perpétuer cette soumission en se livrant à des travaux d'enquête. L'institution, assurant sa reproduction par le recrutement presque exclusif d'étudiants en lettres, se pourvoit d'agents prédisposés, par leur formation, à sanctionner et perpétuer cet ordre. » Bergounioux (1992 : 18).

¹⁸¹ Voir également sur ce point Rastier (2005).

¹⁸² Suard (éd., 1864), *Lettres de Mme De Sévigné*, consultable à l'adresse : <http://www.gutenberg.org/files/43901/43901-h/43901-h.htm>

¹⁸³ De Mme De Sévigné À M. De Pomponne, mercredi 17 décembre 1664.

¹⁸⁴ De Mme De Sévigné À M. De Pomponne, lundi 22 décembre 1664.

Il faut en revenir à la règle des vingt-quatre heures, que Madame de Sévigné interprète ici de façon fort amusante et, oserais-je le dire, avec une logique toute féminine. M. d'Artagnan a visiblement calculé qu'il ferait mieux de rester éveillé jusqu'à l'apparition de la comète. Il ne se sera couché qu'ensuite. Pour lui, le moment où trois heures ont sonné appartient au jour précédent. Madame de Sévigné, au contraire, semble avoir trouvé cette solution quelque peu malcommode. Elle s'est couchée, et s'est levée à trois heures du matin. Pour elle, l'événement fait partie du jour suivant. La comète a bien été vue au même moment¹⁸⁵ par nos deux astronomes, mais cet instant appartient à des jours différents pour chacun d'eux, parce qu'ils ne se sont pas couchés à la même heure. Aucun doute, pour Madame de Sévigné, ce sont les heures passées à dormir qui tracent la frontière entre Passé simple et Passé composé.

En travaillant sur ce type d'écrit, on ne peut pas écarter le risque de ne retrouver dans les données que ce que la tradition grammaticale a bien voulu y mettre. Si on peut penser qu'il s'agit en l'occurrence d'un cas extrême, il sera tout de même difficile de nier que le rôle joué par la norme grammaticale est nettement plus important dans les productions écrites que dans les productions orales, où l'instantanéité de l'énonciation (hormis, naturellement les cas des discours professionnels, qui bénéficient en amont d'une préparation) privilégie les régularités linguistiques au détriment des règles normatives.

Aujourd'hui, alors que la situation a radicalement changé comme en témoigne la multiplication de projets de constitution de corpus oraux¹⁸⁶, tous les problèmes ne se trouvent pas pour autant résolus. L'oral, en tant que tel, n'est pas une panacée.

3.2.1.2. De la représentativité au concept de données situées

Si, ainsi que le rappellent Cappeau & Gadet (2007a), depuis l'émergence de l'approche sur corpus en linguistique s'est posée la question de leur représentativité, la plupart des auteurs¹⁸⁷, depuis Biber (1993), s'accordent sur le caractère problématique de cette notion, comme un horizon lointain. En tant qu'échantillon, le corpus est en effet passible de deux types d'erreurs statistiques qui rendent difficile toute tentative de généralisation, l'incertitude (*random error*) et la déformation (*bias error*) que Biber (1993 : 219-220) présente ainsi :

Random error occurs when a sample is not large enough to accurately estimate the true population; bias error occurs when the selection of a sample is systematically different from the target population it is intended to represent.

¹⁸⁵ Sur ce point, Weinrich s'est trompé, mais cela ne remet pas fondamentalement en cause l'essentiel de son analyse.

¹⁸⁶ Voir, pour une présentation déjà un peu datée des corpus oraux disponibles Cappeau & Gadet (2007b). Voir Branca-Rosoff & al. (2009) pour une présentation du CFPP (Corpus du Français Parlé Parisien) et Gadet (2012) pour une présentation de CIEL-F (Corpus International Ecologique de la Langue Française). Pour une présentation plus complète et plus actualisée des corpus oraux disponibles on peut consulter le site de COCOON (Collections de COrpus Oraux Numériques) à l'adresse <http://cocoon.huma-num.fr/exist/crdo/> et surtout le site de l'IRCOM (Consortium Corpus Oraux et Multimodaux) : <http://ircom.huma-num.fr/site/corpus.php>.

¹⁸⁷ Voir par exemple Habert (2000) et Cappeau & Gadet (2007a).

Comme l'explique Habert (2000), il est en effet hasardeux de spéculer par exemple sur le français mitterrandien à partir d'un extrait de 2000 mots d'une interview de F. Mitterrand par Y. Mourousi ou d'étudier le français journalistique du *Monde* en examinant les articles de la seule rubrique Économie de ce journal et ceci quel que soit le volume textuel rassemblé.

Pour minimiser ces erreurs, un corpus doit donc être à la fois suffisamment large (encore que la question de la taille soit problématique) et diversifié (encore que la variation ne soit pas « le vide-grenier » selon l'expression de Habert 2000). Si on le suit, on constate que la notion de représentativité s'appuie sur au moins trois visions différentes, souvent combinées dans les faits. La première privilégie les conditions de production : c'est le choix retenu en partie par le *British National Corpus* (BNC)¹⁸⁸ qui a sélectionné pour l'oral divers types de conversations spontanées recueillies à partir d'un échantillonnage démographique en termes d'âge, de sexe, de groupe social et de région. La deuxième vision de la représentativité privilégie les conditions de réception. Ainsi le BNC, tout en retenant des choix relatifs à la production, s'appuie, pour la sélection des ressources à l'intérieur d'un échantillon, sur la diffusion effective (par exemple, pour l'écrit de fiction, il a été procédé à une sélection parmi les listes des meilleures ventes, celles de prix littéraires, les indications de prêts en bibliothèque, etc.) pour garantir la meilleure représentation possible des usages majeurs de l'anglais¹⁸⁹. La troisième vision de la représentativité retient comme critère de sélection le type de textes ou de genres. Cette dernière vision est particulièrement problématique : si l'identification des genres et de leurs poids reste difficile pour l'écrit¹⁹⁰, elle l'est encore plus pour l'oral. « Notre connaissance de la « population » des données langagières, écrit Habert (2000) est encore extrêmement fragmentaire. Pour l'oral, par exemple, il n'est pas évident de spécifier les « genres »¹⁹¹ les plus produits ou les plus importants en réception. »

Toutes ces difficultés ont conduit à l'abandon du concept de représentativité au profit de celui de variationnisme. Sur ce point, le projet ESLO était un précurseur qui, loin de

¹⁸⁸ Corpus de 100 millions de mots mêlant écrits divers (textes de fiction à partir de 1960 et textes « informatifs » à partir de 1975) et différents genres de l'oral à hauteur de 10 %, ce qui en fait actuellement le plus gros corpus d'oral au monde.

¹⁸⁹ Voir pour une présentation de ce corpus les références de Habert. Concernant la situation en France, cet auteur affirme (2000) que la plupart des projets de constitution de corpus ne s'inscrivent pas vraiment dans cette perspective de représentation planifiée de la diversité langagière, mais reposent sur l'« assemblage » ou le recyclage de données préexistantes, voire sur des regroupements opportunistes. De Frantext au projet européen Parole (1996-1998), les difficultés juridiques conjuguées à la réticence des éditeurs privilégient de *facto* les textes « libres de droits », ou ceux que l'on doit à la politique offensive d'une société de presse en matière d'accès électronique à l'information (*Le Monde*). Des pans entiers de productions langagières se trouvent ainsi écartés.

¹⁹⁰ Voir Adam (1992).

¹⁹¹ Voir sur cette question Gerstenberg (2010a) et (2010b).

revendiquer une quelconque représentativité¹⁹², a cherché à recenser le mieux possible la variété des usages, aussi bien au niveau diaphasique qu'au niveau diastratique. Le choix d'Orléans constituait un choix par la négative sur le plan scientifique. Si on peut imaginer des justifications diverses allant de la commodité de sa proximité avec Paris à la taille de la ville, dont on pouvait reconstruire aisément la dynamique des formes linguistiques simultanément présentes¹⁹³, l'hypothèse qui ferait du choix d'Orléans celui d'une forme standardisée du français, représenté par une variété régionale non marquée doit être clairement écartée, comme en témoigne la présence parmi les témoins de l'enquête de profils (pieds noirs, parisiens, bretons...) qu'une enquête dialectologique aurait écartés.

Dans le projet ESLO2, nous avons cherché à garantir au maximum cette variété diastratique (une répartition des témoins des entretiens en tranches d'âge, sexes et catégories socio-professionnelles) et diaphasique (différents modules, représentant différentes situations communicationnelles, ont complété les entretiens : repas, école, boulangerie, module 24h, conférences...) ¹⁹⁴.

Données primaires (enregistrements) et données secondaires (transcription et autres annotations éventuelles) se trouvent ainsi systématiquement attachées à des métadonnées chargées de documenter le corpus¹⁹⁵. En renseignant sur les différentes variables concernant aussi bien le profil du locuteur (quand cela est possible, notamment dans le cas des entretiens) que la situation ou le genre de communication, les métadonnées permettent d'explicitier les critères de sélection et d'organisation des données et le bornage du corpus. Or c'est bien l'explicitation des bornes du corpus (conditions de production, de réception, contexte des usages, informations sociologiques sur les producteurs, genre, etc.) qui permet de juger de sa représentativité. Conçu comme un réservoir, ESLO permet au linguiste de constituer son propre corpus en retenant les critères qu'il juge pertinent en fonction de son objet.

Aucun corpus ne représente la langue : ni la langue fonctionnelle qui fait l'objet de la description linguistique, ni la langue historique, qui comprend l'ensemble des documents disponibles dans une langue. En revanche, un corpus est adéquat ou non à une tâche en fonction de laquelle on peut déterminer les critères de sa représentativité et de son homogénéité. La linguistique de corpus peut ainsi être objective, mais non objectiviste,

¹⁹² Blanc & Biggs, qui étaient parmi les initiateurs du projet ESLO, écrivent (1971 : 23) : « Les témoignages ne sont donc pas représentatifs ; nous n'y chercherons pas cet individu mythique, l'Orléanais moyen. »

¹⁹³ Voir Bergounioux & al. (1992).

¹⁹⁴ Une présentation plus détaillée, quoique encore incomplète, se trouve sur le site <http://eslo.huma-num.fr/>.

¹⁹⁵ Pour plus de détails techniques sur les choix opérés concernant la structuration de la documentation d'ESLO (formats, métadonnées Dublin Core, extensions préconisées par le programme OLAC, TEI, etc.), voir par exemple Abouda & Baude (2007) et les références citées.

puisque tout corpus dépend étroitement du point de vue qui a présidé à sa constitution. (Rastier, 2005 : 32).

Le concept de données situées permet de mesurer l'adéquation des données sélectionnées. Si, par exemple, aucune attestation de l'imparfait d'atténuation ne peut être identifiée dans un corpus, aussi vaste soit-il, l'hypothèse la plus probable ne consiste pas à affirmer l'inexistence de cette famille d'emplois dans la langue, mais l'inadéquation du corpus sélectionné pour l'étude de ce type d'emplois. Il convient dans ce cas de sélectionner, voire de constituer, un autre corpus¹⁹⁶.

En bref, ainsi que le note Habert (2000) :

[d]ans tous les cas, améliorer la représentativité d'un corpus consiste à préciser la production et la réception de chacun de ses composants, en lien avec les motifs qui ont conduit à la création du corpus, mais aussi à pouvoir déterminer sur des bases objectivables les différents emplois du langage auxquels on s'intéresse.

Cette perspective s'oppose frontalement à celle que dessine Milner (1989 : 117-118) quand il écrit :

C'est pourquoi il n'est pas d'une importance cruciale que l'exemple soit accompagné de précisions sur le lieu et l'instant où il a été effectivement prononcé. Cette précaution que se sont imposée des auteurs tels que Damourette et Pichon est plus apparente que réelle. Il est très rare qu'elle apporte des informations utiles : en vérité, elle ne s'impose que dans les cas marginaux, où justement le linguiste lecteur pourrait éprouver des doutes à l'endroit de la donnée proposée [...]. En réalité, il ne s'agit pas de documentation, mais d'information touchant les protocoles d'observation. Dans les circonstances ordinaires, le plus utile de l'exemple réside dans l'exemple lui-même.

Sans doute vérifiable dans le domaine de la syntaxe, cette affirmation ne peut qu'être contestée dans le champ sémantique de la temporalité. L'énoncé sur lequel travaille le linguiste se doit d'être situé avec précision. Il en va, ainsi qu'il a été précisé ci-dessus, de sa pertinence et du poids qu'on peut ou non lui accorder à l'intérieur du modèle : un énoncé est loin de délivrer d'emblée au linguiste toutes les informations qu'il recèle ; savoir qui l'a produit et dans quelle condition peut aider.

¹⁹⁶ Le constat a été fait dans Abouda (2004b). Examinant une partie du corpus d'Aix (qu'une collègue du GARS a bien voulu me fournir) et la partie disponible à l'époque d'ESLO j'ai noté (p. 60) : « Les résultats de ces collectes, même s'ils sont riches en enseignements, méritent d'emblée d'être relativisés. D'abord parce que le type d'IMP recherché ne peut se rencontrer que dans des contextes pragmatiques très étroits, et donc finalement peu présents dans les corpus examinés. Ceci est particulièrement vrai des corpus oraux qui, correspondant le plus souvent à des entretiens, n'enregistrent quasiment jamais les micro-situations pragmatiques où il est théoriquement possible de rencontrer un IMP d'atténuation. » Ce type de constat sera à l'origine de la constitution de deux modules ESLO2 (malheureusement non encore disponibles en ligne) : le module « boulangerie » et « guichets ».

Je voudrais terminer en réaffirmant l'importance d'une transcription alignée sur le son¹⁹⁷, qui constitue l'une des manières les plus sûres de situer une production orale. Outre ses vertus économiques¹⁹⁸ – l'abandon du mythe d'une transcription idéale au profit d'une transcription vue comme une manière de naviguer dans le son – elle permet une avancée méthodologique majeure. « En effet, notent Abouda & Baude (2007 : 147), avec des transcriptions alignées sur le signal sonore, l'oral devient physiquement l'objet d'étude et est systématiquement disponible en même temps que la transcription. Le retour aux données peut alors être systématique, ce qui est de nature à faciliter les procédures de vérification, étape essentielle du travail scientifique ».

Si, en effet, ainsi que l'écrit Bergounioux (2004 : 80), « est entendu ce qui est attendu », que dire de ce qui est compris dans ce qui est lu ?¹⁹⁹ L'expérience que j'ai acquise dans l'annotation de corpus oraux, notamment lors des études sur le futur, confirme à quel point il est primordial de ne pas travailler sur la transcription, mais directement sur la matière sonore. Plus long et fastidieux²⁰⁰, ce retour systématique au son s'est avéré de plus en plus nécessaire. Les informations dont on peut disposer à l'écoute, et seulement à l'écoute, des énoncés (et qui peuvent aller de l'intonation, à la pause jusqu'à l'environnement sonore de la situation) s'avèrent souvent déterminantes pour minimiser les risques d'erreur lors de l'annotation. Ainsi, dans l'exemple suivant :

(60) ESLO1_CONF 504

vous voyez que c'est analysé tout en détail n'est-ce pas *je vais vous en remettre*
un exemplaire à chacun

¹⁹⁷ Il n'est pas selon moi exagéré de comparer, au niveau de ses effets méthodologiques sur le travail du linguiste, l'alignement de la transcription avec le signal sonore, à la découverte du magnétophone. Tant qu'il n'est pas encore technologiquement possible de naviguer directement dans le son, cet alignement constitue le meilleur moyen de concilier la navigation dans le corpus (par recherche de chaînes de caractères, lemmes, catégories d'annotation,...) qui n'est actuellement possible que dans la transcription, avec le retour nécessaire à la matière sonore.

¹⁹⁸ Une transcription fine, ainsi que le note Habert (2005 : 57), « peut demander jusqu'à 40 ou 60 fois le temps de l'enregistrement, voire bien davantage si les conditions d'enregistrement sont difficiles ». Que l'on fasse les comptes.

¹⁹⁹ Qui ne se méfie plus des échanges mails dans ce mélange d'instantanéité et d'absence de support qui laisse ouvertes toutes les possibilités interprétatives, y compris les pires ?

²⁰⁰ Mais l'évolution technologique est extrêmement rapide : des outils informatiques améliorent sans cesse l'ergonomie du travail du linguiste. Nous sommes actuellement en train d'évaluer avec Flora Badin, l'ingénieur recherche du LLL, l'opportunité d'utiliser à la place de (ou parallèlement à) TXM (<http://textometrie.ens-lyon.fr/>), le logiciel Sonal développé par Alex Alber (<http://www.sonal-info.com/>). Ce logiciel, s'il n'est pas couplé (pour l'instant) à un étiqueteur morphosyntaxique, permet de travailler directement sur des transcriptions en lien avec le son, ce qui permettrait de gagner en ergonomie et en productivité : l'annotation se fait pour l'instant sur des tableurs CSV extraits de TXM mais l'écoute du son se fait, pour chaque ligne, sur le site du corpus ou sur la base de données interne.

on ne peut comprendre que le futur périphrastique est un futur imminent (donc non remplaçable par un futur synthétique) qu'en écoutant le bruit du papier des exemplaires qui allaient être distribués. La transcription peut déformer le sens : à plusieurs reprises, à la relecture d'annotations que nous avons, Marie Skrovec et moi, effectuées quelques semaines ou quelques mois auparavant, nous nous sommes trouvés dans la situation inconfortable de devoir contester nos propres analyses, pourtant validées à deux. Seule la réécoute de l'exemple a permis à chaque fois de conforter nos annotations initiales. C'est à l'examen de ce type de phénomènes, entre bien d'autres, qu'on « on pourrait s'apercevoir en quoi une linguistique de l'oral nous a fait, à défaut d'être plus savants, moins crédules » (Bergounioux, 1992 : 19).

3.2.1.3. Des corpus fantômes aux corpus disponibles

Si on envisage un instant le nombre de corpus, écrits ou oraux, qui ont dû être constitués par des équipes ou des chercheurs isolés (notamment à l'occasion de recherches doctorales), on mesure le gâchis économique et scientifique à l'échelle d'un pays comme la France. Ces corpus qui, pour un doctorant peuvent prendre jusqu'au tiers du temps total consacré à la thèse, sans doute plus pour certaines enquêtes du terrain, ne sont souvent exploités, le plus souvent partiellement, que pour une étude particulière, à l'issue de laquelle on les range sur un support vite obsolète (à mon époque, sur des disquettes) avec la ferme intention d'y revenir pour rentabiliser le temps qu'on y a consacré. A supposer qu'on y revienne, l'exploitation reste nettement en dessous de ce que permettrait une exploitation partagée. Mais on ne prête pas ses corpus !

Ainsi, Suzanne Lafage, parmi les premiers linguistes à travailler sur le français d'Afrique, a raconté qu'elle recevait dans les années 1980 des requêtes de "prêter ses corpus". Elle y voyait comme sous-entendu qu'elle était juste bonne à les recueillir, et que d'autres sauraient les intégrer dans un ensemble théorique significatif..., séparation des rôles qui perdure de nos jours, tout autant que la dissymétrie de prestige qui accompagne chacune des étapes. (Cappeau & Gadet, 2007a : 102).

Le résultat est connu : ainsi que nous l'écrivions avec Olivier Baude (Abouda & Baude, 2007 : 145), « les corpus sont toujours évoqués dans les travaux, mais ne sont que très rarement diffusés. Ils jalonnent les articles et les thèses comme les fantômes hantent les couloirs et les tours : toujours évoqués comme preuve mais n'apparaissant à nul autre qu'à celui qui en parle. »

Esquissant une typologie des corpus, en fonction de leur disponibilité, nous avons identifié (*ibid.*) quatre cas de figure²⁰¹ :

- la plupart des corpus ne sont pas diffusés par des chercheurs qui souhaitent garder une priorité scientifique sur leur travail de collecte ;
- pour d'autres corpus, constitués dans le cadre de recherches particulières, le manque d'interopérabilité complique leur utilisation par d'autres chercheurs : ce sont les conditions de collecte (et/ou d'annotation) qui ne permettent pas la diffusion de ces données « jetables » ;
- des corpus étayant certains travaux ne sont pas présentés par l'auteur qui ne fournit pas la moindre information concernant la possibilité ou non d'accès aux données, pourtant seule garantie d'un travail scientifique en principe ouvert à la falsification ;
- enfin, certains corpus, parce qu'ils sont disponibles, acquièrent le rang de corpus de référence : cela a été le cas de Frantext, du *Monde*, et du Costo... avec les conséquences qu'on imagine en termes de représentativité.

La solution à cette situation réside d'abord dans la reconnaissance, scientifique et institutionnelle, du travail de constitution de corpus, qui présuppose une prise de conscience collective de la nécessité de cette démarche. Mais ceci nécessite également d'harmoniser formats, standards et outils, afin de garantir l'accessibilité des données et asseoir leur interopérabilité. Ce sont des chantiers qui sont actuellement largement engagés.

Le LLL, qui a obtenu des financements publics pour ESLO, a intégré dès le début du projet la mise à disposition des données au centre de ses objectifs. Le corpus sera ainsi librement disponible et déposé à la BnF. Mais l'accessibilité des données a dépassé actuellement le cadre de la mise à disposition de corpus brut pour questionner la mise à disposition, avec un objectif de falsifiabilité de la version (ou interprétation) du corpus à l'origine d'une recherche donnée. En effet, ainsi que le note Habert (2000) :

[u]ne collection de textes permet d'engendrer de multiples corpus distincts. Chacun de ces corpus peut donner lieu à des annotations variées qui constituent autant d'« interprétations », au sens large : étiquetage morpho-syntaxique, projection de catégories sémantiques, lemmatisation, etc. Chacun de ces traitements produit une version différente du corpus. L'écart entre les versions peut être plus ou moins important.

²⁰¹ Mais je ne sais toujours pas où ranger ce « corpus disponible mais non-accessible », dont m'a parlé un jour un conférencier en guise de réponse à ma question concernant la disponibilité du corpus qu'il présentait et qui me semblait intéressant...

Dans le cadre d'une procédure de vérification, il importe que l'analyse soit reproductible et que le lien entre données et analyses puisse à chaque étape être rétabli. Il est dans ce cas décisif que l'on puisse garder la mémoire non seulement des données mais également des différents traitements qu'elles ont dû subir.

C'est le prochain chantier que la *Revue de Sémantique et Pragmatique*²⁰² entend ouvrir en collaboration avec la très grande infrastructure de recherche (TGIR) Huma-Num²⁰³ afin d'établir la correspondance entre le corpus, s'il en est un, et l'analyse qui l'exploite pour chacun des articles publiés sur le site internet de la revue.

3.2.2. *Etude de la temporalité en corpus*

Je me propose d'illustrer la démarche de travail sur corpus dans le domaine de la temporalité par deux recherches menées sur ESLO, en m'attachant surtout à montrer les difficultés d'une telle entreprise ainsi que les avantages qu'elle semble offrir.

3.2.2.1. *Le présent de l'omelette*

Une première expérience d'étude sur la temporalité en corpus a porté sur le présent de l'omelette (Abouda, 2012), dans un sous-corpus réduit, échantillon d'Eslo1²⁰⁴, composé de données de type « recettes de cuisine » obtenues à partir de la question « Comment est-ce qu'on fait une omelette (chez vous) ? ». D'une durée totale d'environ une heure et quarante minutes, ce corpus comporte 96 échanges question-réponse, variant de 16 secondes à un peu plus de trois minutes.

Si cette première étude a porté sur un corpus relativement réduit, elle a tout de même permis de tirer avantage de données orales et sociologiquement situées. En effet, elle a d'abord montré l'existence de cette famille d'emplois, qui n'avait pas été, me semble-t-il, identifiée dans les recherches antérieures : ni dans les études classiques sur le genre « recettes de cuisine »²⁰⁵ (les recettes, écrites, sont soit à l'infinitif, soit à l'impératif, formes

²⁰² Le retard accumulé pendant des années n'a été absorbé que récemment, ce qui a enfin rendu possible d'envisager une édition en ligne de la revue. Le scénario retenu par l'équipe éditoriale privilégie une mise en ligne des anciens numéros sur le portail de publication électronique de revues scientifiques en sciences humaines et sociales (www.persee.fr) et la publication, avec une période d'exclusion, de numéros plus récents sur le portail www.revues.org. C'est sur ce dernier portail que la mise à disposition des données, parallèlement au texte de l'auteur, est envisagée.

²⁰³ <http://www.huma-num.fr/>

²⁰⁴ Ce mini-corpus a fait l'objet de plusieurs recherches menées par des membres du groupe ESLO qui paraîtront dans un ouvrage collectif dirigé par Bergounioux (à paraître).

²⁰⁵ Notamment Greimas (1983), Revaz (1987) et Adam (2001).

inexistantes ou marginales dans les recettes orales²⁰⁶), ni dans celles portant sur le présent de l'indicatif²⁰⁷, qui, s'appuyant généralement sur des exemples fabriqués ou empruntés à des corpus écrits, n'avaient pas repéré ce type d'emploi malgré ses spécificités. D'autre part, cette étude a permis d'identifier pour le présent trois emplois distincts dont chacun, émergeant d'une combinaison spécifique avec l'un des trois pronoms personnels rencontrés dans le corpus, i.e. *je*, *on* ou *vous*, exhibe des propriétés énonciatives relativement distinctes des deux autres :

- (61) FU200: [rire] effectivement ça m'amuse + [rire] + [pron=pi] *je prends* des œufs déjà + [rire] du sel du poivre enfin l'assaisonnement l'assaisonnement [bb] + [conv] dans la poêle *je mets* de du beurre un petit peu d'huile ça évite que mon omelette attache après avoir bien battu mes œufs *je fais* cuire mon omelette *je la retourne* quand elle est dorée d'un côté # Voilà # (Omelette 007)
- (62) 1135: # oh bah c'est pas difficile là de faire une omelette bah + [rire] *vous mettez* du beurre dans la dans la poêle *vous battez* les œufs *vous les collez* dessus et zou # (Omelette 025)
- (63) 1299: *on casse* d'abord les œufs *on les bat on ajoute* un peu de crème du sel du poivre et puis *on verse* dans une poêle dans laquelle on a mis un peu de matière grasse du beurre ou de l'huile ou bien une poêle TEFAL sans matière grasse (Omelette 027)

En effet, si dans chacune de ces trois configurations, l'emploi du présent a pu sans difficulté être relié à la valeur en langue de ce temps, à savoir une simultanéité entre l'intervalle d'énonciation et celui du procès, j'ai cru, m'inspirant de Gosselin (2005), pouvoir identifier trois mécanismes distincts de lien : une validation paradigmatique pour les recettes en *vous*, où le sujet, ne se confondant pas avec le co-énonciateur, peut correspondre à toute instance susceptible de valider dans le futur la recette ; une validation sous le mode d'une présentation du contenu d'une représentation préexistante pour les recettes en *on*, où le témoin ne fait que « citer » une recette standard (souvent nationale²⁰⁸) ; et une validation par

²⁰⁶ Sur les 88 recettes orales étudiées dans Abouda (2012), 6 sont à l'infinitif et 82 ont comme mode dominant l'indicatif, dont 81 sont au présent.

²⁰⁷ Voir, entre autres, Imbs (1960), Riegel, Pellat&Rioul (1998) et Wilmet (1997).

²⁰⁸ La question, posée par des Anglais et contenant souvent un circonstanciel lui-même ambigu, i.e. *chez vous*, n'est pas sans ambiguïté et n'a visiblement pas été interprétée de la même manière d'un témoin à l'autre : lui demande-t-on sa recette personnelle de l'omelette ou la façon de faire une omelette en France ?

réduplication de l'intervalle énonciatif, qui prend la forme d'une simulation de présence où le témoin se transporte dans une situation où il se montre en train de réaliser son omelette.

Or, il est clairement établi dans cette recherche que l'hypothèse de cette tripartition des emplois ne s'appuie pas uniquement sur des propriétés linguistiques, mais également sur des propriétés métalinguistiques accessibles dans ce corpus en même temps que les données. « Filtrée par des processus communicationnels complexes où le profil sociologique du témoin et son sexe jouent un rôle important, ai-je conclu (2012 : 92), la réception de la question n'a pas été identique chez toutes les personnes interrogées. S'il n'est pas hasardeux de penser qu'elles ont toutes répondu par la production d'une recette, on doit reconnaître que les recettes produites dans ce corpus ne sont pas homogènes : certaines relèvent du genre instructionnel (certaines recettes en *vous*), d'autres du genre procédural (les recettes en *on*, où le témoin décrit les principales étapes de la réalisation d'une omelette standard), tandis qu'avec les recettes en *je* nous passons clairement du comment faire à comment je fais... »

3.2.2.2. *Distribution futur simple / futur périphrastique*

Depuis la mise à disposition d'ESLO, en septembre 2013, des études sur corpus oral de plus grande envergure sont devenues possibles. Si je n'ai pas abandonné l'objectif initial que je m'étais fixé depuis quelque temps, à savoir l'étude en corpus de l'imparfait, il m'a semblé utile d'étudier au préalable certaines autres formes temporelles réputées moins difficiles. L'objectif était d'une part d'acquérir une certaine maîtrise technique des outils que propose le domaine du TAL pour le traitement de corpus numériques, et, de l'autre, de forger une démarche de traitement de données en masse.

Le choix d'étudier la distribution des deux formes, analytique et synthétique, du futur, appelées communément respectivement futur périphrastique (FP) et futur simple (FS), s'est alors imposé comme une évidence. D'une part, parce que seules des études sur corpus de grande envergure semblaient pouvoir départager les deux hypothèses principales qui se disputent le traitement du rapport FP/FS, i.e. l'hypothèse d'une concurrence diachronique devant à terme consacrer l'usage du FP au détriment de celui du FS (Mauger (1968), Togeby (1982), Fleischman (1983), etc.) ; et l'hypothèse d'une complémentarité entre les deux formes parvenant chacune, en se spécialisant, à se maintenir dans des configurations linguistiques spécifiques²⁰⁹. D'autre part, ESLO nous a semblé particulièrement bien placé pour éclairer,

²⁰⁹ Les différents travaux divergent toutefois sur la nature de cette complémentarité, tour à tour traitée en termes aspecto-temporels (Jeanjean, 1988 ; Bilger, 2001), syntaxiques (Emirsonian & Sankoff, 1986), modaux (Confais,

dans le cadre d'une diachronie courte, l'éventuelle évolution historique de la répartition FP/FS, là où les études antérieures se situaient dans une perspective strictement théorique (Fleischman) ou s'appuyaient sur des corpus²¹⁰ qui nous ont paru à la fois relativement restreints sur le plan quantitatif, et surtout, de nature synchronique, peu adaptés à l'étude de la relation diachronique FP/FS.

Nos recherches sur ces deux formes du futur, initiées avec Marie Skrovec il y a maintenant plus de deux ans, si elles avaient initialement été conçues comme une première expérience de travail sur ESLO dont nous pensions qu'elle serait relativement rapide, sont toujours en cours : elles ont donné lieu à 3 communications dans des colloques internationaux, et à deux articles (Abouda & Skrovec, à paraître et soumis), et feront l'objet prochainement de deux autres communications et sans doute d'autres articles.

Je me propose dans ce qui suit de retracer les principales étapes de ces recherches, ce qui complètera utilement les deux articles (Abouda & Skrovec (2015, à paraître & soumis)) joints au second volume de ce dossier.

3.2.2.2.1. Constitution du sous-corpus d'étude

Nous avons construit un sous-corpus échantillonné d'environ 1 million de mots (80 heures et 51 min.), composé, à parts égales, d'extraits d'ESLO1 et d'ESLO2. Il n'était pas en revanche possible, au moment où nous avons sélectionné nos données, d'assurer un équilibre diaphasique entre (et à l'intérieur de) chacun des deux extraits ESLO1/2. Dans un corpus disponible majoritairement constitué d'entretiens, nous avons tout de même réussi à intégrer, à hauteur de 20%, deux genres interactionnels de « contrôle », supposés se situer de part et d'autre d'une échelle diaphasique, i.e. les repas (en famille ou entre amis) et les conférences.

Le tableau suivant récapitule les caractéristiques quantitatives (durée en minutes et nombre de locuteurs) du corpus et sa composition diachronique et diaphasique :

1990 ; Maingueneau, 1994 ; Laurendeau, 2000...), ou encore diaphasiques (entre autres, Arrivé, Gadet & Galmiche, 1986).

²¹⁰ Emirkanian & Sankoff 1986, Jeanjean 1988, Bilger 2001, Fleury & Branca 2010...

(64)

	Genre	ESLO1	ESLO2	Corpus global
Durée (en min.)	Conférences	192	186	378
	Repas	196	201	397
	Entretiens	2042	2034	4076
	Total	2430	2421	4851 (80h et 51min.)
Nombre de locuteurs	Entretiens	30	30	60

A l'intérieur de cette composition, s'est parallèlement posée la question de la sélection des enregistrements notamment pour les entretiens. S'agissant d'un corpus de données situées, nous avons essayé de garantir au mieux un équilibre diastratique (entre ESLO1 et ESLO2 et à l'intérieur de chacun d'entre eux), en termes de profil de locuteurs, choisis selon les variables de sexe (H/F), d'âge (3 tranches) et de catégorie socio-professionnelle (5 classes).

Pour opérer une telle sélection, dont les détails sont précisés en annexe²¹¹, en fonction de ces trois variables²¹², nous avons pris quelques décisions et dû consentir quelques aménagements.

La première décision a été de classer les locuteurs en trois tranches d'âge : 15-35, 35-60 et plus de 60 ans. Lorsque, à l'intérieur d'une tranche d'âge, les autres variables nous laissaient le choix entre plusieurs témoins, nous avons privilégié ceux dont l'âge était le plus éloigné des tranches voisines.

Concernant la variable de sexe, si nous avons souhaité qu'il y ait, globalement et à l'intérieur de chaque classe d'âge et de chaque CSP, autant d'hommes que de femmes, la combinaison des différents critères n'a pas toujours permis cet équilibre. Nous avons alors donné la priorité aux deux autres critères en acceptant que la catégorie sexe ne soit pas

²¹¹ Cf. Annexe 1 : Corpus futur, p. 114 sq.

²¹² D'autres critères de sélection ont été utilisés dans les cas où la conjonction de ces trois variables principales nous laissait le choix entre plusieurs entretiens. Le premier a consisté à privilégier les locuteurs « DIA » (locuteurs interrogés dans ESLO1 et ESLO2 : case locuteur en grisé dans l'annexe, p. 114 sq.). Le second critère a été de choisir l'entretien dont la durée est la plus proche de 60 minutes.

systématiquement équilibrée dans chaque combinaison tranche d'âge-CSP. Ainsi que nous pouvons le voir (dans les tableaux de l'annexe, lignes en grisé et valeur sexe en gras), ce cas de figure s'est présenté trois fois : dans la classe de plus de 60 ans, aussi bien dans ESLO1 que dans ESLO2, au lieu d'avoir pour chaque catégorie sociologique un représentant de chaque sexe, nous avons admis que l'équilibre F/H puisse n'avoir lieu qu'au niveau d'un couple de CSP contiguës²¹³. Il n'a pas été en revanche possible de compenser de la sorte le déséquilibre dans le 3^e cas qui s'est présenté (dans ESLO2_15-35, 2 hommes dans la catégorie E).

De même, afin de nous approcher le plus possible d'un équilibre quantitatif (mesuré en minutes et non en nombre d'entretiens) global entre ESLO1 et ESLO2, nous avons dû accepter qu'il y ait dans le dernier sous-corpus un entretien de plus²¹⁴, ce qui ne pouvait qu'affecter l'équilibre en termes de sexe et de CSP.

Mais les compromis les plus importants, et sans doute les plus discutables, ont concerné les variables sociologiques.

Dans l'enquête initiale, les initiateurs d'ESLO1 ont fait preuve, ainsi que le souligne Bergounioux & al. (1992) d'une « extrême modernité » dans la construction de l'échantillonnage sociologique, en dépassant les catégories INSEE, utilisées initialement mais rapidement abandonnées au profit d'une nouvelle classification qui anticipait largement les reformulations qu'allait être proposées par Bourdieu, avec notamment la notion de *capital culturel*²¹⁵, infléchissant ainsi une sociologie purement économiste au profit d'une prise en compte parallèle des habitudes culturelles des témoins. Clairement identifiable au niveau du questionnaire, la filiation avec les conceptions qui se développaient au Centre de Sociologie Européenne dirigée à l'époque par Bourdieu au sein de l'EHESS, était explicitement revendiquée :

« Le questionnaire sociolinguistique a été établi par les soins de M.B. Vernier, sociologue, élève du Professeur P. Bourdieu [...]. L'hypothèse sous-jacente à ce questionnaire prend appui sur une théorie sociale de la transmission de la culture et de la langue. L'origine sociale et le niveau d'éducation du témoin concourent, avec ses activités socio-professionnelles, à déterminer sa compétence linguistique d'une part, et son attitude envers les normes du langage d'autre part. Selon cette hypothèse, plusieurs types d'attitudes seraient en corrélation avec des niveaux socio- culturels, attitudes qui iraient de l'indifférence à la dévotion. Au sommet comme au bas de l'échelle socio-culturelle, cette attitude se traduirait par

²¹³ Dans Eslo1, deux femmes dans la catégorie C, compensées par deux hommes dans la catégorie D ; dans ESLO2, 2 femmes en D, composées par deux hommes en E.

²¹⁴ Dans ESLO2_Âge + 60, 11 entretiens (au lieu de 10) ont dû en effet être sélectionnés.

²¹⁵ Voir Bourdieu (1979) ainsi que Détrez (2005) pour une présentation critique.

l'indifférence : mais tandis que dans la classe "Supérieure" cette indifférence est la marque d'une liberté par rapport à la norme et est liée à la capacité de jouer avec un grand nombre de registres, au bas de l'échelle, en revanche, cette indifférence est le résultat soit de l'ignorance, soit d'une certaine conscience de classe. Entre ces deux niveaux se situent d'une part ceux qui, ayant défini la norme linguistique, y adhèrent et cherchent à l'imposer ; d'autre part ceux qui, aspirant à une ascension sociale, s'efforcent d'imiter ce qui est pour eux la langue de prestige. » Blanc & Biggs (1971 : 18).

Le résultat fut la conception d'une nouvelle grille comprenant cinq catégories sociales – sténographiées « échelle AM » (du nom de son inventeur, Alix Mullineaux) – dont chacune était identifiée par une lettre de A à E, où l'on tente d'agréger aux CSP identifiées par l'INSEE des caractéristiques culturelles ainsi que des critères relatifs à la mobilité géographique²¹⁶.

L'adaptation nécessaire aux théories sociologiques contemporaines et aux nouvelles réalités sociétales n'a pas permis le maintien dans ESLO2 de la classification AM. C'est l'un des chantiers que l'équipe ESLO n'a pas pu rouvrir. Les locuteurs ont donc simplement été classés, conformément à la nouvelle nomenclature de l'INSEE, en 8 catégories socio-professionnelles²¹⁷.

Lors de la constitution de notre sous-corpus, afin de garantir une certaine comparabilité entre ESLO1 et ESLO2, nous avons repris, en cherchant à les adapter au cas par cas, les catégories AM. L'objectif était de traduire par l'un des agrégats A-E l'ensemble des informations dont nous disposons concernant le profil du locuteur, aussi bien au niveau des métadonnées (profession, études et diplômes, âge) qu'au niveau des informations restituables au sein de l'entretien concernant son appartenance sociale.

Menée sans appui sociologique et avec des critères complexes et non toujours explicités, la tâche était toutefois particulièrement fastidieuse et délicate, notamment concernant les catégories intermédiaires. En effet, s'il nous a paru relativement aisé d'identifier des locuteurs

²¹⁶ Voir Mullineaux & Blanc (1982).

²¹⁷ La nomenclature, qui date de 2003, est détaillée sur le site de l'INSEE : <http://www.insee.fr/fr/methodes/?page=nomenclatures/pcs2003/pcs2003.htm>

- 1 Agriculteurs exploitants
- 2 Artisans, commerçants et chefs d'entreprise
- 3 Cadres et professions intellectuelles supérieures
- 4 Professions Intermédiaires
- 5 Employés
- 6 Ouvriers
- 7 Retraités
- 8 Autres personnes sans activité professionnelle

comparables sur les « extrémités » de l'échelle²¹⁸, pour les catégories B-C-D, la classification nous a paru extrêmement périlleuse. Quid par exemple de la femme au foyer, titulaire d'un bac + 5 en Droit, mariée, 49 ans, habitant Olivet – village aisé en périphérie d'Orléans – et parlant anglais, allemand et italien ?

Au-delà de ces problèmes inhérents à toute tentative de classification, il nous a paru que les difficultés proviennent surtout des nouvelles réalités sociales qui ont rendu nettement caduques les critères qui fondaient l'échelle AM. Le critère de diplôme, par exemple, ne pourrait avoir qu'une pertinence toute relative : un bac obtenu en 2015 ne peut avoir ni les mêmes débouchés professionnels ni la même valeur sociale qu'un bac de 1970. La féminisation de certains métiers, le déclassement de certains autres, l'émergence de nouvelles élites, l'apparition d'une nouvelle classe populaire d'origine immigrée qui semble cumuler les handicaps économiques et culturels, etc. sont autant de facteurs qui rendent aujourd'hui nécessaire une nouvelle nomenclature sociale critique. De même, les locuteurs classés tout en bas de l'échelle AM dans ESLO1 ne sont pas réellement comparables aux E de ESLO2. On pourrait d'ailleurs penser que les vrais E²¹⁹ de notre époque (chômeurs n'ayant jamais cotisé ou très peu, jeunes déscolarisés, etc.) ne sont même pas présents dans ESLO2. Malgré les efforts de l'équipe et la mobilisation de divers réseaux, ce type de profils est quasiment absent de l'enquête, notamment dans la catégorie des 15-35 ans. Le problème a été identifié dès ESLO1 où les ouvriers (hormis les délégués syndicaux, habitués à la prise de parole en public) étaient déjà sous-représentés...

Toutes ces difficultés montrent que le sous-corpus constitué ne peut en aucun cas être considéré comme sociologiquement représentatif : il s'agissait simplement de réduire les variables entre les deux extraits ESLO1/ESLO2 en sélectionnant des profils pouvant offrir la meilleure comparabilité qualitative possible. Dans le cadre de nos recherches sur le futur, le paradigme le plus pertinent était celui qui opposaient les données sur le plan micro-diachronique.

3.2.2.2.2. TXM

Pour exploiter au mieux un tel corpus numérique, nous avons exploré avec Flora Badin, ingénieure de recherche au LLL, différentes solutions logicielles disponibles. Notre choix est

²¹⁸ D'une part, la catégorie A (cadres et professions intellectuelles supérieures), et, de l'autre, la catégorie E (sans profession, ouvrier, niveau d'études : primaire).

²¹⁹ On peut se demander, par exemple, si on a raison de classer le locuteur BV1, jeune ouvrier sans diplôme et sans qualification, en E, ne serait-ce parce qu'il est en couple avec une jeune étudiante en master et qu'il a connu une mobilité géographique...

tombé sur TXM²²⁰, qui semblait bien placé pour répondre aux besoins spécifiques de notre recherche sur le futur. Ce logiciel permet en effet de travailler sur des corpus de transcription de l'oral²²¹ et de constituer, à partir du corpus global préalablement importé, des sous-corpus et des partitions, ce qui rend plus aisé d'appliquer des calculs de contraste entre différents groupes constitués de données (par exemple : Eslo1 vs Eslo2, Entretiens vs Conférences, groupes de locuteurs A vs B, ainsi que différentes combinaisons possibles entre ces partitions). D'autre part, s'il faut, semble-t-il pour des raisons de licence, avoir préalablement installé l'étiqueteur morphosyntaxique et lemmatiseur TreeTagger²²², l'application TXM est livrée avec un plugin pour différentes langues, dont le français, ce qui permet d'extraire directement grâce à des requêtes CQL²²³ toutes les occurrences du futur simple et du futur périphrastique.

Les listes d'occurrences ont par la suite été exportées dans des tableaux CSV, où elles ont été corrigées manuellement²²⁴ avant d'être annotées (sous-spécifiées) pour un certain nombre de traits.

L'étape finale de cette démarche a consisté à réinjecter sous TXM les listes d'occurrences (corrigées et annotées) ainsi que les étiquettes d'annotation forgées. L'objectif est d'explorer les données en textométrie en croisant approches quantitative et qualitative. Les tendances quantitatives dégagées, si elles étaient quelquefois significatives en elles-mêmes, nous ont souvent obligés, pour être correctement interprétées, à un retour à des analyses qualitatives où il a été question d'identifier les éventuelles variables distributionnelles et co-textuelles (négation, personne, adverbes, etc.) pertinentes dans la distribution FP/FS²²⁵.

Comme on vient de le constater, l'annotation d'un corpus numérique outillé permet, dès lors qu'il s'agit d'un corpus documenté (précisant notamment le type de genre interactionnel,

²²⁰ Heiden & al. (2010) et Heiden (2010), voir également le site : <http://textometrie.ens-lyon.fr/>

²²¹ La transcription encode les locuteurs et des points de synchronisation temporelle avec le son, ce qui rend possible de retrouver plus facilement l'extrait sonore. Mais, comme il a été dit ci-dessus, il paraît aujourd'hui possible et souhaitable qu'il y ait un accès plus direct et plus ergonomique aux fichiers sonores.

²²² Là non plus il ne s'agissait pas du seul étiqueteur disponible en open source, mais de celui que nous avons choisi en raison de sa robustesse pour l'oral et sa compatibilité avec TXM. Il est téléchargeable sur le site : <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

²²³ La syntaxe de la requête CQL pour le FS : [frpos="VER:futu"] et celle pour le FP : [frlemma="aller"&frpos="VER:pres"][] {0,5} [frpos="VER:infi"].

²²⁴ La correction visait à écarter les erreurs d'étiquetage et, en ce qui concerne le FP, à exclure les occurrences d'*aller* au sens lexical plein (verbe de mouvement).

²²⁵ Cette distribution s'éclaire quelquefois mieux lorsqu'on croise ces informations avec des métadonnées relatives au genre interactionnel et à l'activité de communication en cours. Par exemple, le futur périphrastique historique n'apparaît dans notre corpus que dans ESLO2 : loin de correspondre à une quelconque évolution micro-diachronique, cet usage du FP s'explique le fait que l'une des conférences d'ESLO2 est une conférence d'histoire. Voir Abouda & Skrovec (à paraître, § 3.4).

l'activité de communication en cours – par exemple, dans une conférence, il peut y avoir en plus de l'exposé, une interaction questions-réponses – mais aussi la thématique, qui peut avoir un rôle explicatif important dans le domaine de la temporalité), de situer les données linguistiques (à la fois en interne, sur le plan syntagmatique, et au niveau des métadonnées situationnelles et diastématiques) et d'accéder aisément à des hypothèses expliquant les tendances quantitatives observées dans des grands ensembles. Malgré les progrès, inimaginables il y a encore peu de temps, rendus possibles par l'appui de l'informatique, les problèmes techniques que soulève l'annotation sémantique sont encore nombreux.

Il convient tout d'abord de préciser que la tâche d'annotation des types d'emplois et sous-spécifications a été réalisée entièrement manuellement. L'attribution d'une valeur pour chacune des étiquettes des 3370 occurrences de FS ou FP identifiées dans le corpus est une tâche à la fois longue (le temps de travail peut être estimé en centaines d'heures) et fastidieuse (manipulations techniques nombreuses et répétitives). Conjuguées à la complexité d'analyse, sur laquelle je reviendrai ci-dessous, ces difficultés ont de quoi décourager la démarche. La pérennité de celle-ci passe par un gain en ergonomie et en productivité.

Deux points en particulier méritent d'être développés. D'abord, ainsi qu'il a été expliqué ci-dessus, l'accès au son (et pas uniquement à la transcription), nécessaire dans le cadre de notre démarche, n'est pas directement possible dans la version utilisée de TXM, et nécessite un détour par un autre environnement logiciel pour écouter l'extrait sonore. Il est nécessaire soit d'œuvrer pour une adaptation logicielle de TXM, soit d'opter pour une autre application²²⁶.

Un autre problème technique mérite d'être évoqué ici : dans une version antérieure de TXM, il n'était pas possible d'exploiter les métadonnées du corpus. Lorsque, sur notre demande, l'ajout d'un plugin a rendu possible cet accès, les tours de parole ont disparu, ce qui, entre autres, pose problème à la fois pour un balisage rapide des occurrences au niveau des transcriptions (qui perdent du coup en lisibilité) et à l'exportation des exemples. Il semblerait que sur ce point ce ne soit pas TXM qui est en cause mais la structuration du corpus ESLO où les métadonnées, au lieu d'être attachées à chaque tour de parole, sont attachées aux entretiens comme ensembles. Certes, il est toujours possible de passer par

²²⁶ Par exemple, nous avons évoqué ci-dessus (note 200, p. 101) le logiciel Sonal (<http://www.sonal-info.com>) qui, s'il permet de travailler sur des transcriptions alignées sur le signal sonore, n'est pas, du moins pour l'instant, couplé à un étiqueteur morphosyntaxique. Mais les progrès dans le domaine sont rapides et la communauté des développeurs à l'écoute.

d'autres chemins pour garantir l'accès à toutes les données pertinentes, mais la tâche perd en productivité et en ergonomie (souvent, nous avons dû travailler sur deux versions parallèles du corpus : une version annotée et une version structurée en tours de parole).

Si une amélioration technique rendant accessible sur une même interface toutes les données nécessaires, du son aux lemmes étiquetés et situés syntagmatiquement et par rapport à leurs métadonnées, semble à portée de main, elle n'est envisageable que pour une prochaine étude, à moins qu'elle ne permette de récupérer le travail de correction et d'annotation déjà accompli. Le choix du prochain outil, au même titre que celui du prochain sous-corpus constitué (qui devra être mieux conçu en termes de « profilage » de ses composants et de ses locuteurs), ne pourra pas ne pas tenir compte du bilan de cette première expérience.

3.2.2.2.3. Annotation

Dans le sous-corpus constitué, 3370 occurrences de futur (FP et FS) ont été identifiées après correction. Elles se répartissent de la manière suivante :

(65)

	ESLO1	ESLO2	ESLO
FS	985	466	1451
FP	715	1204	1919
FS+FP	1700	1670	3370

L'annotation des occurrences a été menée en deux étapes successives, la première visant à identifier le type d'emploi et la seconde à sous-spécifier l'occurrence selon un certain nombre de traits, qui varient en partie en fonction du type d'emploi.

Lors de la première couche d'annotation, spécifiant le type de valeur sémantique exprimée, chaque occurrence reçoit une des trois valeurs possibles correspondant à l'une des trois classes d'emplois distinguées : en plus des deux emplois habituellement reconnus²²⁷, i.e. les

²²⁷ Encore que l'extension exacte de la classe des emplois temporels et celle des emplois modaux varie selon les auteurs. Pour nous, l'emploi temporel est un emploi qui situe le procès dans un moment ultérieur à un repère référentiel, qui est généralement le moment de l'énonciation, mais qui peut aussi correspondre à un repère distinct, comme c'est le cas pour le futur historique :

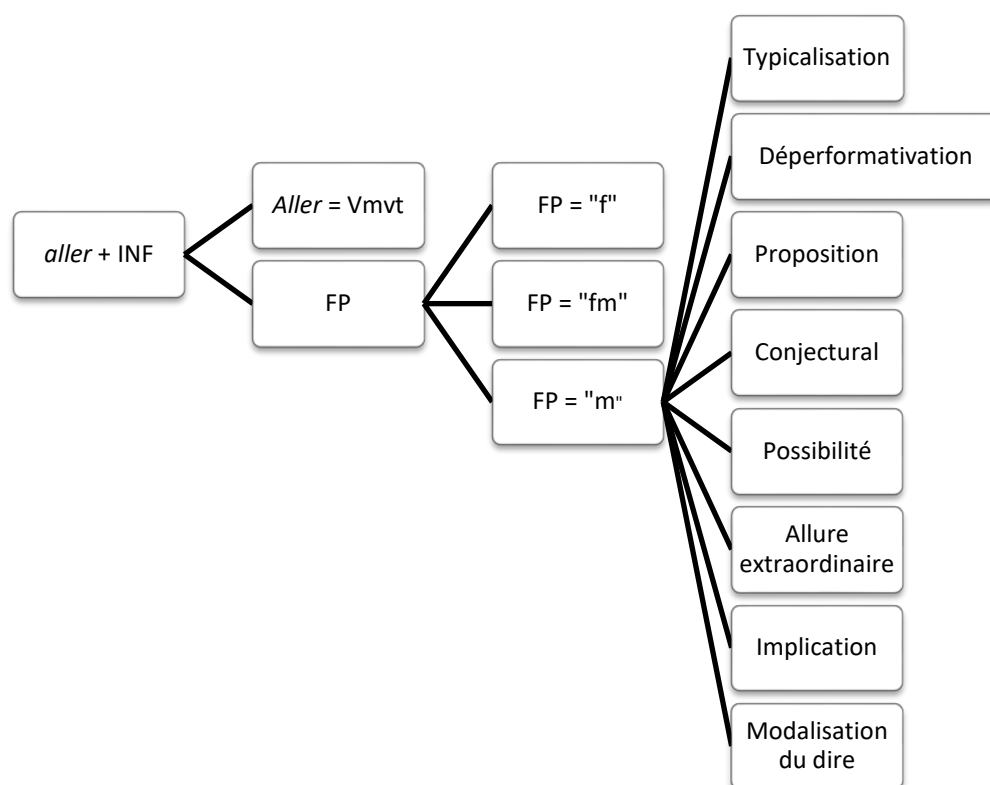
ESLO2_CONF_4CPMEb_C_11 :

le récit se clôt là où il avait commencé à la frontière franc- franco-espagnole en Hendaye l'endroit où le père de Jacques avait disparu et où Jacques mourra aussi incapable de dépasser les clivages linguistiques et culturels que sa mère avait vainement essayé de supprimer en lui

emplois modaux (étiquetés “m”), et les emplois temporels (étiquetés “f”), nous avons admis une troisième classe, mixte, d’emplois (étiquetés “fm”)²²⁸.

La deuxième étape d’annotation diffère selon le type d’emploi. Pour les emplois étiquetés « m », l’annotation (qui a été menée sur la totalité des emplois de ce type, soit 176 occurrences de FS et 455 occurrences de FP) a visé à identifier le type de modalité exprimée par chaque occurrence. L’ensemble de l’annotation jusqu’à cette étape peut être présentée (pour le FP, mais le FS ne s’en distingue que sur quelques détails) sous la forme arborescente suivante :

(66)



On trouvera dans Abouda & Skrovec (2015, à paraître) une présentation des valeurs modales les plus importantes (en nombre d’occurrences et/ou en pertinence en termes d’évolution micro-diachronique). Celles-ci, identifiées dans le cadre d’une approche qui

ou encore pour le futur relatif :

ESLO2_CONF_5FLb_C_3 :

Platon mais aussi d’autres évoquent le symposium grec où les convives se passent un cratère où dans lequel il y a de l’eau et du vin

et ils font l’éloge de l’amour mais pas seulement n’est-ce pas lors de ces banquets

et en quelque sorte ils se lancent des défis

et ils essayent de finalement de faire le l’éloge qui **sera** le le plus apprécié par tous hein

²²⁸ voir, pour une justification, Abouda & Skrovec (2015, à paraître).

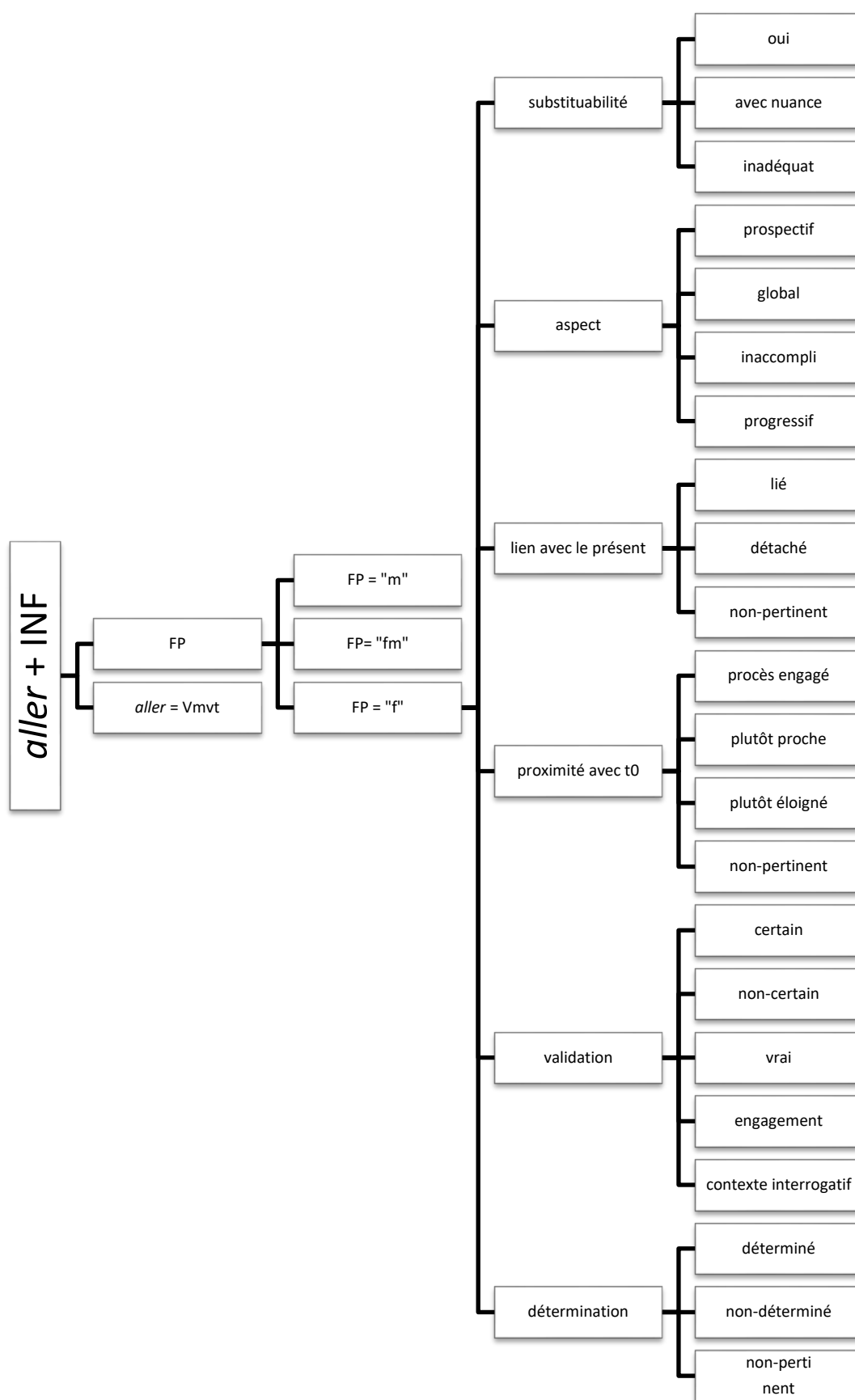
s'inspire de la méthodologie dite guidée par les données (*datadriven* pour Biber 2009), ne recoupe pas toujours les valeurs modales les plus habituellement reconnues.

Pour les emplois étiquetés « f », l'annotation n'a pu être menée à son terme que pour environ 30% des occurrences concernées (595 occurrences sur 2010), sélectionnées d'une manière qui garantisse le meilleur équilibre possible entre genres interactionnels (conférences, entretiens et repas)²²⁹. C'est que le nombre des étiquettes pour ce type d'emploi est important, et la détermination des valeurs particulièrement délicate, ce qui rend la tâche particulièrement longue.

L'architecture d'ensemble de cette annotation peut être présentée sous la forme arborescente suivante (pour le FP, mais, là non plus, le FS ne s'en distingue que sur certains points de détail) :

²²⁹ Mais l'objectif, à très court terme, est de mener cette annotation sur la totalité des occurrences : les résultats partiels ont été présentés dans une communication aux 8^e JLC en septembre 2015. Pour les deux communications prévues en novembre 2015 (voir CV, rubrique « Communications dans des colloques à comité de sélection » (2015b) et (2015c)), j'aimerais que la totalité des occurrences étiquetées « f » soit annotée.

(67)



Parce qu'elles n'ont pas encore fait l'objet d'une publication²³⁰, il me paraît indispensable de présenter brièvement les six étiquettes renseignées, en illustrant les principales valeurs par des exemples du corpus.

1. Substituabilité FP/FS

Il s'agit de voir si la forme du futur peut ou non être remplacée par l'autre forme dans le contexte occurrent. Trois cas semblent devoir être distingués.

(i) La substitution (FP-FS ou FS-FP) s'avère impossible²³¹

- FP non remplaçable par un FS :

(68) ESLO1_ENT_001_C_3

Loc1 : et alors qu'est-ce que vous faites exactement comme travail en ce moment ?

Loc2 : je suis boucher

Loc1 : oui oui oui et euh en quoi est-ce que ça consiste exactement ?

Loc2 : eh bien je vais vous dire euh comment vous voyez ça ? pour moi c'est un mon cas est peut-être un petit peu spécial pas parce que **je vais vous expliquer** euh j'ai une boucherie bien sûr euh qui m'appartient en propre mais en plus de ça depuis une dizaine d'années euh on ne peut pas dire le hasard mais enfin les relations ont fait que j'ai un de mes amis qui a ouvert deux supermarchés à Orléans et alors il m'a donné la gestion si vous voulez de ses deux rayons de boucherie de supermarché

- FS non remplaçable par un FP :

(69) ESLO1_CONF_503_B_23

Loc1 : je suis bien de votre avis plus on commencera tôt ce dressage et moins on pourra toujours on pourra de dispenser à ce niveau-là de chercher des motivations qui n'existent finalement pas oui n'est-ce pas hein c'est-à-dire que quand on **commencera** à apprendre l'anglais partout à huit ans eh bien vous o- vous organisez la actuel de ce qu'on appelle la la sixième qui avec qui vous pourrez chercher de vraies motivations

(ii) La substitution (FP-FS ou FS-FP) est possible avec une nuance de sens

- FP remplaçable par un FS avec glissement de sens :

(70) ESLO2_REP_14_C_5

Loc1 : en plus euh elle tu peux avoir super confiance hein puis elle est contente du coup [rire du locuteur] parce que ça ça lui évitera de la mettre sur Ebay et tout elle dit que ça la fait chier

Loc2 : eh je veux juste des informations hein pour le moment

Loc1 : hm mais ouais hm je **vais demander** euh à Pauline parce que hm j'ai vu la voiture que vite fait mais oui c' est une cinq portes c' est le euh hm deuxième modèle de Clio là celle qui est un petit peu arrondie là

²³⁰ La communication présentée aux JLC 2015 (voir note 230, p. 118) figure parmi les communications sélectionnées par le comité scientifique pour figurer dans un numéro spécial « Linguistique de corpus » qui paraîtra début 2016 dans la revue en ligne *CORELIA* (version écrite attendue pour le 15 novembre 2015...).

²³¹ Il s'agit, non pas d'une impossibilité agrammaticale, mais d'une inadéquation avec le contexte.

- FS remplaçable par un FP avec nuance :

(71) ESLO2_REPAS_22_C

Loc1 : non mais franchement non non **j'arriverai** à la vendre hm vu l'état qu'e- enfin c'est une petite voiture euh tout enfin tout est refait

- (iii) La substitution semble possible sans provoquer un changement de sens détectable

- FP remplaçable par un FS :

(72) ESLO1_ENT_009_C_27

Loc1 : ils sont encore dans les jupons de papa vous vous rendez compte ils n'ont pas encore de ce de métier alors ça a été jusqu'au bac ou je sais pas j'ai jamais osé demander si ça avait des examens ben alors qu'est-ce que ça fera ces jeunes-là quand les parents **vont plus être** là?

- FS remplaçable par un FP :

(73) ESLO2_REP_11_C_3

Loc1 : bah euh tu vas prendre du ketchup avec ton riz et ça **fera** de la sauce pour les deux

2. Aspect

Il a été ici question de déterminer la propriété aspectuelle de l'emploi dans son contexte. Trois valeurs principales ont été identifiées : prospectif, inaccompli (procès présenté en cours d'accomplissement et constituant souvent le cadre d'un autre procès, généralement ponctuel), et global (ou aoristique : le procès est perçu dans sa globalité).

- (i) Prospectif

Cette valeur n'a été identifiée que pour le FP :

(74) ESLO2_CONF_1BGb_C_1

Loc1 : et ces verriers vont émigrer dans toute l'Europe en apportant leur savoir-faire et ce savoir-faire qui va faire des allers-retours vu qu'on va le retrouver après en Italie parce qu'il y a bon y a tout un tas de contacts et je **vais parler** de l'un d'entre eux qui s'est établi à Orléans qui est donc Bernardo Perroto plus connu sous le nom de Bernard Perrot

- (ii) Inaccompli

Si le modèle de Gosselin²³² prévoit l'existence de FP inaccompli, nous n'avons trouvé, dans la partie déjà annoté de notre corpus, que des cas au FS :

(75) ESLO2_REP_01_02_C_11

Loc1 : quand il **sera** en troisième il sera dans la voix deux

²³² *A huit heures, il va (être en train de) manger* (Gosselin, 1996).

(iii) Global

- FP :

(76) ESLO1_ENT_003_C_17

Loc1 : dès qu'il va rentrer dans le primaire je **vais le mettre** à l'école euh publique

- FS :

(77) ESLO2_REP_11_C_1

Loc1 : bah mets-lui ça on lui en **remettra** après si elle en reveut

(iv) Progressif

A ces trois valeurs de base, on a identifié des emplois qu'on peut rapprocher de l'aspect progressif²³³, et qui ne se présentent qu'avec le FP :

(78) ESLO1_ENT_001_C_4

Loc1 : euh la vente sous cellophane malheureusement **va s'étendre** puisque c'est dans les traditions maintenant

3. Lien avec le présent

Ce trait cherche à vérifier l'une des hypothèses proposées dans la littérature selon laquelle le FP exprimerait un procès en lien avec le Présent. Pour annoter, il fallait définir avec précision la notion de *lien* qui semble avoir des acceptions relativement divergentes. Sera considéré ici comme lié au Présent tout procès qui se trouve soit déjà engagé (*Je vais rentrer déjà chez moi* (j'ai déjà fait quelques pas)), soit dans sa phase préparatoire (*Je vais rentrer chez moi* (je suis en train de ramasser mes affaires)), soit il s'agit d'un procès dont les conditions de réalisation se trouvent remplies dans t_0 (*C'est ce qu'on va manger ce soir* (décision déjà prise) ; *Elle va avoir un bébé* (elle est déjà enceinte)).

(i) Procès lié au Présent

- FP :

(79) ESLO1_ENT_008_C_9

Loc1 : comment est-ce qu'on fait une omelette chez vous ?

Loc2 : elle allait en faire une

Loc 1 : ah c'est pourquoi je vous pose la question

Loc 2 : elle allait en faire une justement c'est ce qu'on **va manger** ce soir

²³³ Une dizaine d'emplois qui se présentent avec des verbes comme *durer, s'étendre, se prolonger, s'éliminer, en arriver à, continuer...* Des vérifications plus poussées semblent nécessaires puisque ce type d'emploi n'a pas, me semble-t-il, été identifié en français (par exemple, Lansari (2009) n'en parle pas).

- FS

(80) ESLO2_CONF_5FLb_C_1

Loc1 : et puis euh j'ai bien compris que il se fait tard et que euh comme en géographie d'ailleurs on lie toujours l'utile à l'agréable et euh lorsqu'on est géographe ruraliste et quand en plus on travaille sur la vigne et le vin généralement ça se termine toujours non pas en chanson mais dans des caves et donc avec force dégustation c'est ce que nous **ferons** ce soir voilà donc la géographie est très heureuse de vous accueillir pour ces travaux pratiques

(ii) Procès détaché

- FS :

(81) ESLO1_CONF_503_B_23

Loc1 : je suis bien de votre avis plus on commencera tôt ce dressage et moins on pourra toujours on pourra de dispenser à ce niveau-là de chercher des motivations qui n'existent finalement pas oui n'est-ce pas hein c'est-à-dire que quand on **commencera** à apprendre l'anglais partout à huit ans eh bien vous o- vous organisez la actuel de ce qu'on appelle la la sixième qui avec qui vous pourrez chercher de vraies motivations

- FP :

(82) ESLO1_ENT_009_C_27

Loc1 : ben alors qu'est-ce que ça fera ces jeunes-là quand les parents **vont plus être** là ?

4. Proximité

Il s'agit d'un trait relativement proche de la notion de *lien*, mais qui a été distingué en ce qu'il cherche à mesurer la distance, strictement chronologique, plus ou moins importante qui sépare le procès au futur de t_0 . Là aussi, on reconnaît une hypothèse classique qui cherche à expliquer l'opposition FP/FS en termes de proximité avec la situation d'énonciation, qu'il est ici question de vérifier.

(i) Procès engagé

Ce cas de figure n'a été identifié que pour le FP :

(83) ESLO2_ENT_1010_C_16

LOC1 : donc je me suis dit ah j'ai peut-être pas choisi le le bon secteur

LOC2 : hm hm le bon filon

LOC 1 : et en fait c'est vraiment ça quoi c'est à dire que là au bout au bout de quatre ans on est plus que seize mille

LOC2 : en quatre ans ça a été d'accord ouais ouais ouais donc s-

LOC 1: seize mille et euh c'est que l- voilà hein donc on est dans l'ère où ça **va** encore **continuer** de de réduire

(ii) Procès plutôt proche

- FP :

(84) ESLO2_ENT_1010_C_17

Loc1 : en fait le plan **va avoir** lieu là dans peu de temps donc ça sera trop tard

- FS :

(85) ESLO2_CONF_1074_C_1

Loc1 : l'après-midi va va durer euh un certain temps hein puisque nous avons prévu euh cinq euh interventions euh vous avez la possibilité euh bien sûr d'aller prendre un café euh des petits gâteaux pour vous restaurer et puis vous sortir un petit peu bon hm voilà entre deux entre deux conférences on va dire euh on **fera** pas vraiment de pauses euh complètes qui risquent de faire dériver le le le temps

(iii) Procès plutôt éloigné

- FP :

(86) ESLO1_CONF_504_A_22

Loc1 : vous savez que ces programmes dans l'idée de la commission sont quadriennaux tous les quatre ans ça ne veut pas dire que dans quatre ans on **va déchirer** tout ça et puis inventer d'autres choses n'est-ce pas

- FS :

(87) ESLO1_REPAS_272_B_18

Loc1 : ils ont eu une chance en soixante-huit euh ça se **reproduira** pas avant avant vingt ou trente ans

5. Validation

Il n'a pas été ici question de distinguer toutes les valeurs modales épistémiques (possible, probable, certain, contingent...), mais uniquement deux valeurs essentielles : certain et non-certain, afin de vérifier l'une des hypothèses quelquefois proposées selon laquelle le FS exprimerait un procès moins certain que le FP. A ces deux valeurs essentielles, trois autres valeurs ont dû être ajoutées : non-pertinent, vrai, et engagement (direction d'ajustement allant de l'énoncé au monde, cf. Gosselin 2010)

(i) Procès présenté comme certain

- FP :

(88) ESLO1_ENT_003

JR: et est-ce que vous avez des projets pour les vacances d'été ?

DJ: ah pour les vacances d'été **je vais** aller euh **je vais** aller comme tous les ans au Guilvinec qui est un petit port euh breton où y a beaucoup plus de d'autochtones que de touristes

- FS :

(89) ESLO2_CONF_4CPMEb_C_8

Loc1 : euh ce roman a reçu le sixième prix Vargas Llosa de roman en Espagne il **sera** publié à la rentrée prochaine

(ii) Procès présenté comme non-certain

- FS :

(90) ESLO2_ENT_1010

ch_CD 2: à voir quoi il est prévu pour quand ?

GK: nous on l- il est prévu pour deux mille treize je crois

ch_CD d'accord oui c'est [inintelligible]

GK [inspiration] mise en service donc y a encore euh trois trois bonnes années de [bruit de bouche] d'ennui [rire du locuteur] pour se garer

ch_CD 2: 0:08:44 [rire du locuteur] c'est ça c'est ça

GK 11: 0:08:46 mais euh sinon euh ouais mais je pense qu'on **sera** plus là quand le tram euh lui sera là

- FP :

(91) ESLO2_REP_11_C_3

Loc1: je vais te le couper

Loc2: je crois pas que je **vais manger** les deux tranches

Loc1: mais alors ah bah si ça fait trop je peux en enlever là hein

(iii) Procès présenté comme vrai (futur historique)

- FP :

(92) ESLO2_CONF_1BGb_C_5

Loc1 : Bernard Perrot c'est un Orléanais il **va laisser** des tra- un certain nombre de traces dans Orléans dont les traces les plus importantes sont dans les roses des cathédrales n- nord euh des transepts nord et sud de la cathédrale d'Orléans comme euh nous avons pu le montrer récemment

- FS :

(93) ESLO2_CONF_4CPMEb

Loc1 : le récit se clôt là où il avait commencé à la frontière franc- franco-espagnole en Hendaye l'endroit où le père de Jacques avait disparu et où Jacques **mourra** aussi incapable de dépasser les clivages linguistiques et culturels que sa mère avait vainement essayé de supprimer en lui

(iv) Engagement relatif à la réalisation du procès

- FP :

(94) ESLO2_REP_11_C_2

Loc1: tu te souviens dans la voiture je te disais j'avais soif

Loc2: ah oui c'est vrai je **vais** t'en **donner** tout de suite moi aussi j'ai soif

- FS :

(95) ESLO1_CONF_503_B_2

Loc1 : on serait tenté de croire que l'apprentissage d'une langue étrangère c'est quelque chose d'assez semblable d'assez allogue euh d'assez parallèle à l'apprentissage de la langue maternelle euh vous me direz ce que vous en pensez pour ma part je vous le **dirai** aussi mais enfin euh là-dessus la discussion reste ouverte euh

6. Détermination

Avec cette étiquette, il est question de savoir si le procès est présenté comme déterminé (occurent) ou indéterminé (plus virtuel). En plus des cas de non-pertinence, deux valeurs sont distinguées (procès déterminé, ou non déterminé).

(i) Procès présenté comme non-déterminé

- FS :

(96) ESLO1_ENT_009 :

G258: on pense pas bon ben c'est ma taille on dit oui ou si on leur dit si ça ne va pas vous **reviendrez**

- FP :

(97) ESLO1_ENT_009_C_28

Loc1 : et sa fille ben sa fille elle a carrément dit à son père dis-donc moi c'est trop tenu je vais pas faire dactylo je veux mon temps je veux me promener bon ben la fille maintenant elle est elle est toujours célibataire et elle compte trop sur papa maman c'est pas ça qui **va** nous **sauver**

(ii) Procès présenté comme déterminé

- FS :

(98) ESLO2_CONF_5FLb_C_1

Loc1 : et puis euh j'ai bien compris que il se fait tard et que euh comme en géographie d'ailleurs on lie toujours l'utile à l'agréable et euh lorsqu'on est géographe ruraliste et quand en plus on travaille sur la vigne et le vin généralement ça se termine toujours non pas en chanson mais dans des caves

et donc avec force dégustation c'est ce que nous **ferons** ce soir voilà donc la géographie est très heureuse de vous accueillir pour ces travaux pratiques

- FP :

(99) ESLO1_ENT_003_C_17

Loc1 : dès qu'il va rentrer dans le primaire je **vais** le **mettre** à l'école euh publique publique

**** **

A l'examen de ce parcours rapide des principales catégories d'annotation pour les emplois temporels du futur (FP et FS, étiquetés « f »), on constate que deux points, au moins, méritent d'être discutés.

D'abord, les catégories semblent en partie présenter sinon une certaine redondance, du moins une certaine intersection : c'est notamment le cas, d'une part, des traits de proximité et de lien, et, de l'autre, des traits de détermination et de validation, voire entre ces couples de traits. Cela constitue une réponse au fait que ces étiquettes ne sont généralement pas (ou peu) définies et s'accommodent aisément d'acceptions divergentes, et plus ou moins métaphoriques : il est par exemple classique, y compris au sein d'un même travail, que la notion de proximité vacille entre une acception strictement chronologiques (dans la représentation du Temps objectif) et une acception psychologisante (proximité dans le Temps psychologique), ce qui fait glisser la notion de proximité vers celle de lien, voire vers celles de détermination et de validation²³⁴. L'annotation d'énoncés oraux authentiques, s'inspirant en l'occurrence d'une démarche dite guidée par les données, obligent, dans cet instant strictement descriptif, à distinguer ces différents traits (ou ces différentes facettes d'un trait plus général mais qui ne pourrait être identifié que dans une phase ultérieure), là où une démarche par papillonnage finira, en écartant des données, par trouver les illustrations recherchées.

²³⁴ On trouvera chez Maingueneau (1981 : 75) un exemple explicite de ce type de traitement : «[le] FP met le procès dans le prolongement immédiat du présent d'énonciation, alors que le FS apparaît plus "coupé" de ce présent, s'applique mieux aux procès éloignés dans le temps. Mais cette différence de "distance" qu'impliquent ces deux futurs va au-delà du simple éloignement temporel et doit également se lire comme une prise en charge plus ou moins forte de son énoncé par l'énonciateur : le FS est plutôt associé aux circonstances solennelles, aux engagements, prophéties, rêves, etc., tandis que le FP fonctionne plutôt dans ce qui relève du quotidien, des intentions inscrites dans une réalité proche. Cela explique l'emploi privilégié du FP dans la conversation courante ». Voir également, pour une autre illustration de ce type de démarche, Confais (1990 : 279 sq.).

L'autre difficulté inhérente à ce type de démarche consiste à distinguer ce qui relève du contexte de ce qui relève du marqueur. Ce point, déjà discuté ci-dessus (§ 2.3.2.3), est sans conteste le plus difficile : comment parvenir à distinguer ce qui relève du contexte, de ce qui relève du sens holiste de l'énoncé, et de ce qu'on laissera au marqueur en guise de sens contextuel spécifique ?

Confrontée à chaque instant à cette question²³⁵, l'annotation, aussi difficile et délicate, me paraît nécessaire et désormais plus accessible, dans le cadre de corpus numériques et outillés et sur des données situées.

Si en effet le risque est réel de confondre ce qui relève du marqueur de ce qui relève du contexte, et même si la solution n'est pas simple, la démarche suivie s'est entourée d'un certain nombre de précautions. D'abord, la double annotation à laquelle nous avons procédé limite l'effet de l'intuition purement individuelle. Ensuite, les catégories utilisées, même s'il n'a pas été question de faire totale abstraction de ce qui est proposé dans le domaine de la temporalité, ont été conçues comme des catégories « de travail », qui émergent des données, et non des catégories préalables que l'adoption d'un modèle aurait théoriquement privilégiées²³⁶. Dans cette démarche d'essai et de découverte, il n'a pas été rare de redéfinir certaines catégories et d'en abandonner d'autres si leur pertinence se trouve déstabilisée face aux données. Les étapes les plus délicates ont été discutées, et l'adoption de solutions a été argumentée avec à l'appui l'usage des procédures habituelles du travail linguistique (usage des tests de suppression de certains marqueurs cotextuels, de substitution du marqueur par une autre forme, soit plus attendue (un FS par un présent), soit synonymique (*on va dire* par un conditionnel), etc.). Si de telles précautions ne garantissent pas l'exactitude des analyses, le risque ne me paraît pas plus élevé que dans le cadre d'une démarche hypothético-déductive où on est amené tout de même à trier et hiérarchiser des emplois, à identifier les marqueurs responsables de tel ou tel effets de sens, etc., décisions descriptives qui reviennent à isoler ce qui relève du marqueur.

Pas plus risquée qu'une démarche hypothético-déductive, l'annotation semble offrir par rapport à cette dernière l'avantage de ne pas forcer les emplois réels à entrer en conformité avec les emplois attendus, fussent-ils déviants. Conjuguée à la fabrication d'exemples ou au

²³⁵ Voir Ducrot (1984 : 50) et Gosselin (1996 : 188) déjà cités ci-dessus.

²³⁶ Les étiquettes retenues correspondent quelquefois à des étiquettes déjà proposées (atténuation, allure...), et quelquefois à des étiquettes forgées par redéfinition et retraçage d'extension (élargissement ou rétrécissement), comme c'est le cas de la modalisation du dire par un verbe de parole, le déperformatif, et la typicalisation...

papillonnage sur corpus, une pure démarche hypothético-déductive risque de ne trouver que les résultats qu'elle a préalablement injectés (les emplois listés et décrits depuis longtemps par la grammaire).

Ce travail de forcing aurait eu différentes conséquences, et notamment un appauvrissement et une simplification des occurrences, qui se concrétisent de différentes manières :

- a. proposer, notamment dans le cadre des traits binaires (par exemple aoristique et prospectif), un trait simplement parce qu'il est moins pire que l'autre ;
- b. proposer un trait parce qu'il ne paraît pas incompatible avec le sens de l'occurrence, mais sans certitude concernant son statut définitoire ;
- c. sacrifier une ou plusieurs facettes sémantiques de l'occurrence : dans notre démarche, les données nous ont souvent incités à reconnaître des catégories mixtes (par exemple « fm ») ou des emplois mixtes (par exemple, certains emplois modaux ont reçu plus d'une valeur de sous-spécification : typicalisation et cadre, etc.), voire des occurrences silencieuses concernant un trait pertinent pour d'autres occurrences.

Il convient enfin de préciser que ces recherches sur le futur n'ont pas eu pour objectif d'identifier la valeur en langue d'un marqueur donné (en l'occurrence le FP et le FS) en partant des effets de sens qui se dégagent des occurrences observées. C'est qu'une démarche purement inductive de ce type, en plus des difficultés signalées, viderait le marqueur en langue de sa substance sémantique. Si une démarche hypothético-déductive paraît ainsi mieux placée pour une telle entreprise, les hypothèses qu'elle avance peuvent trouver dans les corpus documentés, situés et annotés un terrain privilégié de falsification.

Tout cela revient à dire que la méthode hypothético-déductive, tout aussi nécessaire et légitime, peut ainsi être complétée, à l'autre bout de la chaîne, par un retour aux données.

Conclusion générale

J'ai cherché dans cette synthèse à restituer les étapes qui me paraissent emblématiques de ma trajectoire de chercheur, en les illustrant par des analyses et des positions théoriques ayant fait l'objet de travaux significatifs.

Dans la première partie, en partant de la question classique du statut catégoriel du conditionnel – temps ou mode –, j'ai brièvement rappelé ma contribution à ce débat, dans un cadre syntaxique. Par l'examen des différentes structures syntaxiques incluant un conditionnel, il s'agissait de vérifier, au moyen d'un test de substitution, si l'on était confronté à une structure exclusive ou d'une structure partagée avec l'un ou l'autre des deux autres modes personnels, i.e. indicatif et subjonctif.

Les quatre cas de figure identifiés ont été présentés (§ 1.1.2) et discutés afin de reconnaître les paramètres et leur nature (syntaxique ou non) qui interfèrent dans la sélection de cette forme. A l'issue de cette discussion, il semble se confirmer que le conditionnel, partageant toutes ses structures avec (au moins) une forme de l'indicatif, doit être considéré comme un temps de l'indicatif.

Le reprise de cette question a permis parallèlement de revenir sur une hypothèse, plus générale, celle de la visibilité syntaxique des modes (§ 1.2). Elle part d'un constat simple, qui se conçoit mieux par la comparaison avec les phénomènes temporels : contrairement aux formes temporelles, qui semblent constituer des objets logico-sémantiques invisibles pour la syntaxe, les modes verbaux, obéissant à des principes syntaxiques, ont une visibilité structurelle qui les rend à la fois descriptibles et prédictibles en syntaxe. La démonstration commençait par une distinction de domaine, la phrase pour la sélection modale et, pour la sélection temporelle, l'énoncé et jusqu'au texte. Il a été établi que les chemins syntaxiques qu'empruntent les traits d'héritage de mode relèvent de la domination stricte, là où les chemins suivis par les traits de percolation de temps se conforment à des contraintes extra-phrastiques ou à des chemins qui relèvent de l'adjonction. La distinction du temps et du mode s'éclaire également dans l'opposition entre impossible modal, impossible grammatical, et inacceptable temporel, résultant de l'inadéquation entre une forme temporelle et la coprésence distributionnelle ou situationnelle d'un élément porteur d'instructions sémantiques contradictoires, que des procédures de réinterprétation peuvent tenter de sauver. La concordance des temps, qui semble contredire cette vision, a été discutée (§ 1.2.2.3) avant d'être abandonnée au profit de l'hypothèse de l'ancrage référentiel, qui, en inversant le sens de la relation (ce n'est pas le verbe principal qui transmet des traits temporels au verbe de la

subordonnée mais le verbe subordonné qui, parce qu'il a besoin d'un repère référentiel, peut le récupérer dans la principale), semble mieux expliquer les nombreuses transgressions observées et le fait que le phénomène ne concerne qu'une partie des temps qui apparaissent dans la subordonnée, ceux, justement, qu'on appellera polychroniques. Indéterminés, ils recherchent le repère qu'ils n'ont pas dans t_0 et qu'ils peuvent éventuellement trouver dans la principale.

Dans la deuxième partie, il a été question, après quelques préliminaires destinés à préciser une série de choix théoriques, d'esquisser une vision de l'indicatif.

Si, comme l'écrivait Saussure (1916 : 162), la « plus exacte caractéristique [d'un signe] est d'être ce que les autres ne sont pas », la détermination du statut catégoriel du conditionnel n'est possible que dans la restitution du réseau d'oppositions dans lequel se déterminent différenciellement ses propriétés. Celles-ci doivent d'une part l'inscrire dans un paradigme, qui s'oppose lui-même à d'autres paradigmes, et, d'autre part, le différencier de tous les autres marqueurs du même paradigme.

L'appartenance catégorielle du conditionnel au mode indicatif prélude ainsi à de nouveaux questionnements, inépuisables, concernant le système de l'indicatif et les réseaux d'oppositions entre les différentes formes temporelles qui le composent.

Cette synthèse ne propose pas une vision exhaustive du mode indicatif. Quelques propositions ont été avancées qui s'inscrivent dans le cadre d'une longue tradition, celle d'une bipartition (§ 2.2) entre deux groupes de formes temporelles (encore que cette bipartition ait été considérée de différentes façons). Les considérations morphologiques ont été mises au premier plan. Par l'examen de la structure morphologique du conditionnel, je me suis démarqué d'une conception structuraliste qui ferait de cette forme verbale l'unique cas d'une coexistence de deux morphèmes temporels. Une alternative revenait à faire l'hypothèse que la structure du conditionnel n'a de spécifique que le double marquage phonologique de deux morphèmes qui coexistent dans tout le paradigme de l'indicatif. La catégorie traditionnelle du « temps » devient une catégorie générique, amalgamant systématiquement deux catégories distinctes, selon le schéma morphologique (simplifié, puisque l'on écarte pour la clarté de la démonstration, le Mode et l'Aspect) :

$$(100) \quad \text{Radical} + X + Y + \text{PN}$$

A partir de ce schéma, a été esquissée l'une des lectures possibles qui paraît intéressante sur les plans descriptif et explicatif. Une telle interprétation implique des décisions théoriques, qui vont du statut du morphème zéro, à des hypothèses concernant l'opposition Présent/Imparfait (§ 2.2.2.2). Décrite de façon extensive en termes de décrochage du repérage par rapport à <Moi, Ici, Maintenant>, comme procédant de nombreux traitements classiques et actuels, cette solution n'est ni une esquivé, ni une annulation du signifié de l'imparfait ; c'est une manière de prendre acte d'une sous-détermination dans la langue elle-même. C'est dans son contexte d'emploi, et il en a toujours un, que l'imparfait se décline en lecture atténuative, hypothétique, ludique... ou chronologique. Le traitement de ce dernier emploi, qui semble constituer un point d'achoppement pour les approches non-temporalistes, a été étayé sur une hypothèse empiriquement vérifiable : dans son emploi temporel, l'imparfait possède systématiquement un antécédent temporel, qui est fourni par le co-texte ou le contexte.

Ce n'est pas la seule hypothèse qui doive être vérifiée empiriquement par l'examen de données pertinentes. Parvenu au terme d'une organisation paradigmatique hiérarchisée du système indicatif, le traitement doit encore affronter une difficulté récurrente : la polysémie temporelle.

L'hypothèse toute théorique de la polychronie (§ 2.3.2), qui semble à même de fournir un cadre adapté au traitement de certains types de polysémie temporelle, peut être justifiée et argumentée par le recours à des données authentiques disponibles en masse (*Big Data*). Ce point est abordé dans la dernière partie.

La permanence de la thématique de la temporalité dans mes recherches ne s'explique pas uniquement par un intérêt personnel, dont j'ignore sans doute les véritables raisons, mais également par l'objet lui-même qui offre un vaste champ de recherche et lance des défis multiples et complexes. Lorsque, par lassitude, il me prend de vouloir changer de domaine, c'est un autre objet temporel, ou le même objet, envisagé d'un autre angle de vue, qui se présentent. S'il y a un désir qui se fixe sur d'autres objets (par exemple, une étude sur le parler tunisien, une enquête sociolinguistique sur les juifs tunisiens en France et en Tunisie, sur le maltais...), les questions liées à la temporalité restent au centre de mes perspectives de recherche.

Avec l'accès facilité par les outils du TAL à des masses de données authentiques, un nouvel horizon s'ouvre qui renouvelle le débat scientifique. En effet, pour des raisons à la fois

historiques, sociales et épistémologiques, liées au prestige des emplois scripturaux et à la disponibilité des matériaux, les références à l'écrit ont longtemps prévalu dans les travaux conduits en sciences du langage. Au français parlé, lieu par excellence de la variation diatopique, diastratique et diaphasique, ont été substituées des attestations scripturales standardisées, souvent littéraires, dans le cadre de travaux qui perpétuaient une tradition qui se démarquait peu de la philologie. L'accès à des données orales authentiques remodèle notre conception des faits.

La linguistique *en corpus* pour laquelle plaide cette synthèse voit dans l'avènement des humanités numériques une chance pour un nouvel essor de la linguistique, sans renier les acquis d'une science cumulative. S'il serait dommage de ne voir dans la croissance exponentielle de données orales accessibles qu'une simple opportunité d'« élargissement du répertoire des exemples à citer ou à inclure dans les démonstrations »²³⁷, d'autres démarches revendiquant un rapport différent aux données existent, qui ont montré leur efficacité.

D'où la position relativiste, mais non paradoxale, défendue concernant le fait linguistique. Si les travaux sur des données orales authentiques, documentées et quantitativement significatives sont de nature à renouveler profondément le débat scientifique, le fait linguistique, du morphème à l'énoncé, déborde l'attesté. Toujours construit, ne serait-ce que parce qu'il est saisi, la frontière du fait linguistique s'étend de l'existant (qui peut être, ainsi que nous l'avons vu, impossible) au possible (qui est peut-être inexistant).

²³⁷ Descriptif de la journée d'étude organisée le 10 avril 2015 par Paul Cappeau « Une grammaire sous influence de l'oral ? »

Annexe 1 : Corpus futur

Sélection des locuteurs des entretiens d'ESLO1

Âge 15-35					
Enregistrements	Durée	Code locuteur	Sexe	Année de naissance	Échelle AM
121	125	RF 211	H	1947	A
029	89	1268	F	1948	A
017	75	QB 100	F	1945	B
141	86	VU 469	H	1941	B
048	98	CF 4	H	1943	C
015	57	NS 571	F	1949	C
149	104	YR 399	H	1942	D
115	64	PY 94	F	1943	D
050	45	CI 58	F	1942	E
044	50	BF 69	H	1950	E
Total	793				

Âge 35-60					
Enregistrement	Durée	Code locuteur	Sexe	Année de naissance	Échelle AM
003	55	DJ39	H	1932	A
088	47	JX 318	F	1934	A
055	53	CW 739	F	1934	B
008	54	GJ 131	H	1921	B
150	43	YT 387	H	1928	C
065	53	ES 187	F	1920	C
009	91	JG 258	F	1923	D

001	60	BA725	H	1912	D
096	69	LN 438	F	1917	E
129	79	TS 165	H	1910	E
Total	604				

Âge + 60					
Enregistrements	Durée	Code locuteur	Sexe	Année de naissance	Échelle AM
058	75	DH59	H	1906	A
072	60	FZ 61	F	1907	A
022	47	WL 512	F	1898	B
052	50	CP7	H	1904	B
132	66	UN 412	F	1899	C
067	65	FH 717	F	1903	C
045	105	BP 35	H	1908	D
043	45	BD 704	H	1909	D
081	60	IC 346	F	1899	E
160	72	4003	H	1903	E
Total	645				
TOTAL entretiens eslo1	2042 min (34h)				

Sélection des locuteurs des entretiens d'ESLO2

Âge 15-35						
Enregistrements	Durée	Code locuteur	Sexe	Année de naissance	Cat. INSEE	Équivalence échelle AM
ENT_1010	58	GK 11	H	1979	Cadres/ profession intell sup	A
ENT_1071	46	HT398	F	1977	Cadres/ profession intell sup	A
ENT_1081	34	SF018	F	1979	Prof. interméd.	B
ENT_1041	72	BC41	H	1980	Prof. interméd.	B
ENT_1017	50	BT17	H	1982	Prof. interméd.	C
ENT_1044	58	DE44	F	1978	Prof. interméd.	C
ENT_1013	74	ZD13	H	1983	employés	D
ENT_1062	29	JP62	F	1977	Employés	D
ENT_1082	60	MX953	H	1981	employés	E
ENT_1001	69	BV1	H	1987	ouvriers	E
Total	550 min.					

Âge 35-60						
Enregistrement	Durée	Code locuteur	Sexe	Année de naissance	Cat. INSEE	Équivalence Échelle AM
ENT_1028	55	QF28	H	1952	Cadres/ profession intell. sup.	A gastro- entérologue
ENT_1056	56	TJ56	F	1974	Cadres/ profession intell. sup.	A psychiatre
ENT_1002	90	RL2	F	1964	Prof. intermédiaire	B bac+2 éducatrice

						spécialisée
ENT_1049	130	AU49	H	1964	employé	B Bac +1 enquêteur et gestionnaire d'enquête auprès des entreprises
ENT_1029	103	WC29	H	1956	Prof. intermédiaire	C équivalent bac technique (cinéma) CAP opérateur projectionniste marionnettiste réalisateur
ENT_1033	56	LK33	F	1966	Artisan, commerçant	C Bac Commerce alimentaire
ENT_1057	79	UZ67	F	1967	Employé	D Lycée incomplet Agent d'accueil
ENT_1025	65	LF25	H	1955	Employé	D secondaire incomplet/ boulangier grande surface
ENT_1046	78	WZ46	H	1952	ouvrier	E scolarité primaire
ENT_1085	56	RN488	F	1961	employée	E femme de ménage brevet des collèges
Total	768					

Âge + 60						
Enregistrements	Durée	Code locuteur	Sexe	Année de naissance	Cat. INSEE	Éq. échelle AM
ESLO2_DIA_003	54	DJ39	H	1932	retraité (médecin)	A
ESLO2_DIA_017	53	QB100	F	1945	Cadre hospitalier	A
ESLO2_DIA_121	47	RF211	H	1946	Professeur Formateur	A
ENT_1015	71	EW15	F	1937	responsable agence bancaire	B
ENT_1009	71	LX10	F	1945	retraîtée institutrice	B
ESLO2_DIA_150	65	YT387	H	1928	retraité	C
ESLO2_DIA_115	40	PY94	F	1943	Retraîtée Contrôleur PTT	C
ENT_1024	64	GC24	F	1925	retraîtée maraîchère	D
ENT_1023	72	YR23	F	1942	Retraité ouvrière usine, vendeuse épicerie, vêtements, ATSEM	D
ESLO2_DIA_149	99	YR399	H	1942	Ouvrier retraité	E
ESLO2_DIA_048	80	CF4	H	1943	ajusteur	E
Total	716					
TOTAL ENTRETIENS ESLO2	2034 (34h)					

Annexe 2 : Extrait d'un tableau CSV en cours d'annotation (avant réinjection sous TXM)

Capture d'écran 2 : annotation du FP

ContexteGauche	Pivot	frpos	fp	mvt	amb	fs	sub	aspect	lien	prox	val	det	scat
on me disait bah tu	vas mettre	VER:pres			amb								
faites une omelette alors je	vais vous donner	VER:pres	f				inad	prospectif	lie	plus	eng	plus	
surprenant comme question bah je	vais vous la détailler	VER:pres	f				inad	prospectif	lie	plus	eng	plus	
aime pas ça que je	vais faire	VER:pres	m										typical
pas mal c'est ça	va me je vais pouvoir	VER:pres	f				oui	global	det	non	non	moins	
si j'ai envie je	vais pouvoir	VER:pres	f				oui	global	det	non	non	moins	
et puis tout le quartier	va être	VER:pres	f				nua	global	lie	non	cert	non	
de traverser la Loire ça	va pas être	VER:pres	fm										
une école maternelle bon ça	va revenir	VER:pres	f				nua	global	lie	plus	eng	plus	
question qu'est-ce que tu	vas faire	VER:pres	f				inad	global	det	non	interr	moins	
dit bon pourquoi pas on	va essayer	VER:pres	f				inad	global	lie	plus	eng	moins	
donc euh donc je je	vais en Italie euh hm voir	VER:pres		mvt									
je sais pas comment ça	va évoluer	VER:pres	f				inad	global	lie	non	interr	moins	
les ficelles faciles euh qui	vont fonctionner	VER:pres	f				inad	global	det	non	cert	non	
type de culture qui r-	va rester	VER:pres	f				oui	global	lie	non	cert	non	
votre accord donc hein ça	va je vous dis euh constituer	VER:pres	f				inad	global	lie	non	cert	non	
hein hein et et on	va utiliser	VER:pres	f				nua	global	det	non	cert	non	
d'ailleurs bah écoutez je	vais je vous r- je vais arrêter	VER:pres	f				inad	prospectif	lie	engag	non	plus	
ans donc je suis on	va dire	VER:pres	d										
attache pour eux ça ça	va être	VER:pres	m										typical
Vignat rue Eugène Vignat ça	va faire	VER:pres	f				inad	global	lie	plus	non	moins	
les gens après l'entraînement	vont pas euh boire	VER:pres		mvt									
tram ouais en fait ça	va tout changer	VER:pres	f				inad	global	lie	plus	cert	moins	
sais pas ce que ça	va changer	VER:pres	f				nua	global	lie	plus	interr	moins	
vraiment je sais qu'il	va peut-être y avoir	VER:pres	f				oui	global	det	plus	N-cert	moins	
même moins de voitures ils	vont réduire	VER:pres	f				inad	global	lie	plus	cert	moins	
moi des produits que je	vais présenter	VER:pres		mvt									
à chaque fois que je	vais voir	VER:pres		mvt									
donc y ré- voilà on	va dire	VER:pres	d										

...

Annexe 3 : Exemple de requête sous TXM (après réinjection de l'annotation)

Capture d'écran 1 : résultats d'une requête sous TXM (croisant pour l'emploi temporel du FP valeur aspectuelle et validation)

The screenshot shows the TXM software interface. On the left is a file explorer with a tree view containing various corpora like GRAAL, VOEUX, ATELIER TXM, CONDANOTE, CORPUSANNOTÉS JUIN 2014, CORPUS FUTUR, CORPUS FUTUR 170215, CORPUS FUTUR ANNOTE, CORPUS FUTUR ANNOTE 1106, CORPUS FUTUR ANNOTE FPF, CORPUS TEST, DISCOURS, ESSAI 2, FUTUR ANNOTE MAI 2014, FUTUR CORPUS ORIGINAL, and FUTUR TEMPS ASPECT 270815. The main window displays the search results for the query `[fp="f"&aspect="prospectif"&val="eng"]`. The results are shown in a table with columns: text_id, Contexte gauche, Pivot, and Contexte droit. The table contains 20 rows of results, each showing a text ID, a left context, a pivot word, and a right context.

text_id	Contexte gauche	Pivot	Contexte droit
ESLO1_CONF_503_B	et auditifs d'autre part autrement dit nous	allons	nous demander je je me suis demandé euh et je crois c
ESLO1_CONF_504_A	évidemment leur donner une de notions logiques vous	allez	voir que c'est précisé cette idée-là dans le B O donnai
ESLO1_CONF_504_A	avez vu à la fin des commentaires je	vais	vous le montrer euh tel est le processus prévu par le pr
ESLO1_ENT_001_C	un petit peu spécial pas parce que je	vais	vous expliquer euh j'ai une boucherie bien sûr euh qui
ESLO1_ENT_001_C	n'est-ce pas ? normale eh bien je	vais	vous expliquer ce que je ressens et ce que je pense qu
ESLO1_ENT_001_C	tuer il est comprenez ? euh comment je	vais	expliquer ça je voudrais pourtant bien expliquer ça eul
ESLO1_ENT_003_C	haut vous tracassez pas bon merci bah je	vais	essayer de quand même de vous retenir le moins possi
ESLO1_ENT_003_C	y a certaines questions plus précises que je	vais	oui essayer de de en vitesse oui même très précises trè
ESLO1_ENT_008_C	ah bon va chercher le ah bon je	vais	vous montrer va va ben tu vas chercher le hm hm le
ESLO1_ENT_008_C	hm hm le bloc là hm hm je	vais	vous montrer bon comme ça vous saurez mieux et qua
ESLO1_ENT_009_C	répond oui elle est belle attends moi je	vais	te la faire tenir debout je lui dis non tu peux pas
ESLO1_ENT_015_C	bon alors je je je recommence et je	vais	essayer de débrouiller cette question d'accord il y a to
ESLO1_ENT_015_C	français ? oui oui ensuite euh bon je	vais	tout de même fermer la machine parce que
ESLO1_REPAS_272_B	ah bein en effet bon bein écoutez je	vais	y aller hein oui d'accord oui bein je je lui demanderais
ESLO1_REPAS_272_B	pas tout le fromage annie bon bein je	vais	préparer le dessert mais non un petit peu
ESLO1_REPAS_273_B	vierge sont beaucoup plus foncés Régis Marie Agnès	va	vous donner un morceau de pain si vous avez faim et p
ESLO1_REPAS_274_B	il a pas de femme ? Richard je	vais	vous faire voir grosso modo arrête arrête arrête
ESLO1_REPAS_275_B	a plus de vapeur bon ben attends je	vais	l'ouvrir charbon de bois oh la cuisinière que je fais bor
ESLO1_REPAS_275_B	ça saute hm tu piges ben oui je	vais	te faire de la place alors ça j'ai ma poubelle elle
ESLO2_CONF_1074_C	présence et euh sans plus attendre donc je	vais	laisser la place à nos intervenants alors euh donc euh p
ESLO2_CONF_1BGB_C	a tout un tas de contacts et je	vais	parler de l'un d'entre eux qui s'est établi à
ESLO2_CONF_1BGB_C	siècle alors les les premières études que je	vais	montrer c'est une étude qui a été faite y a deux
ESLO2_CONF_1BGB_C	de choses euh de ce genre-là là on	va	rester vraiment à la physique et à la chimie on va laisse

Bibliographie

- ABEILLE, A. (1993). *Les Nouvelles syntaxes Grammaire d'Unification et analyse du français*. Paris, Armand Colin, col. 'Linguistique'.
- ABNEY, S. (1996). Statistical Methods and Linguistics, in J. Klavans & Ph. Resnik (eds), *The Balancing Act*, Cambridge, The MIT Press, 1-26.
- ADAM, J.-M. (1992). *Les textes : types et prototypes. Récit, description, argumentation, explication et dialogue*, Paris, Nathan.
- ADAM, J.-M. (2001). Types de textes ou genres de discours ? Comment classer les textes qui disent de et comment faire ?, *Langages*, 141, 10-27.
- ANSCOMBRE, J.-C. (2004). Quelques remarques sur l'existence et le fonctionnement d'un *si* concessif en français contemporain, in M. L. Donaire (éd), *Dynamiques concessives*, Madrid, Arrecife, 41-74.
- ANSCOMBRE, J.-C. (2005). Le ON-locuteur : une entité aux multiples visages, in J. Bres, P.P. Haillet, S. Mellet, H. Nolke, L. Rosier (éds.), *Dialogisme et polyphonie, Approches linguistiques*, Bruxelles, De Boeck-Duculot, 75-94.
- ANSCOMBRE, J.-C. (2006a). Stéréotypes, gnomie et polyphonie : la voix de son maître, in L. Perrin (éd.), *Le sens et ses voix, Dialogisme et polyphonie en langue et en discours*, Metz, Université Paul Verlaine, *Recherches linguistiques*, 28, 349-378.
- ANSCOMBRE, J.-C. (2006b). Polyphonie et classification des énoncés sentencieux, *Le français moderne*, 1, 87-99.
- ANSCOMBRE, J.-C. (2009). La comédie de la polyphonie et ses personnages, *Langue française*, 164, 11-31.
- ARNAULD, A. & LANCELOT, Cl. (1660). *Grammaire générale et raisonnée de Port-Royal*, Paris, Slatkine. éd. 1972.
- ARRIVE, M ; GADET, F. & GALMICHE, M. (1986). *La Grammaire d'aujourd'hui*, Paris, Flammarion.
- AUROUX S. (1998). *La raison, le langage et les normes*, Paris, PUF.
- AUTHIER-REVUZ, J. (1982). Hétérogénéité montrée et hétérogénéité constitutive : éléments pour une approche de l'autre dans le discours, *DRLAV*, 26, 91-151.
- AUTHIER-REVUZ J. (1995). *Ces mots qui ne vont pas de soi*, Paris, Larousse (2 vol.).
- AUTHIER-REVUZ, J. (1996). Hétérogénéités en discours : remarques sur la catégorie de l' « îlot textuel », *Cahiers du français contemporain*, 3, 91-115.
- AZZOPARDI, S. (2011). *Le futur et le conditionnel : valeur en langue et effets de sens en discours. Analyse contrastive espagnol/français*, Thèse de doctorat, Université Paul-Valéry – Montpellier 3.
- BAKHTINE, M. (1977). *Le Marxisme et la philosophie du langage*, Paris, Minuit.
- BAKHTINE, M. (1978). *Esthétique et théorie du roman*, Paris, Gallimard.
- BAEYENS, L. (2012): Le conditionnel épistémique dans les textes juridiques du XIV^e siècle, mémoire de maîtrise, Université d'Anvers.

- BARCELO, G. J. (2007). Le(s) futur(s) dans les langues romanes : évolution linéaire ou cyclique ?, *Sémantique et diachronie du système verbal français*, Cahiers Chronos, 16, Amsterdam, New York, NY, 47-62.
- BARCELÓ, G. & BRES, J. (2006). *Les temps de l'indicatif en français*, Paris, Ophrys.
- BATTISTELLI, D. (2009). *La temporalité linguistique. Circonscrire un objet d'analyse ainsi que des finalités à cette analyse*, HDR, Paris 10.
- BAUDE, O. (coord.). (2006). *Corpus oraux. Guide des bonnes pratiques*, Paris, Cnrs éditions – Orléans, PUO.
- BEGUELIN, M.-J. (1998). Le rapport écrit-oral. Tendances dissimilatrices, tendances assimilatrices, *Cahiers de linguistique française*, 20, 229-253.
- BENNINGER, C., CARLIER, A. & LAGAE, V. (éds.). (2003). *Temps et texte*, Presses universitaires de Valenciennes.
- BENMAKHLOUF, A. (2002). *Frege, le nécessaire et le superflu*, Paris, Vrin.
- BENVENISTE, E. (1946). Structure des relations de personne dans le verbe, in *Problèmes de linguistique générale*, I. Paris, Gallimard, 1966, 225-236
- BENVENISTE, E. (1956). La Nature des pronoms, in *PLG*, I, 251-257.
- BENVENISTE, E. (1959). Les Relations de temps dans le verbe français, in *PLG*, I, 237-250.
- BENVENISTE, E. (1970). L'appareil formel de l'énonciation, in *PLG*, II, 79-88.
- BERGOUNIOUX, G. (1992a). Linguistique et variations : repères historiques, *Langages*, 108, 114-125.
- BERGOUNIOUX, G. (1992b). Les enquêtes de terrain en France, *Langue française*, 93, 3-21.
- BERGOUNIOUX, G., BARADUC, J. & DUMONT, C. (1992). L'Etude sociolinguistique sur Orléans (1966-1991), 25 ans d'histoire d'un corpus, *Langue française*, 93, 74-93.
- BERGOUNIOUX, G. (1994). *Aux origines de la linguistique française*, Agora / Les classiques / Pocket.
- BERGOUNIOUX, G. (2004). *Le moyen de parler*, Verdier.
- BERGOUNIOUX, G. (2005). Saussure sans sémantique ou le signifié ne fait pas sens, *Revue de Sémantique et Pragmatique*, 18, 69-85.
- BERGOUNIOUX, G. (à paraître). *Une étude de cas en linguistique de corpus*, Champion, 2016.
- BERRENDONNER, A. (1981). *Éléments de pragmatique linguistique*, Paris, Minit.
- BERTHONNEAU, A.-M. & KLEIBER, G. (1993). Pour une nouvelle approche de l'imparfait. L'imparfait, un temps anaphorique méronomique, *Langages* 112, 55-73.
- BERTHONNEAU, A.-M. & KLEIBER, G. (1994). Imparfais de politesse. Rupture ou cohésion ?, *Travaux de linguistique*, 29, 59-92.
- BERTHONNEAU, A.-M. & KLEIBER, G. (2006). Sur l'imparfait contrefactuel, *Travaux de linguistique*, 53, 7-64.
- BIBER, D. (1993). Using register-diversified corpora for general language studies, *Computational Linguistics*, 19(2), 243-258.

- BIBER, D. (2009). Corpus-Based and Corpus-driven Analyses of Language Variation and Use, in Bernd Heine and Heiko Narrog (ed.), *The Oxford Handbook of Linguistic Analysis*, Oxford University Press.
- BILGER, M. (2001). Retour sur le futur dans les corpus de français parlé, *RSFP*, 16, Université de Provence, 177-189.
- BLANC, M. & BIGGS, P. (1971). L'enquête socio-linguistique sur le français parlé à Orléans, *Le Français dans le Monde*, 85, 16-25.
- BLANCHE-BENVENISTE, C. (1997). *Approches de la langue parlée en français*, Paris, Ophrys.
- BLANCHE-BENVENISTE C., DEULOFEU J., STEFANINI J. & VAN DEN EYNDE K. (1984). *L'Approche pronominale et son application au français*, Paris, SELAF.
- BLANCHE-BENVENISTE C. & JEANJEAN C. (1986). *Le français parlé*, Edition et transcription, Paris, Didier-Erudition.
- BOURDIEU, P. (1979). *La Distinction : critique sociale du jugement*, Paris, Minuit.
- DETREZ, C. (2005). Le capital culturel, *Idées*, 142, 6-13.
- BOUROVA, V. & DENDALE, P. (2013). "Les deux conditionnels à valeur évidentielle en français, leur place dans l'ensemble des emplois du conditionnel et leur parcours diachronique", *Cahiers Chronos*, 26, 183-200.
- BRANCA-ROSOFF, S. , FLEURY, S. , LEFEUVRE, F & PIRES, M. (2009). Constitution et exploitation d'un corpus de français parlé parisien, consultable à l'adresse <http://cfpp2000.univ-paris3.fr/CFPP2000-05062009.pdf>
- BRES, J. (2003). Mais oui, il était un joli temps du passé comme les autres, le petit imparfait hypocoristique, *Langue française*, 138, 111-125.
- BRES, J. & al. (éd.) (2005). *Dialogisme et polyphonie : approches linguistiques*, Actes du colloque de Cerisy, septembre 2004, Bruxelles, De Boeck-Duculot.
- BRES, J. (2009). Dialogisme et temps verbaux de l'indicatif, *Langue française*, 163, 21-39.
- BRES, J. (2014). Dialogisme, médiativité : le jeu dialogique du futur et du conditionnel français dans le marquage d'une source indirecte par ouï-dire et par conjecture, in J.-C. Anscombre, E. Oppermann-Marsaux, A. Rodriguez Somolinos (éds.), (2014), *Médiativité, polyphonie et modalité en français : études synchroniques et diachroniques*, Paris, Presses Sorbonne nouvelle, 19-34.
- BRES, J. & MELLET, S. (2009). Une approche dialogique des faits grammaticaux, *Langue française*, 163, 3-20.
- BRUNOT, F. (1905). *Histoire de la langue française des origines à nos jours*, éd. 1966, Paris, Armand Colin.
- BRUNOT, F. (1922). *La Pensée et la langue*, Paris, Masson et Cie.
- BRUNOT, F. & BRUNEAU, Ch. (1956). *Précis de grammaire historique de la langue française*, 4^e édition, Paris, Masson.
- BYBEE, J.L., PAGLIUCA, W. & PERKINS, R.D. (1991). Back to the future, in. Traugott, E.C. & Heine, B. (eds). *Approaches to the grammaticalization*, II, Amsterdam/Philadelphia, John Benjamins Publishing Company, 17-58.

- BYBEE, J., PERKINS, R. & PAGLIUCA, W. (1994). *The Evolution of Grammar : Tense, Aspect and Mood in the Languages of the World*. Chicago : University of Chicago Press.
- CAMUSSI-NI, M.-A. (2006). *Analyse formelle et conceptuelle des formes verbales du français contemporain : A la croisée du passé simple et de l'imparfait, du futur et du conditionnel, les concepts « $\pm$ potentiel » et « défini »*, Thèse de l'Université Rennes 2
- CAPPEAU, P. & GADET, F. (2007a). L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale, *Revue française de linguistique appliquée*, 12, 99-110.
- CAPPEAU, P. & GADET, F. (2007b). Où en sont les corpus sur les français parlés ?, *Revue française de linguistique appliquée*, 12, 129-133.
- CAREL, M. & DUCROT, O. (2009). Mise au point sur la polyphonie, *Langue française*, 164, 33-43.
- CAUDAL, P. & VETTERS, C. (2005). Un traitement conjoint du conditionnel, du futur et de l'imparfait : les temps comme des fonctions d'acte de langage, *Cahiers Chronos*, 12, Temporalité et attitude, 109-124.
- CHOMSKY, N. (1995). *The Minimalist Program*, MIT Press.
- COHEN, M. (1965)2. *Le Subjonctif en français contemporain*. Paris, SEDES.
- COMRIE, B. (1985). *Tense*, Cambridge University Press.
- COMRIE, B. (1986). Tense in indirect speech, *Folia Linguistica*, 20, 265-296.
- CONDAMINE, A. (éd.) (2005). *Sémantique et corpus*, Paris, Hermès.
- CONFAIS, J.-P. (1990). *Temps, Mode, Aspect*, Toulouse, Presses Universitaires de Mirail.
- CORI, M. & DAVID, S. (2008). Les corpus fondent-ils une nouvelle linguistique ?, *Langages*, 171, 111-129.
- COSERIU, E. (1980). Aspect verbal ou aspects verbaux ? Quelques questions de théorie et de méthode, in J. David & R. Martin (dir.), Paris, Klincksieck.
- CULIOLI, A. (1980). Valeurs aspectuelles et opérations énonciatives: l'Aoristique, in David, J. et R. Martin, éd. (1980), 181-193.
- CULIOLI, A. (1990). *Pour une linguistique de l'énonciation. Opérations et représentations*, Paris, Ophrys, t. 1.
- DAMOURETTE, J. & PICHON, E. (1911-1940). *Des Mots à la pensée : essai de grammaire de la langue française*, Paris, éd. Artrey, 1970.
- DAVID, J. & MARTIN, R. (dir.). (1980). *La Notion d'aspect. Actes du colloque de Metz*, Paris, Klincksieck.
- DAVIDSEN-NIELSEN, N. (1988). Has English a future ? Remarks on the future tense *Acta Linguistica Hafniensia*, 21, 5-20.
- DE MULDER, W. (2012). Un sens épistémique pour l'imparfait et le passé simple ?, *Langue française*, 173, 99-113.
- DE SAUSSURE, L & STHIOUL, B. (2005). Imparfait et enrichissement pragmatique, *Cahiers Chronos*, 14, 103-210.

- DEULOFEU, J. (1991). La notion de dépendance syntaxique dans l'approche pronominale, *L'information grammaticale*, 50, 19-24.
- DEULOFEU, J. (1999). Questions de méthode dans la description morphosyntaxique de l'élément *que* en français contemporain, *Recherches sur le français parlé*, 15, 163-198.
- DEULOFEU, J. (2002). Comment présenter les types de phrase dans une grammaire basée sur corpus ?, *Recherches sur le français parlé*, 17, 137-165.
- DEULOFEU, J. (2003). L'approche macrosyntaxique en syntaxe : un nouveau modèle de rasoir d'Occam contre les notions inutiles ?, *Scolia*, 16, 77-95.
- DENDALE, P. (2014). Le conditionnel de reprise : son apparition en français et son traitement dans les grammaires du XVI^e au XX^e siècle, in PPERMANN-MARSAUX, RODRIGUEZ, ANSCOMBRE (éd.): *Médiativité, polyphonie et modalité en français : études synchroniques et diachroniques*, Presses Sorbonne nouvelle, pp.243-260.
- DONAIRE, M. L. (1995). Subjonctif, négation et polyphonie, *Hermes, Journal of Linguistics*, 15, 155-177.
- DONAIRE, M. L. (1998). La mise en scène du conditionnel ou quand le locuteur reste en coulisses, *Le français moderne*, 66, 2, 204-227.
- DONAIRE, M. L. (2006). Imparfait anaphorique, imparfait polyphonique ?, in Riegel M., Schnedecker C., Swiggers P. et Tamba I. (éd.), *Aux carrefours du sens*, Leuven, Peeters, 353-366.
- DUCROT, O. (1984). *Le Dire et le dit*, Paris, Minuit.
- DUCROT, O. & al. (1980). *Les Mots du discours*, Paris, Minuit.
- EMIRKANIAN, L. & SANKOFF, D. (1986). Le futur simple et le futur périphrastique dans le français parlé, in *Morphosyntaxe des langues romanes*, Actes du XVII^e Colloque international de linguistique et de philologie romanes, Aix-en-Provence, 397-407.
- FEUILLET, J. (1988). Problématique de l'auxiliation, in *Cahiers du CERLICO, La Question de l'auxiliaire*, 1-38.
- FLEISCHMAN, S. (1983). From pragmatics to grammar: diachronic reflections on complex pasts and futures in Romance, *Lingua*, 60 : 183-214.
- FLEISCHMAN, S. (1992). *The Future in Thought and Language. Diachronic Evidence from Romance*, Cambridge : Cambridge University Press.
- FLEURY, S. & BRANCA, S. (2010). Une expérience de collaboration entre linguiste et spécialiste de TAL : L'exploitation du corpus CFPP 2000 en vue d'un travail sur l'alternance Futur simple / Futur périphrastique, *Cahiers AFLS*, Volume 16(1).
- FOURNIER, N. (1986). L'imparfait chez les grammairiens classiques, de Maupas (1607) à Restaut (1730), in P. Le Goffic (1986, éd.), 13-41.
- FUCHS, C. (1986). L'ambiguïté et la paraphrase en psycho-mécanique : l'exemple de l'imparfait, in P. Le Goffic (1986, éd.), 43-54.
- FUCHS, C. & LEONARD, A.-M. (1979). *Vers une théorie des aspects : les systèmes du français et de l'anglais*. Paris, Mouton.

- GADET, F. (1989). *Le français ordinaire*, Paris, Armand Colin.
- GADET, F. (2012). Un grand corpus de français parlé : le CIEL-F. Choix épistémologiques et réalisations empiriques, *Revue française de linguistique appliquée*, 17-1, 39-54.
- GENETTE, G. (1972). *Figures III*, Paris, Seuil, col. 'Poétique'.
- GERSTENBERG, A. (2010a). Analyse sociolinguistique d'un corpus oral par regroupement hiérarchique, in Sergio Bolasco, Isabella Chiari & Luca Giuliano (eds.), *Statistical Analysis of Textual Data*, Proceedings of the 10th International Conference 'Journées d'Analyse statistique des Données Textuelles', 9-11 June 2010 - Sapienza University Rome, Milano, LED, 939–950.
- GERSTENBERG, A. (2010b). Question de genre, question de style. Quelques remarques sur la base d'un corpus oral, in Maria Iliescu, Heidi Siller-Runggaldier & Paul Danler (eds.), *Actes du XXVe Congrès International de Linguistique et de Philologie Romanes*, Innsbruck 2007, Berlin, de Gruyter, 417–426.
- GIRARD, G. (1747). *Les vrais principes de la langue françoise*, Genève, Paris, Droz.
- GOSSELIN, L. (1996). *Sémantique de la temporalité en français. Un modèle calculatoire et cognitif du temps et de l'aspect*, Louvain-la-neuve, Duculot, col. Recherches.
- GOSSELIN, L. (2001). Relations temporelles et modales dans le "conditionnel journalistique". In *Le conditionnel en français*, eds. Patrick DENDALE et Liliane TASMOWSKI. Metz: Université de Metz, p.57.
- GOSSELIN, Laurent, (2005), *Temporalité et modalité*. (postface de B. Victorri), Bruxelles, Duculot-de Boeck, col. Recherches.
- GOSSELIN, Laurent (2010), *La validation des représentations. Les modalités en français*, Amsterdam-New York, Rodopi, coll. Etudes Chronos.
- GOUGENHEIM, G., MICHÉA, R., SAUVAGEOT, A. et RIVENC, P. (1964). *L'élaboration du Français Fondamental*, Paris, Didier.
- GREIMAS, A. J. (1983). La soupe au pistou ou la construction d'un objet de valeur, in *Du sens II : essais sémiotiques*, Seuil, 157-169.
- GREVISSE, M. (2011). *Le Bon Usage*, Bruxelles, De Boeck-Duculot.
- GUILLAUME, G. (1929). *Temps et verbe*, Paris, Honoré Champion. [Réimprimé en 1965].
- HABERT, B. (2000). Des corpus représentatifs : de quoi, pour quoi, comment ?, in M. Bilger (éd.), *Linguistique sur corpus. Études et réflexions*, Perpignan, Presses Universitaires de Perpignan, 11-58. Version en ligne : http://icar.univ-lyon2.fr/ecole_thematique/idocora/documents/Habert_des_corpus_representatifs.pdf
- HABERT, B. (2005). *Instruments et ressources électroniques pour le français*, Paris/Gap, Ophrys, col. « L'Essentiel Français ».
- HABERT, B., NAZARENKO, A., & SALEM, A. (1997). *Les linguistiques de corpus*, Paris, Armand Colin et Masson.
- HAILLET, P.P. (1992). *Emplois « journalistiques » du conditionnel en français contemporain : essai de linguistique descriptive*, thèse de doctorat, Université de Toronto.

- HAILLET, P.P. (1995). *Le conditionnel dans le discours journalistique : essai de linguistique descriptive*, Neuville, Québec, Bref.
- HAILLET, P.P. (2002). *Le conditionnel en français : une approche polyphonique*, Gap, Ophrys.
- HAILLET, P.P. (2003a). Représentations discursives, point(s) de vue et signifié unique du conditionnel, *Langue Française*, 138, 35-47.
- HAILLET, P.P. (2003b). Polyphonie et modalisation, *Thélème, Revista Complutense de Estudios Franceses*, Número Extraordinario, 95-108.
- HAILLET, P. (2007). *Pour une linguistique des représentations discursives*, Bruxelles, Duculot-De Boeck.
- HAILLET, P. P. (2009). Approche polyphonique des attitudes du locuteur : constructions de type [comme si A], *Langue française*, 161, 135-145.
- HALLE, M. & MARANTZ, A. (1993), Distributed morphology and the pieces of inflection, In Kenneth Hale, Samuel Jay Keyser (eds.), *The View from Building 20*, Cambridge, Mass. : MIT Press, pp. 111–176.
- HALLE, M. & MARANTZ, A. (1994). Some key features of Distributed Morphology, in Andrew Carnie, Heidi Harley, Thomas Bures (eds.), *MIT Working Papers in Linguistics*, Vol. 21 : Papers on phonology and morphology, Cambridge, 275-288.
- HEIDEN, S., MAGUE, J-P., PINCEMIN, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. In I. C. Sergio Bolasco (Ed.), *Proc. of 10th International Conference on the Statistical Analysis of Textual Data - JADT 2010*, Vol. 2, 1021-1032, Edizioni Universitarie di Lettere Economia Diritto, Roma, Italy.
- HEIDEN, S. (2010). The TXM Platform: Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. In K. I. Ryo Otoguro (Ed.), *24th Pacific Asia Conference on Language, Information and Computation - PACLIC24*, 389-398, Institute for Digital Enhancement of Cognitive Development, Waseda University, Sendai, Japan.
- HEINE, B. & KUTEVA, T. (2002). *World Lexicon of Grammaticalization*, Cambridge University Press.
- HUOT, H. (1986). Le Subjonctif dans les complétives : subjectivité et modalisation. In M. Roant & D. Couquaux (éd.). *La Grammaire modulaire*, 81-111. Paris, Minuit.
- HUOT, H. (1987). Morphosyntaxe du verbe français et inversion du clitique sujet. *Travaux de linguistique*, Gand, 14/15, 159-176.
- HUOT, H. (1991). Quelques aspects syntaxiques de la non-assertion. Actes du XVIIIe Congrès International de Linguistique et de Philologie Romanes. Université de Trèves (Trier). Tübingen, Max Niemeyer Verlag, II, 389-401
- IMBS, P. (1960). *L'Emploi des temps verbaux en français moderne*, Paris, Klincksieck.
- JAUSS, H.-R. (1978). *Pour une esthétique de la réception*, Paris, Gallimard.
- JEANJEAN, C. (1988). Le futur simple et le futur périphrastique en français parlé : Étude distributionnelle, in C. Blanche-Benveniste et al. *Grammaire et histoire de la grammaire : Hommage à la mémoire de Jean Stéfanini*, PU de Provence, 235-257.

- KLEIBER, G. (1993). Lorsque l'anaphore se lie aux temps grammaticaux, in Vetters (éd.), *Le Temps, de la phrase au texte*, Lille, PUL, 117-166.
- KORZEN, H. et NØLKE, H. (1990). Projet pour une théorie sur les emplois du conditionnel, *Actes du XI^e Congrès des Romanistes Scandinaves*, Trondheim, Université de Trondheim, 301-312.
- KRISTEVA, J. (1969). *Séméiotikè, Recherches pour une sémanalyse*, Paris, Seuil.
- KRONNING, H. (2005). Polyphonie, médiation et modalisation : le cas du conditionnel épistémique, in Bres, J., Haillet, P. P., Mellet, S. Nølke, H. & Rosier, L. (éds), *Dialogisme et polyphonie : approches linguistiques*, Bruxelles, De Boeck Duculot, 297-312.
- KRONNING, H. (2012). Le conditionnel épistémique : propriétés et fonctions discursives, *Langue française*, 173, 83-97.
- LABELLE, M. (1994). Acquisition de la valeur des temps du passé par les enfants francophones, *Revue québécoise de linguistique*, 23, 99-121.
- LAMPITELLI, N. (2015). La structure phonologique des marqueurs flexionnels de l'imparfait de l'indicatif du français, *Langages*, 198, 91-108.
- LANGACKER, R. W. (2009). The English Present: Temporal coincidence vs. epistemic immediacy, in R. W. Langacker (ed.), *Investigations in Cognitive Grammar*, Berlin/New York, Mouton de Gruyter, 185-218.
- LANLY, A. (1958). Nous avons à parler maintenant du futur. *Le Français moderne*, 26, 16-46.
- LANLY, A. (2002). Plaidoyer pour le conditionnel en -rais. *L'information grammaticale*, 95, 20-22.
- LANSARI, L. (2009). *Linguistique contrastive et traduction. Les périphrases verbales aller + infinitif et be going to*, Ophrys.
- LAURENDEAU, P. (2000). L'alternance futur simple/futur périphrastique : une hypothèse modale, *Verbum*, tome 22, n°3, PU de Nancy, 277-292.
- LEBAUD, D. (1993). IMP : indétermination aspectuo-temporelle et changement de repère, *Le Gré des langues*, 5, 160-176.
- LE BIDOIS, G. & LE BIDOIS, R. (1935-1938). *Syntaxe du français moderne. Fondements historiques et psychologiques*, 2^e édition, 1967, Paris, A. Picard.
- LE GOFFIC, P. (éd.). (1986). *Points de vue sur l'imparfait*, Caen, Centre de Publications de l'Université de Caen.
- LE GOFFIC P. (1995). La double incomplétude de l'imparfait, *Modèles linguistiques*, 31, 133-149.
- LE GOFFIC, P. (éd.). (2001). *Le présent en français*, *Cahiers Chronos*, 7, Rodopi.
- LEMARECHAL, A. (1990). Du statut de certaines marques zéro, *Linx*, 22, 105-128.
- LEMARECHAL, A. (2001). Typologie pré-greenbergienne, morphologie et cognition : « flexionalismes » et « flexionalité », *Linx*, 45, Invariants et variables dans les langues. *Études typologiques*, 51-58.
- LYONS, J. (1970). *Linguistique générale*, Paris, Larousse.

- MAINGUENEAU, D. (1981). *Approche de l'énonciation en linguistique française*, Paris, Hachette.
- MAINGUENEAU, D. (1993). *Éléments de linguistique pour le texte littéraire*, Dunod.
- MAINGUENEAU, D. (1994). *L'énonciation en linguistique française*, 2e édition (1999), Hachette.
- MAINGUENEAU, D. (2000). *Analyser les textes de communication*, Nathan.
- MANNING, Ch. (2003). Probabilistic Syntax, in Bod R., Hay J., Jannedy S. (eds), *Probabilistic Linguistics*, Cambridge-London, The MIT Press, 289-341.
- MARANDIN, J.-M. (2001). (2001). Avant-Propos, *Cahier Jean-Claude Milner*, Verdier, 7-18.
- MAUGER, G. (1968). *Grammaire pratique du français d'aujourd'hui*, Paris, Hachette.
- MELLET, S. (1981). L'aspect verbal chez G. Guillaume et ses disciples, *L'Information Grammaticale*, 9, 6-12.
- MILNER, J.-C. (1989). *Introduction à une science du langage*, Paris, Seuil.
- MOESCHLER, J. ; REBOUL, A. ; LUSCHER, J.-M. & JAYEZ, J. (1994). *Langage et pertinence. Référence temporelle, anaphore, connecteurs et métaphore*, Presses Universitaires de Nancy.
- MOLENDIJK, A. (1993). Présuppositions, implications, structure temporelle, in C. Vetter (éd.), *Le Temps, de la phrase au texte*, Lille, PUL, 167-191.
- MOLENDIJK, A. (1996). Anaphore et imparfait : la référence globale à des situations présupposées ou impliquées, *Cahiers Chronos*, 1, 109-124.
- MONDADA, L. (2005). L'analyse de corpus dans la perspective de la linguistique interactionnelle : des analyses de cas singuliers aux analyses de collections, in A. CONDAMINE (éd.), *Sémantique et corpus*, Paris, Hermès, 76-108.
- MONVILLE-BURSTON, M. & BURSTON, J. (2005). Retour à « remise de peine » : l'imparfait, un toncal à faible marquage, *Cahiers Chronos*, 14, 135-156.
- MULLINEAUX, L.A. & BLANC, M. H. A. (1982). The problem of classifying the population sample in the socio linguistic survey of Orléans (1969) in terms of socio-economic, social and educational categories, *Review of Applied Linguistics*, 55, 3-37.
- NØLKE, H. (1994). La dilution linguistique des responsabilités. Essai de description polyphoniques des marqueurs évidentiels *il semble que* et *il paraît que*, *Langue française*, 102, 84-94.
- NØLKE, H. (2001). *Le regard du locuteur 2. Pour une linguistique des traces énonciatives*, Paris, Kimé.
- NØLKE, H. (2006). Connecteurs pragmatiques. Apport de quelques connecteurs à la structure polyphonique, *Le français moderne*, 1, 32-42.
- NØLKE, H. (2009). Types d'êtres discursifs dans la scapoline, *Langue française*, 164, 81-96.
- NØLKE H., FLØTTUM K., NORÉN C. (2004). *ScaPoLine, La théorie scandinave de la polyphonie linguistique*, Paris, Kimé.
- NYROP, K. (1930). *Grammaire historique de la langue française*, VI, Copenhague, 4^e édition revue, Genève, Slatkine Reprints, 1979.

- PATARD, A. (2007). L'un et le multiple. L'imparfait de l'indicatif en français : valeur en langue et usages en discours, Thèse de l'université Paul-Valéry, Montpellier III.
- PATARD, A. & DE MULDER, W. (2012). L'évolution des usages du conditionnel en français, *Faits de langues*, 40, p. 29-36.
- PATARD, A., GRABAR, N. & DE MULDER, W. (à paraître). Étude diachronique du conditionnel passé ou l'origine de la contrefactualité, *Journal of French Language Studies*.
- PERRIN, L. (2005). Polyphonie et séquence écho, in J. BRES & al. (éd.), 173-185.
- PERRIN, L. (2006a). Voix et points de vue étrangers dans le discours, *Le français moderne*, 1, 22-31.
- PERRIN L. (éd.). (2006b). *Le sens et ses voix. Dialogisme et polyphonie en langue et en discours*, Recherches linguistiques 28, Metz : Université de Metz.
- PERRIN, L. (2006c). Autophonie et formules modales. Verbes de parole et d'opinion à la première personne, *Travaux de linguistique*, 52, 41-55.
- PERRIN, L., VINCENT, D. (1997). L'interprétation du verbe dire en contexte conversationnel : des clauses performatives aux clauses descriptives, *Revue de sémantique et pragmatique*, 2, 201-217.
- PINCHON, J. & COUTE, B. (1981). *Le Système verbal français*. Paris, Nathan.
- POIBEAU, T. (2014). La linguistique est-elle soluble dans la statistique ? *Revue Sciences/Lettres* [En ligne], 2 | 2014, mis en ligne le 07 octobre 2013, consulté le 21 juillet 2015. URL : <http://rsl.revues.org/402>.
- POLLARD, C. and SAG, I. (1994). *Head-Driven Phrase Structure Grammar*, Chicago, University of Chicago Press & Stanford, CSLI Publications.
- RASTIER, F. (2005). Enjeux épistémologiques de la linguistique de corpus, in G. Williams (dir.) *La Linguistique de corpus*, Rennes, Presses Universitaires de Rennes, 31-45.
- REICHENBACH, H. (1980). *Elements of Symbolic Logic*, New-York, Dover Publications (1ère éd. 1947).
- RENDULIC, N. (en cours). *Le discours représenté dans les interactions orales. De l'étude des structures en contexte vers la construction de l'image du locuteur*, Thèse en cours de l'université d'Orléans.
- REVAZ, F. (1987). Du descriptif au narratif et à l'injonctif, *Pratiques*, 56, 18-38.
- REVAZ, F. (1997). *Les textes d'action*, Metz, Klincksieck.
- RIEGEL, M., PELLAT, J.-C. & RIOUL, R. (1998). *Grammaire méthodique du français*, PUF, 4e édition.
- SABANEEVA, M. (1993). Mode verbal et problèmes connexes. Essai d'une formule nouvelle, *La Linguistique*, 29, 1, 55-65.
- SABIO, F. (1995). Micro-syntaxe et macro-syntaxe. L'exemple des compléments antéposés en français, *Recherches sur le français parlé*, vol. 13, 112-155.
- SABIO, F. (2011). *Syntaxe et organisation des énoncés - Observations sur la grammaire du français parlé*, mémoire d'habilitation à diriger des recherches, Université Aix-Marseille 1.

- SAUSSURE (de), F. (1916). *Cours de linguistique générale*, rééd. Payot, 1972.
- SKROVEC, M. (2014). *Répétitions, entre syntaxe en temps réel et rhétorique ordinaire*, Freiburg, Rombach/FRIAS.
- TASMOWSKI-DE RYCK, L. (1985). L'Imparfait avec et sans rupture, *Langue française*, 67, 59-77.
- TOGEBY, K. (1982). *Grammaire française*, Copenhague, Akademisk Forlag, vol.2-3.
- VET, C. (1980). *Temps, aspect et adverbess de temps en français contemporain : essai de sémantique formelle*, Genève, Droz.
- VET, C. (1981). La Notion de 'monde possible' et le système temporel et aspectuel du français, *Langages*, 64, 109-123.
- VET, C. (1985). Univers de discours et univers d'énonciation : les temps du passé et du futur, *Langue française*, 67, 38-57.
- VET, C. (1996). Anaphore et deixis dans le domaine temporel, *Cahiers Chronos*, 1, Anaphores temporelles et (in)cohérence, 147-163.
- VET, C. (2005). L'imparfait : emploi anaphorique et emplois non anaphoriques, *Cahiers Chronos*, 14, 33-44.
- VETTERS, C. (éd.) (1993a). *Le Temps, de la phrase au texte*, Lille, PUL.
- VETTERS, C. (1993b). Temps et deixis, in Vetters (éd.), *Le Temps, de la phrase au texte*, Lille, PUL, 85-115.
- VETTERS, C. (1996). *Temps, aspect et narration*, Amsterdam, Rodopi.
- VUILLAUME, M. (1990). *Grammaire temporelle des récits*, Paris, Minuit.
- VUILLAUME, M. (1993). Le Repérage temporel dans les textes narratifs, *Langages*, 112, 92-105.
- WAGNER, R.-L. (1939). *Les Phrases hypothétiques introduites par si dans la langue française des origines à la fin du XVIe siècle*, Paris, Droz.
- WAGNER, R.-L. & PINCHON, J. (1962). *Grammaire du français classique et moderne*, Paris, Hachette.
- WEINRICH, H. (1973). *Le Temps*, Paris, Le Seuil.
- WEIST, R.M. (1986). Tense and aspect: temporal systems in child language, in P. Fletcher & M. Garman (ed.), *Language acquisition: studies in first language development*, Cambridge, Cambridge University Press, 356-374.
- WEIST, R.M. (1989). Time concepts in language and thought: filling the piagetian void from two to five years, in I. Levin & D. Zakay (ed.), *Time and Human Cognition*, Elsevier Science Publishers (North-Holland), 63-118.
- WILKS, Y. (2011). Computational Semantics requires computations, *Proceedings of the 24th FLAIRS Conference*, Melbourne (FL), 34, vol. 4.
- WILMET, M. (1996). L'Imparfait : le temps des anaphores ?, *Cahiers Chronos*, 1, 199-215.
- WILMET, M. (1997). *Grammaire critique du français*, Louvain-la-Neuve, Duculot.
- YVON, H. (1951). Convient-il de distinguer dans le verbe français des temps relatifs et des temps absolus ? *Le français moderne*, 19, 4, 265-276.

YVON, H. (1952). Faut-il distinguer deux conditionnels dans le verbe français ? *Le Français moderne*, 20, 4, 249-265.

Liste des travaux

- ABOUDA, L. (1997a). Recherches sur la syntaxe et la sémantique du conditionnel en français moderne, Thèse, Université Paris 7.
- ABOUDA, L. (1997b). Le conditionnel : temps ou mode ? Arguments syntaxiques, *Revue Romane*, 32, 2, 179-198.
- ABOUDA, L. (1998a). Vers une localisation syntaxique des modes verbaux, *Cahiers Chronos*, 3, Variations sur la référence verbale, 293-322.
- ABOUDA, L. (1998b). De la variation modale dans les complétives, *Le Langage et l'Homme*, vol. XXXIII, n° 1, 121-136.
- ABOUDA, L. (1999). Les modes verbaux dans les phrases indépendantes, *Travaux de linguistique*, 38, Sémantique, interprétation et effets syntaxiques, 89-108, juin 1999, Louvain-La-Neuve, Duculot.
- ABOUDA, L. (2001). Les emplois journalistique, polémique, et atténuatif du conditionnel. Un traitement unitaire, in P. Dendale & L. Tasmowski (éds.), *Le conditionnel en français*, Metz, Université de Metz / Paris, Klincksieck, Collection 'Recherches Linguistiques', n° 25, 277-294.
- ABOUDA, L. (2002). Négation, interrogation et alternance Indicatif-Subjonctif, *Journal of French Language Studies*, 12.1, 1-22, Cambridge University Press.
- ABOUDA, L. (2003). De la distorsion temporelle, in Benninger, Carlier & Lagae (eds.), *Temps et texte*, Presses universitaires de Valenciennes, 13-28.
- ABOUDA, L. (2004a). L'imparfait d'atténuation : entre justification et déperformativité, in S. Mejri (dir), *Espace euro-méditerranéen : une idiomaticité partagée*, Actes du colloque international (tome 1), Université de Tunis, ENS, Etudes de littérature, langue et civilisation (vol. 12) : 7-23.
- ABOUDA, L. (2004b). Deux types d'imparfait atténuatif, *Langue française*, n°142 : 58-74.
- ABOUDA, L. (2004c). De la polyphonie à la polychronie. Un modèle de traitement de la polysémie temporelle. Application au futur, in Javier Suso Lopez & Rodrigo Lopez Carrillo (coord.), *Le français face aux défis actuels. Histoire, langue et culture*, vol. 1, Universidad de Granada, APFUE – GILEC : 149-160.
- ABOUDA, L. & BAUDE, O. (2007). Constituer et exploiter un grand corpus oral : choix et enjeux théoriques. Le cas des Eslo, in F. Rastier et M. Ballabriga (dir.), *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*, Actes du XXVIIe Colloque d'Albi « Langages et Signification », 161-168.

- ABOUDA, L. & BAUDE, O. (2009). Du français fondamental aux Eslo, *Les Cahiers de Linguistique de Louvain*, 33, 2, 131-146.
- ABOUDA, L. (2010). De la visibilité syntaxique des modes, de l'invisibilité syntaxique des temps, in Álvarez Castro, Camino / Bango de la Campa, Flor M^a / Donaire, María Luisa (éds), *Liens linguistiques. Etudes sur la combinatoire et la hiérarchie des composants*, Collection « Sciences pour la Communication », Vol. 90, 319-333, Peter Lang.
- ABOUDA, L. (2012). Le temps des omelettes. Une nouvelle valeur pour le présent ?, *Revue de Sémantique et Pragmatique*, 32, 71-94.
- ABOUDA, L. & Skrovec, M. (2015a) (à paraître). Du rapport entre formes synthétique et analytique du futur. Étude de la variable modale dans un corpus oral micro-diachronique, *Revue de Sémantique et Pragmatique*, n° 38, décembre 2015.
- ABOUDA, L. & Skrovec, M. (2015b) (soumis). Du mouvement au figement : pragmatization de la forme on va dire. Etude micro-diachronique sur un corpus oral, *Actes du colloque de Grenade "Langage et analogie"*.
- LEFEUVRE, A. ; ANTOINE, J.-Y. ; SAVARY, A. ; SCHANG, E. ; ABOUDA, L. MAUREL, D. & ESHKOL, I. (2014). Annotation de la temporalité en corpus : contribution à l'amélioration de la norme TimeML, *taln*, 21, 562-567.
- RENDULIC, N. & ABOUDA, L. (2015). (à paraître). La mise en scène dans la représentation des interactions autres, *Actes du Xe CILF*, Lambert Lucas (automne 2015).

Index des noms

A

ADAM	106, 112
ANSCOMBRE	27, 150
AUROUX	36, 92, 93, 94

B

BERGOUNIOUX	46, 100, 101, 102, 103, 106, 109, 112, 117
BERTHONNEAU	59, 60, 66, 67, 94, 97
BIBER	105, 124
BRES	58, 148, 155

C

CAPPEAU	96, 105, 110, 140
CHOMSKY	31, 49
CONFAIS	45, 56, 61, 82, 83, 114, 133
COSERIU	17, 52, 55
CULIOLI	48, 55

D

DAMOURETTE & PICHON	12, 17, 33, 43, 52, 53, 55, 56, 58, 73, 75, 78, 96, 108
DONAIRE	78, 81, 82, 148, 161
DUCROT	41, 57, 73, 77, 78, 79, 80, 81, 83, 87, 94, 134

F

FLEISCHMAN	14, 75, 114
FUCHS	43, 57

G

GADET	19, 59, 96, 105, 110, 114
GOSSELIN	10, 35, 38, 39, 40, 44, 45, 46, 47, 48, 50, 54, 61, 62, 63, 64, 66, 73, 74, 75, 76, 82, 85, 86, 87, 113, 127, 130, 134

H

HABERT	89, 97, 98, 99, 105, 106, 108, 111, 153
HAILLET	27, 35, 37, 47, 72, 77, 78, 82, 95, 148, 155

M

MAINGUENEAU	52, 114, 133
MILNER	27, 28, 31, 32, 36, 48, 92, 93, 94, 96, 108, 156

N

NØLKE	78, 80, 95, 155
-------	-----------------

P

PATARD 14, 57, 58, 59, 60, 66, 72, 75, 82, 98

S

SAUSSURE (Ferdinand de) 44, 46, 48, 103, 138, 149

V

VETTERS 18, 37, 40, 53, 77, 155, 156, 158

W

WEINRICH 20, 45, 52, 60, 80, 104
WILMET 46, 52, 53, 59, 62, 75, 83, 84, 112

Index des notions

A

Annotation 45, 66, 70, 105, 107, 116, 117, 118-130

C

Corpus 10, 20, 21, 23, 24, 31, 38, 39, 45, 46, 56, 63, 65, 66, 70, 86 *sq.*, 93, 94, 95, 96-130, 131

E

ESLO 10, 63, 66, 96-108, 110, 111, 114, 118

M

Modalité 8, 19, 44, 46, 119

Mode 30-38, 43, 44, 53, 59, 66, 84, 93, 109, 133, 134

et statut du conditionnel 8-29

Morphème 15, 16, 17, 18, 19, 45, 48, 49, 50, 134, 135, 136

Morphologie 8, 16, 42, 44, 48, 53, 151

O

Oral 10, 20, 24, 35, 38, 45, 51, 63, 65, 66, 70, 86-89, 92-108, 109, 110, 116, 129, 136

P

Paradigme 9, 11, 12, 15, 17, 21, 25, 29, 43 *sq.*, 72, 96, 115, 134

Polychronie 75, 76, 79, 83, 135

Polyphonie 9, 75, 76, 78, 79

Polysémie 9, 10, 46, 47, 71, 73, 79, 80, 82, 135

S

Sémantique 10, 12, 18, 25, 29, 31, 32, 36, 39, 40, 42, 44, 45, 46, 47, 48, 54, 55, 56, 60, 71, 87, 88, 90, 91, 92, 104, 117, 118, 131

Structures syntaxiques 20, 21, 22, 29, 32, 133

Syntaxe 21, 22, 25, 30-40, 43, 48, 56, 59, 104, 116, 133

Exemple en syntaxe 86, 87, 88, 90, 91, 92

T

TXM 105, 116, 117, 118

V

Visibilité syntaxique 8, 30, 31, 35, 38, 43, 133

Z

Zéro 9, 16, 45, 80, 135

Table des matières

Remerciements	Erreur ! Signet non défini.
Sommaire	7
Introduction	10
1. Du statut catégoriel du conditionnel à l'hypothèse de l'identification syntaxique des modes : à la recherche du paradigme	14
1.1. Temps ou Mode ?	15
1.1.1. Les arguments classiques	15
1.1.2. Arguments syntaxiques	23
1.2. La visibilité syntaxique des modes	33
1.2.1. Le modèle théorique	33
1.2.2. De la visibilité des modes, et de l'invisibilité des temps en syntaxe	34
2. Une vision de l'indicatif : propositions	44
2.1. Préliminaires	46
2.1.1. Paradigmes de paradigmes	46
2.1.2. Mode vs temps VS Modalité vs temporalité	47
2.1.3. Traitement de la polysémie temporelle	49
2.2. Deux groupes de formes temporelles	51
2.2.1. Morphologie de l'indicatif	51
2.2.2. Arguments pour une bipartition	54
2.3. De la polyphonie à la polychronie	78
2.3.1. Polyphonie linguistique	79
2.3.2. De la polyphonie à la polychronie : un modèle de traitement de la polysémie temporelle ?	82
3. Retour aux données	88
3.1. La linguistique par l'exemple	91
3.1.1. <i>Pierre, Marie et la pomme</i>	91
3.1.2. De Stendhal à Martin	95
3.2. La linguistique en corpus	99
3.2.1. <i>Les ESLOs</i>	99
3.2.2. <i>Etude de la temporalité en corpus</i>	111
Conclusion générale	135
Annexe 1 : Corpus futur	140
Sélection des locuteurs des entretiens d'ESLO1	140
Sélection des locuteurs des entretiens d'ESLO2	142

Annexe 2 : Extrait d'un tableau CSV en cours d'annotation (avant réinjection sous TXM).....	145
Annexe 3 : Exemple de requête sous TXM (après réinjection de l'annotation)	146
Bibliographie.....	147
Liste des travaux	159
Index des noms	161
Index des notions	163