

# **L'expression de la sous-détermination dans les procédés de référence en français contemporain**

Thèse présentée à la Faculté des Lettres et Sciences Humaines

Institut des Sciences du langage et de la communication

Linguistique française

Université de Neuchâtel

Pour l'obtention du grade de docteur ès Lettres

Par

**Laure Anne JOHNSEN**

**Sous la direction de**

**Prof. Marie-José BEGUELIN** (Université de Neuchâtel)

**Prof. Alain BERRENDONNER** (Université de Fribourg)

**Membres du jury :**

**Patricia CABREDO HOFHERR**, chargée de recherche, CNRS, Université Paris 8

**Francis CORNISH**, professeur émérite, Université Toulouse Jean Jaurès

**Catherine SCHNEDECKER**, professeure, Université de Strasbourg

**SOUTENUE LE 19 MAI 2017**



**IMPRIMATUR**

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de Mme Marie-José Béguelin, co-directrice de thèse, professeure honoraire, Université de Neuchâtel ; M. Alain Berrendonner, co-directeur de thèse, professeur honoraire, Université de Fribourg ; Mme Patricia Cabredo Hofherr, chargée de recherches, CNRS, Université Paris 8 ; M. Francis Cornish, professeur émérite, Université de Toulouse ; Mme Catherine Schnedecker, professeure, Université de Strasbourg autorise l'impression de la thèse présentée par Mme Laure Anne Johnsen en laissant à l'auteur la responsabilité des opinions énoncées.

Neuchâtel, le 19 mai 2017

Le doyen  
Hédi Dridi*p.o. Crani*



## Résumé :

Le propos de cette thèse est de confronter des faits de sous-détermination référentielle en français aux théories existantes dans le domaine de la référence en linguistique. Notre visée est de proposer une modélisation suffisamment générale pour le traitement des expressions référentielles en discours. Les faits de sous-détermination (via, par exemple, les marqueurs vagues comme (*tout*) *ça*, *ils* non introduit, l'adverbe *là*, les dénominations postiches comme *le truc*, *le machin*, etc.) remettent en effet en question certains postulats bien implantés dans la tradition sémantique, tels que la vocation identificatoire des désignateurs ou encore la notion d'anaphore conçue dans une optique aussi bien textuelle que cognitive. L'accès aux bases de données fournit de nos jours une ressource opportune pour renouveler la réflexion dans le domaine – en particulier les faits d'oral – là où la recherche en français se fonde très majoritairement sur des critères d'acceptabilité à l'égard d'exemples fabriqués ou sur l'examen de données consensuelles provenant de l'écrit travaillé. Après un bilan critique des théories de la référence en vigueur, nous proposons une typologie et un inventaire des faits de sous-détermination. Puis, nous présentons deux études empiriques portant sur des phénomènes productifs dans les genres « improvisés » de l'oral : le cas de *tout ça* dans les énumérations comme *la France est un pays qui adore le progrès la technique les avions tout ça* ainsi que le cas de *ils* à référence sous-déterminée comme dans *ils ont annoncé du mauvais temps*. Nous montrons que les situations de sous-détermination référentielle peuvent non seulement être le fait d'une ignorance du locuteur, mais également résulter d'une optimisation de la pertinence, laquelle se traduit par des stratégies référentielles aux rendements discursifs variés. Le choix d'une expression référentielle, loin d'être déterminé univoquement par un segment antécédent ou un statut cognitif du référent, repose ainsi crucialement sur les besoins communicationnels des participants de l'interaction en cours.

## Mots clés :

Référence, sous-détermination, anaphore, désignateur, mémoire discursive, objet-de-discours, pronom

**Abstract :**

The purpose of this doctoral thesis is to confront referential underdetermination facts in French to existing theories in the field of linguistic reference in order to put forward a modelling approach to generalize the treatment of referential expressions in discourse. Indeed, underdetermination facts (*e.g.*, by means of vague markers like *(tout) ça*, non-introduced *ils*, the adverb *là*, placeholder items like *le truc*, *le machin*, etc.) challenge certain assumptions deeply anchored in the semantic tradition, such as the identification mission of designators or the notion of anaphora in a textual, as well as in a cognitive, approach. Access to databases, in particular to spoken facts, provides nowadays opportune resources to renew the thought in the field, where research in French is mainly based on acceptability judgments regarding constructed examples or on the observation of consensual data from formal written discourse. After a critical overview of the current reference theories, we offer a typology and an inventory of underdetermination facts. Then we present two empirical studies on productive phenomenon in improvised spoken genres : the case of *tout ça* ending enumerations like in *la France est un pays qui adore le progrès la technique les avions tout ça* and the case of the underdeterminate value of *ils* like in *ils ont annoncé du mauvais temps*. We show that referential underdetermination situations result not only from the speaker's ignorance, but also by optimizing relevance, which is achieved through diverse referential strategies and discourse outputs. The choice for a referential expression then, far from being exclusively determined by a textual antecedent or a cognitive referential status, crucially rests upon the communication needs of the participants in the ongoing interaction.

**Keywords :**

Reference, underdetermination, anaphora, designator, discourse memory, discourse object, pronoun

## Remerciements

Je tiens à exprimer ma profonde reconnaissance envers Marie-José Béguelin et Alain Berrendonner, mes co-directeurs de thèse, pour leur soutien constant, leur disponibilité et leurs remarques constructives. Leurs travaux précurseurs et leurs enseignements dans le domaine de la référence ont été une source d'inspiration déterminante pour cette recherche. Je les remercie de la confiance qu'ils m'ont témoignée et qui m'a permis de mener à bien ce travail.

Je remercie également Patricia Cabredo Hofherr, Catherine Schnedecker et Francis Cornish d'avoir accepté de faire partie du jury de thèse. Leur lecture attentive et leurs suggestions lors du colloque préliminaire m'ont aidée à améliorer sensiblement la qualité scientifique du manuscrit final.

Grâce au soutien du *Fonds National Suisse de la recherche scientifique*, j'ai eu l'occasion de passer un séjour de recherche enrichissant à l'Université Toulouse Jean Jaurès au sein de l'équipe ERSS, que je remercie pour son accueil chaleureux. Un grand merci à Francis Cornish qui a rendu possible ce séjour dans les meilleures conditions et qui a toujours été disponible.

Je suis également reconnaissante envers mes collègues linguistes des Universités de Neuchâtel et de Fribourg pour leur soutien et les échanges stimulants partagés au bureau ou lors de nos expéditions scientifiques. Merci en particulier à Mathieu Avanzi, Marine Borel, Virginie Conti, Gilles Corminboeuf, François Delafontaine, Frédéric Gachet, Alexander Guryev, Pascal Montchaud et Géraldine Zumwald Küster. Je souhaite également remercier mes collègues actuels de l'*Institut de langue et civilisation françaises* pour leur encouragement et leur compréhension durant les dernières étapes de ma thèse.

Je tiens pour finir à exprimer ma gratitude envers ma famille et mes amis pour leur soutien sans faille tout au long de ces années. Un merci tout spécial à ma mère et à Florent pour leur contribution technique. Enfin, j'aimerais remercier Adèle et Olivier, qui ont partagé cette aventure malgré eux avec une compréhension et un enthousiasme admirables.



# TABLE DES MATIERES

## INTRODUCTION GENERALE ..... 15

1.	Objet d'étude .....	15
2.	Méthodologie et données.....	17
3.	Organisation du contenu.....	19

## *Première Partie : Référence, anaphore pronominale et sous-détermination : état des lieux..... 21*

## CHAPITRE I GENERALITES SUR LA REFERENCE ET LES EXPRESSIONS REFERENTIELLES ..... 23

1.	Introduction .....	25
2.	Notions fondamentales .....	27
	2.1 Référence, existence et monde réel .....	27
	2.2 Sens et référence .....	29
	2.3 Référence et prédication.....	35
	2.4 Bilan.....	37
3.	Typologie des référents .....	41
	3.1 Les ordres de Lyons (1977).....	41
	3.2 Deuxième et troisième ordres : tentative de classement.....	43
	3.3 Objets « indiscrets ».....	48
	3.4 Bilan.....	50
4.	Les expressions référentielles : catégories majeures.....	55
	4.1 Les noms propres .....	55
	4.2 Les SN définis.....	60
	4.3 Les SN démonstratifs .....	63
	4.4 Les pronoms personnels.....	66
	4.5 Les SN indéfinis.....	71
	4.6 Approche cognitive des expressions référentielles .....	73
	4.6.1 La théorie de l'Accessibilité.....	74
	4.6.2 La Hiérarchie du Donné .....	75
	4.6.3 Bilan .....	77
5.	Modes de désignation référentielle : <i>anaphore</i> et <i>deixis</i> .....	79
	5.1 Origine des notions d'anaphore et de deixis.....	79
	5.2 Conception traditionnelle de l'opposition anaphore vs deixis .....	80
	5.3 L'anaphore sous l'angle textualiste .....	81
	5.4 Limites d'une approche textualiste de l'anaphore .....	83
	5.5 Conception cognitive de l'opposition anaphore vs deixis .....	85
	5.6 Bilan.....	88
6.	Vers un modèle constructiviste du discours.....	91
	6.1 Les référents discursifs.....	91
	6.2 Une représentation du discours .....	92
	6.3 Le modèle du discours fribourgeois .....	94
	6.3.1 Principes méthodologiques .....	94
	6.3.1.1 Le rapport oral-écrit .....	94
	6.3.1.2 Le statut des données .....	95
	6.3.1.3 Une vision non déterministe du rapport monde-langue.....	97

6.3.2	Discours et texte .....	100
6.3.3	Les articulations du discours.....	100
6.3.4	Mémoire discursive et nature des objets-de-discours.....	102
6.3.5	Opérations de pointage .....	104
6.3.6	Stratégies référentielles .....	106
7.	Conclusion.....	109
<b>CHAPITRE II FONCTIONNEMENT DE L'ANAPHORE PRONOMINALE .....</b>		<b>111</b>
1.	Introduction .....	113
2.	Propriétés morphologiques et sémantiques du pronom clitique de 3 <sup>e</sup> personne.....	115
2.1	La personne.....	117
2.2	Le cas .....	118
2.3	Le nombre .....	119
2.4	Le genre .....	125
2.5	Bilan.....	131
3.	Approches de l'anaphore pronominale .....	133
3.1	Conception substitutive.....	133
3.2	Conception textualiste.....	136
3.3	Conception cognitive .....	137
3.4	Bilan.....	139
4.	L'anaphore pronominale indirecte ou associative.....	143
4.1	L'anaphore associative.....	143
4.1.1	Conception étroite.....	144
4.1.2	Conception large.....	146
4.1.3	Bilan .....	150
4.2	Contextes d'apparitions des pronoms indirects .....	151
4.2.1	Restrictions d'emploi.....	151
4.2.2	Régularités d'emploi.....	156
5.	Conclusion.....	161
<b>CHAPITRE III LA SOUS-DETERMINATION REFERENTIELLE .....</b>		<b>163</b>
1.	Introduction .....	165
2.	Caractéristiques de la sous-détermination .....	167
2.1	Sous-détermination et pertinence .....	167
2.2	Une notion scalaire.....	167
2.3	Types de sous-détermination.....	168
3.	Notions voisines .....	173
3.1	Référence indéfinie .....	173
3.2	Approximation .....	173
3.3	Ambiguïté .....	175
3.4	Ambivalence .....	177
3.5	Interprétation transparente vs opaque.....	178
4.	Emergence et traitement de la sous-détermination .....	179
4.1	Motivations de la sous-détermination .....	179
4.2	Sous-spécification sémantique.....	182
4.3	Traitement cognitif de la sous-spécification.....	183
5.	Inventaire des moyens d'expression de la sous-détermination .....	185
5.1	Ressources lexicales.....	185
5.1.1	Hyperonymes.....	185
5.1.2	N sous-spécifiés « capsules ».....	186

5.1.3	Termes postiches .....	188
5.2	Ressources non lexicales.....	189
5.2.1	Anaphores zéro .....	189
5.2.2	Pronoms conjoints .....	190
5.2.2.1	Pronoms sujets : <i>ça, ils</i> .....	190
5.2.2.2	Pronoms régimes : <i>le, y, en</i> .....	191
5.2.2.3	Pronoms régimes dans les « aphorismes lexicalisés » .....	192
5.2.3	Pronoms disjoints.....	193
5.2.3.1	Les démonstratifs <i>ça, ceci, cela</i> .....	193
5.2.3.2	Les pronoms <i>l'un... (l'autre), les uns... les autres</i> .....	196
5.2.4	L'adverbe <i>là</i> .....	197
6.	Conclusion.....	199

***Deuxième Partie Manifestations de la sous-détermination : deux études empiriques ..... 201***

***AVANT-PROPOS : LA CONSTITUTION DES DONNÉES..... 203***

***CHAPITRE IV FONCTIONNEMENT REFERENTIEL ET PRAGMATIQUE DE TOUT ÇA..... 205***

1.	Introduction .....	207
2.	Les emplois de <i>ça</i> .....	209
2.1	<i>ça</i> conjoint.....	210
2.1.1	<i>ça</i> pointe sur un objet catégorisé.....	210
2.1.1.1	<i>ça</i> redouble un sujet .....	210
2.1.1.2	<i>ça</i> et la référence humaine.....	211
2.1.2	<i>ça</i> pointe sur un objet non catégorisé : .....	213
2.1.2.1	<i>ça</i> se rapporte à l'expérience du locuteur .....	213
2.1.2.2	<i>ça</i> « impersonnel » .....	213
2.1.2.3	<i>ça</i> renvoie à un procès ou à un objet « indiscret » .....	214
2.1.2.4	<i>ça</i> renvoie à une quantité.....	216
2.2	<i>ça</i> disjoint.....	216
2.2.1	<i>ça</i> complément.....	216
2.2.2	<i>ça</i> détaché .....	218
2.3	Bilan sur les emplois de <i>ça</i> .....	220
3.	Les emplois de <i>tout ça</i> .....	221
3.1	Le quantificateur <i>tout</i> .....	221
3.2	<i>tout ça</i> régi .....	223
3.3	<i>tout ça</i> détaché .....	223
3.4	<i>tout ça</i> appositif régissant un circonstant .....	224
3.5	<i>tout ça</i> dans les listes.....	225
3.6	Bilan sur les emplois de <i>tout ça</i> .....	225
4.	Etude de <i>tout ça</i> dans les listes à l'oral .....	227
4.1	Travaux antérieurs.....	228
4.1.1	Les particules d'extension.....	228
4.1.2	<i>tout ça</i> particule d'extension .....	229
4.1.3	Bilan .....	229
4.2	Remarques sur les données .....	230
4.3	Propriétés des listes.....	231
4.3.1	Aspects lexico-syntaxiques.....	231
4.3.2	Aspects sémantico-pragmatiques.....	233
4.3.3	Aspects prosodiques .....	235

4.4	Fonctionnement référentiel .....	239
4.4.1	Un pointage résomptif ?.....	240
4.4.2	L'inférence d'un ensemble ?.....	241
4.5	Fonctionnement pragmatique.....	242
4.5.1	Marquage de l'intersubjectivité .....	243
4.5.2	Rôle de ponctuant .....	244
4.5.3	Approximation de paroles rapportées .....	246
4.5.4	Euphémisme .....	247
4.5.5	Emphase .....	248
4.5.6	Association à d'autres éléments.....	249
4.5.6.1	<i>machin tout ça</i> .....	249
4.5.6.2	<i>tout ça mais</i> .....	249
5.	Conclusion.....	251

**CHAPITRE V FONCTIONNEMENT REFERENTIEL ET PRAGMATIQUE DE *ILS* A VALEUR SOUS-DETERMINEE..... 253**

1.	Introduction .....	255
2.	Travaux antérieurs .....	257
2.1	Le traitement de ils dans les grammaires .....	257
2.2	Le traitement de ils en linguistique française .....	261
2.2.1	Un référent sous-déterminé.....	261
2.2.2	Indices morpho-sémantiques .....	264
2.2.2.1	Le nombre pluriel.....	264
2.2.2.2	Le genre masculin .....	265
2.2.2.3	Le trait [+humain] .....	266
2.2.3	Bilan .....	266
2.3	Traitement de <i>they</i> en psycholinguistique .....	267
2.4	Traitement de la 6 <sup>e</sup> personne en grammaire générative et en typologie des langues .....	269
2.4.1	Interprétation arbitraire .....	269
2.4.2	Typologie des pronoms R-impersonnels.....	271
2.4.3	L'hypothèse de la grammaticalisation .....	274
2.4.4	Bilan .....	276
2.5	Synthèse générale sur les travaux antérieurs .....	277
3.	Etude de <i>ILS</i> dans une collection de données en français .....	279
3.1	Remarques sur les données .....	280
3.2	Classement des données .....	281
3.2.1	Anaphore indirecte.....	281
3.2.1.1	Inférence collectif-classe.....	281
3.2.1.2	Inférence lieu-classe.....	283
3.2.1.3	Inférence scénario-actant.....	288
3.2.1.4	Bilan.....	291
3.2.2	Anaphore indirecte ou variable non instanciée ?.....	292
3.2.2.1	Ambiguïté d'analyse .....	292
3.2.2.2	Quiproquos interprétatifs.....	294
3.2.3	Variable non instanciée.....	297
3.2.3.1	Promotion du procès .....	297
3.2.3.2	Verba dicendi .....	300
3.2.3.3	Incrimination feinte.....	302
3.2.3.4	Prédicat non typant.....	303
3.2.3.5	Bilan.....	304

3.3	Discussion.....	304
3.4	ILS dans le paradigme des constructions à agent sous-déterminé.....	307
3.4.1	<i>ILS</i> vs le passif.....	307
3.4.2	<i>ILS</i> vs <i>ON</i> .....	314
3.4.3	Bilan.....	319
4.	Conclusion.....	321
<b>CONCLUSION GÉNÉRALE .....</b>		<b>323</b>
1.	Synthèse des principaux résultats.....	323
2.	Pistes de recherche.....	326
<b>BIBLIOGRAPHIE.....</b>		<b>329</b>

## LISTE DES FIGURES

Figure 1 : Triangle sémiotique d'après Ogden & Richards (1923) .....	33
Figure 2 : Échelle d' <i>Accessibilité</i> d'après Ariel (1990) .....	75
Figure 3 : La <i>Hierarchie du Donné</i> d'après Gundel <i>et al.</i> (1993) .....	76
Figure 4 : Schéma des modalités référentielles d'après Halliday & Hasan (1976) .....	81
Figure 5 : Echelle de <i>phoricité</i> d'après Cornish (2010a) .....	86
Figure 6 : Pluralité interne d'après Curat (1988).....	120
Figure 7 : Pluralité externe d'après Curat (1988).....	120
Figure 8 : Prosogramme 1 .....	238
Figure 9 : Prosogramme 2 .....	239
Figure 10 : Lectures des 3pl sans antécédent d'après Cabredo Hofherr (2014).....	271

## LISTE DES TABLEAUX

Tableau 1 : Flexion des pronoms personnels conjoints en français .....	116
Tableau 2 : Tableau synthétique des traits oppositifs de <i>ils</i> , <i>on</i> et $\emptyset$ .....	320

# Introduction générale

*Ja, kann man ein unscharfes Bild immer mit Vorteil durch ein scharfes ersetzen ? Ist das unscharfe nicht oft gerade das, was wir brauchen ?*  
(Wittgenstein 1953 = 2011 : §71)

## 1. Objet d'étude

La présente étude s'inscrit dans le domaine linguistique de la référence et des expressions référentielles. Ce domaine s'intéresse traditionnellement à la manière dont « le langage entre en relation avec le réel » (Searle 1985 : 236). Les expressions référentielles y apparaissent ainsi comme les moyens d'établir cette relation avec les objets ou référents du monde. Dans les études sur la référence, la plupart largement inspirées des travaux en philosophie, leur fonction est décrite dans une visée essentiellement identificatoire : « [il s'agit] de chercher à comprendre comment, dans un énoncé du type de *j'ai perdu un/le/mon chat* les différents groupes ou syntagmes nominaux dans lesquels apparaît le nom *chat* peuvent ou non conduire à l'identification d'un exemplaire particulier de chat » (Charolles 2002 : 9). Les expressions y sont appréhendées selon des critères d'adéquation et de conformité au réel pour accomplir cette tâche à laquelle elles sont vouées.

Or, ces principes rencontrent des limites face à des actes référentiels courants, comme des faits de sous-détermination (*e.g.* les désignateurs vagues comme (*tout ça, ils* non introduit, l'adverbe *là*, les dénominations postiches comme *le truc, le machin*, etc.). Notre objectif est de confronter ces faits aux théories et postulats existants afin de proposer un modèle plus général pour le traitement des expressions référentielles, qui dépasse les situations « consensuelles » par la prise en compte des productions réelles des usagers. La sous-détermination qui se manifeste en discours, critiquée depuis toujours par les philosophes et grammairiens, soulève en effet d'intéressantes questions à l'égard du traitement de la référence.

Par sous-détermination référentielle, nous entendons la situation où un référent évoqué n'est pas clairement identifié (indistinction de ses attributs, de ses contours, absence de dénomination, de propriété catégorisante, etc.). En examinant ce genre de cas, nous visons ainsi à remettre en question l'idée que les objets que manipulent les locuteurs dans leurs échanges sont ceux de la réalité et proposons qu'il s'agit d'entités construites dans et par le discours. Dans cette perspective, nous montrons que ces « objets-de-discours » ne sont pas toujours bien délimités ni déterminés et que les expressions référentielles ne reflètent pas une adéquation au réel ou aux seuls intérêts de l'interprète à des fins identificatoires, mais plutôt une adéquation à des objectifs communicationnels variés que nous tâcherons de décrire. Les exemples suivants, tirés respectivement de l'oral non planifié et de l'écrit scientifique, donnent un aperçu du genre de faits qui nous intéressent :

- (1) et ce qui me frappe beaucoup c'est que -<sup>1</sup> elles arrivent pas à décoller de l- - de de la même idéologie - c'est-à-dire que elles ont l'impression d'avoir gagné une certaine liberté mais - c'est une liberté qui encore fonctionne dans le quotidien - et je pense que m- - malheureusement je pense **ça** - mais je suis peut-être pessimiste - je pense que le quotidien actuellement ici - maintenant - c'est un truc si frustrant si aliénant - que personne peut épanouir sa personnalité - donc bon c'est c' c'est vrai que c'est pessimiste (oral, ctfp)
- (2) Le rapport « individu-société » est tel que leur dissociation ne peut se concevoir. Cependant, ce rapport a fait problème et fait encore problème pour beaucoup. Comment est-ce possible ? On peut répondre – et **cette réponse** paraît cohérente – que ce problème est un fait sociologique. (Gurvitch, G., *Traité de sociologie* : t. 2, 1968, <Frantext>)

Ces exemples illustrent l'usage d'expressions qui renvoient à des entités sous-déterminées au moment du pointage référentiel – on sait qu'il s'agit respectivement d'un objet de pensée (*je pense ça*) et d'un acte illocutoire (*cette réponse*) – dont la caractérisation est fournie par le discours dans une étape communicative ultérieure. Par ailleurs, cette caractérisation est créée par le discours, et en cela, les objets ne préexistent pas à l'énonciation mais en sont au contraire fondamentalement dépendants.

Les expressions sous-spécifiées, par leur pauvreté sémantique, représentent à cet égard de bons candidats à la diffusion de la sous-détermination : elles sont susceptibles de renvoyer à des référents non catégorisés et « instables ». Leurs traits peu discriminants (et ici leur nature démonstrative) obligent l'interprète à s'appuyer sur des indices de nature contextuelle plutôt que sur leur conformité à des objets de la réalité tangible. Nous avons pour intention d'inventorier les ressources linguistiques, lexicales ou pronominales, spécialisées dans la sous-détermination référentielle et d'en observer les circonstances d'emploi ainsi que les rendements discursifs. Parmi l'ensemble de ces formes, nous avons choisi d'examiner plus en détail deux types de faits représentatifs de cette situation, particulièrement productifs dans les productions spontanées, à savoir l'emploi de *tout ça*, notamment dans les configurations d'énumération, et l'emploi de *ils* à valeur sous-déterminée, souvent rapproché dans la littérature de l'emploi de *on* dit indéfini. En voici deux exemples représentatifs :

- (3) d'accord ces jeunes | \_ |<sup>2</sup> jouent le jeu ils sont | \_ | ils sont enthousiastes ils ils ont de l'énergie ils ont des idées **tout ça** (oral, ofrom)
- (4) ouais c'était vraiment réputé pour être des pistes difficiles | \_ | je sais pas si toi tu as eu mais peut-être plus jeune quand même hein | \_ / **ils ils** travaillaient pas les pistes encore euh c'était vraiment des des champs de bosses tu vois (oral, ofrom)

<sup>1</sup> Selon les conventions de transcription de Blanche-Benveniste *et al.* (2002), ce signe indique une pause brève et lorsqu'il est redoublé, une pause plus longue.

<sup>2</sup> Indication d'une pause dans la base OFROM (Avanzi, Béguelin & Diémoz 2012-2015).

L'emploi de *tout ça* permet au locuteur de rester évasif sur le référent ébauché par l'énumération des qualités. Quant à l'emploi de *ils*, il met en jeu un agent dont l'identité n'est pas pertinente pour l'enjeu de la communication. Les deux expressions ont suscité quelque attention de la part des chercheurs<sup>3</sup>, sans toutefois que la question de la référence soit au cœur des préoccupations ou que des conséquences en soient tirées sur la description des procédés référentiels en général. Notre intention, à l'inverse, consiste précisément à en évaluer les répercussions pour une théorie de la référence. On voit à cet égard d'emblée que l'emploi de *ils* ci-dessus se révèle difficilement conciliable avec la notion d'anaphore telle qu'elle est traditionnellement conçue.

Ce travail offre ainsi l'occasion de revenir sur les notions d'*anaphore* et de *deixis* systématiquement exploitées dans l'étude des procédés référentiels, malgré les défauts de généralité de leur définition. Nos données interrogent aussi bien la conception textuelle de l'anaphore, fondée sur la présence d'un antécédent dans le contexte verbal, encore fortement ancrée dans les grammaires et la littérature en linguistique tous domaines confondus (sémantique, TAL, psycholinguistique, typologie, grammaire générative, etc.) qu'une conception cognitive de l'opposition deixis/anaphore (nouveau vs continuité référentielle). Malgré un vif débat sur l'anaphore au cours des dernières décennies du XX<sup>e</sup> siècle en linguistique française, reflétant des tentatives de renouvellement de la réflexion sur le sujet, les modèles dominants actuels semblent peu enclins à élargir le champ d'observation aux données désormais à disposition. En effet, l'étude des expressions référentielles en français n'a semble-t-il pas profité de l'accès de nos jours facilité aux bases de données orales et écrites pour remettre en question ses propres fondements, contrairement à des domaines comme la syntaxe, la lexicographie, la prosodie, etc. qui ont su en tirer parti pour ajuster leurs modèles ou méthodologies respectives. A l'ère des corpus numérisés, il nous paraît donc opportun de faire le point sur l'état des théories en matière de référence.

## 2. Méthodologie et données

La démarche adoptée dans ce travail est empirique, consistant à recueillir des faits de langue exclusivement attestés en vue d'une analyse minutieuse tenant compte de leur contexte d'occurrence (à l'exception d'exemples empruntés à des auteurs que nous discutons). La méthode se veut qualitative, pour le choix des exemples retenus : ceux-ci ont été récoltés au cours de notre recherche pour leur valeur heuristique, c'est-à-dire pour leur capacité à révéler des aspects pertinents sur le fonctionnement de la référence en français. Nos données ne représentent donc pas à proprement parler un corpus au sens strict du terme, mais une collection d'exemples soigneusement rassemblés sur des critères qualitatifs. En cela, notre travail n'appartient pas à la linguistique *de* corpus, mais constitue une étude linguistique *sur* une sélection de données attestées. Nous faisons en outre le choix de recourir à un ensemble

---

<sup>3</sup> Bilger (1989), Ferré (2011), Secova (2014) pour *tout ça*, Kleiber (1992b), Cabredo Hofherr (2003, 2014), Siewierska (2010, 2011), Siewierska & Papasthati (2011) sur *ils* ou les indices de 3<sup>e</sup> personne du pluriel à travers les langues.

diversifié et ouvert de données. En effet, l'examen des expressions référentielles, dès lors que l'on tient compte des expressions nominales, ne permet tout simplement pas d'exploiter l'extraction automatique de données. D'ailleurs, en se concentrant uniquement sur des faits extractibles, on court le risque de se priver de tout un ensemble de phénomènes qui manifestent des régularités de comportements irréductibles à un repérage automatique (Corminboeuf 2014 : 2378). Nous partons donc dans un premier temps d'une vaste observation à partir de sources diversifiées.

Pour les parties plus empiriques de cette thèse, nous faisons cependant le choix de recourir en priorité à des données d'oral spontané, car ce type de contexte de production reste un champ très peu exploré dans le domaine des expressions référentielles en français. Cela s'explique peut-être par la « grande pauvreté des chaînes conversationnelles qui sont réduites à l'alternance d'un désignateur et de pronoms », comme le relève Corblin (2005 : 253). Mais cette « pauvreté » lexicale des désignateurs en conversation nous paraît au contraire susceptible d'en dire long sur les pratiques référentielles des locuteurs... Pour l'oral, nous recourons aux bases suivantes<sup>4</sup> :

- OFROM (Avanzi, Béguelin & Diémoz 2012-2015) <http://www.unine.ch/ofrom>
- PFC (Durand *et al.* 2002, 2009) <http://www.projet-pfc.net>
- CFPP (Branca-Rosoff, Fleury, Lefeuvre, Pires 2012) <http://cfpp2000.univ-paris3.fr/>
- CTFP (Blanche-Benveniste *et al.* 2002)
- CRFP (Delic 2004)

Les extraits retenus sont reproduits avec leurs conventions d'origine, une tâche d'uniformisation se révélant trop fastidieuse. Consciente que les exemples sélectionnés puissent parfois paraître artificiellement coupés de leur contexte, nous indiquons cependant toujours la source à partir de laquelle est – en principe<sup>5</sup> – accessible le contexte élargi (texte et audio). Nous reproduisons également ponctuellement, pour leur pertinence, des extraits de conversation recueillis « au vol ».

Concernant l'écrit, ce sont nos lectures quotidiennes en tous genres qui constituent d'abord notre source d'observation, à savoir textes de presse (papier ou web), littéraires, pratiques, administratifs, publicitaires, etc. mais également des extraits provenant de genres plus proches de « l'immédiat » (Koch & Oesterreicher 1985) comme les réseaux sociaux, forums, blogs,

---

<sup>4</sup> Le projet ORFEO (<http://www.projet-orfeo.fr/>), encore non accessible, va rassembler et mettre à la disposition du public courant 2017 une masse de données échantillonnées sans précédent en français parlé et écrit.

<sup>5</sup> À l'exception de CRFP, qui n'est pas en accès libre, mais dont nous n'avons reproduit que trois exemples et dont nous pouvons fournir un contexte élargi sur demande.

courriels, SMS<sup>6</sup>. Nous avons également recours à la base *Frantext* ([www.frantext.fr](http://www.frantext.fr)) pour des recherches de formes particulières.

Dans un second temps, une fois ciblés les objets d'étude spécifiques, la récolte des données peut évidemment bénéficier de l'extraction automatique dans les bases respectives (e.g. recherches de (*tout*) *ça*, *ils*, *on*, etc.). Mais les requêtes aboutissant à un nombre trop élevé de résultats pour une analyse contextuelle exhaustive de chaque occurrence, ce sont des critères qualitatifs qui orientent notre tri. En effet, selon nos principes descriptifs, chaque résultat requiert pour son observation la prise en compte d'un contexte large qui peut équivaloir dans certains cas à l'enregistrement entier (ou à des recalibrages de contexte fastidieux dans *Frantext*). Les données sélectionnées ne peuvent donc prétendre à un reflet représentatif de la distribution des occurrences étudiées, quoique nous fournissions ponctuellement quelques comptages effectués sur des sous-corpus restreints. Mais de manière générale, nous ne nous attardons pas sur des faits jugés consensuels ou typiques, et favorisons la prise en compte et l'examen de données laissées en marge des descriptions par les chercheurs, sous prétexte de rareté, de manque de pertinence, voire de déviance. Si certains types de sources paraissent plus crédibles à étudier que d'autres selon des préjugés ambiants, nous accordons à l'inverse à tous les faits la même valeur scientifique. D'ailleurs, malgré l'hétérogénéité des sources, les données révèlent des régularités évidentes sur le fonctionnement de la langue, qu'une épuration pourrait conduire à occulter.

### 3. Organisation du contenu

Cet ouvrage est organisé en deux parties ; la première, constituée de trois chapitres, consiste en un état des lieux des notions clés de référence, d'anaphore et de sous-détermination ; la seconde, composée de deux chapitres, présente deux études de cas.

Les trois chapitres de la première partie représentent des bilans critiques i) sur le traitement vérifonctionnel de la référence et des expressions référentielles par les modèles existants, à la suite de quoi est exposé le cadre théorique dans lequel se situe notre approche ; ii) sur le pronom de 3<sup>e</sup> personne et son rôle unanimement reconnu d'anaphorique ; iii) sur la problématique de la sous-détermination et les ressources à la disposition des locuteurs pour l'exprimer. Dans chacun de ces chapitres, nous confrontons l'état de la recherche à des données authentiques et gênantes pour les théories existantes, afin de mettre en évidence les aspects qui nécessitent à nos yeux plus ample réflexion dans la perspective d'un modèle général de la référence.

Les deux chapitres de la partie empirique portent respectivement sur l'usage de *tout ça*, en particulier dans les structures d'énumération (Chapitre IV) et sur l'emploi de *ils* dont la valeur demeure sous-déterminée (Chapitre V). Ces deux chapitres ont pour enjeu de dégager les

---

<sup>6</sup> Swiss SMS corpus : <https://sms.linguistik.uzh.ch> (Stark *et al.* 2009-2014) ; 88milSMS : <http://88milsms.humanum.fr/index.html> (Panckhurst R., Détrie C., Lopez C., Moise C., Roche M., Verine B. 2014).

facteurs contextuels, accidentels ou stratégiques, qui favorisent la mise en œuvre des procédés de sous-détermination.

\*

Cette étude sur la sous-détermination, qui vise à une meilleure compréhension des mécanismes référentiels en contexte, intéressera les linguistes travaillant dans le domaine de la référence, de la cohérence, de la structure informationnelle et du français parlé. Plus généralement, le sujet se situe aux croisements de plusieurs disciplines, comme la sémantique, la pragmatique, l'analyse du discours et la linguistique interactionnelle. Cette thèse trouvera également des échos parmi les didacticiens de la langue. En effet, on constate dans les manuels ou usuels de grammaire l'insuffisance des ressources disponibles pour l'enseignement des procédés de cohérence, en particulier des expressions anaphoriques, invariablement abordées en termes de dépendance à une séquence textuelle. L'ouvrage pourra encore intéresser les chercheurs en TAL qui travaillent sur la résolution automatique d'anaphores et d'expressions coréférentielles par ordinateur. En effet, la plupart des faits que nous étudions échappent aux algorithmes développés par les ingénieurs en vue de la reconnaissance des expressions et constituent donc une difficulté majeure dans leur travail. Une analyse fouillée du phénomène de la sous-détermination permettra, nous l'espérons, d'apporter de l'eau au moulin des recherches sur la référence.

## **Première Partie :**

*Référence, anaphore pronominale et sous-détermination : état des lieux*



# **Chapitre I**

## **Généralités sur la référence et les expressions référentielles**



## 1. Introduction

Lorsqu'un usager de la langue participe à une interaction verbale, il entraîne généralement dans son discours des objets auxquels il attribue des propriétés (on dit qu'il prédique sur eux) et les met en relation les uns avec les autres. Dans le cadre de cette activité locutoire, l'acte de référence consiste précisément à évoquer dans le discours, par des moyens verbaux, les entités souhaitées. Le choix pour un type d'expression varie selon différents paramètres (l'opération accomplie, la prise en compte du savoir partagé, etc.) :

(5) j'ai perdu **un/le/mon chat** (Charolles 2002 : 9)

L'enjeu des travaux en sémantique référentielle consiste principalement à étudier les modalités de référence à un objet selon l'expression sélectionnée. Ressortissant au domaine de la *parole*, la problématique de la *référence* a d'abord été écartée de l'analyse linguistique « immanentiste » dont l'objet essentiel était la *langue* en tant que système de signes fermé. C'est cependant depuis des siècles un thème incontournable pour les philosophes, qui cherchent à « comprendre comment le langage entre en relation avec le réel » (Searle 1985 : 236). L'intérêt pour la notion de *référence* est toutefois apparu dans le champ de la linguistique avec le développement des études sur l'*énonciation*, autrement dit sur l'actualisation de la langue en contexte par les usagers, et avec l'essor de la *pragmatique* et de la *sémantique* comme disciplines dès lors non plus restreintes au domaine de la philosophie.

Ce chapitre vise à examiner la notion de référence et ses notions voisines (§2), tout en montrant qu'elles restent largement tributaires des conceptions philosophiques, fondées notamment sur l'idée d'une conformité des signifiés linguistiques aux objets de la réalité tangible. Nous présentons les tentatives déployées en sémantique pour classer ces référents (§3) et nous dressons une synthèse critique sur la manière dont les principales expressions référentielles sont traitées en linguistique (§4). Puis nous abordons, toujours d'un œil critique, les différentes modalités de désignation régulièrement invoquées dans l'acte de référence, à savoir l'*anaphore* et la *deixis* (§5). Face aux approches dominantes en matière de référence, nous présentons pour finir une approche alternative et constructiviste de la question, incarnée par le modèle que nous prenons pour cadre d'analyse (§6).



## 2. Notions fondamentales

### 2.1 *Référence, existence et monde réel*

D'après Frege (1892=1971) et ses continuateurs, un nom propre n'a de référence (ou *dénotation*, dans sa terminologie) que s'il désigne un objet pour lequel on suppose une existence avérée dans le monde (*cf. infra* §2.2). Autrement dit, les référents constituent les objets existants du monde, qu'il est possible de prendre comme objets du discours en y référant par des moyens linguistiques. De la sorte, le nom *Vénus* renvoie bien à un objet extralinguistique, la planète ainsi nommée et observable, mais il n'en va pas de même pour le nom *Ulysse*, qui n'a pas de référence établie dans le monde qui nous entoure, donc pas de référence du tout selon la conception frégréenne.

Dans une optique résolument inverse, certains considèrent que tout objet que l'on évoque dans le discours implique son existence : « whatever is referred to must exist » (Searle 1969 : 77). C'est ici l'acte de référence linguistique même qui confère une existence, dès lors indéniable, à l'objet visé.

Le compromis généralement adopté face à ces positions antagonistes consiste à considérer que des noms tels qu'*Ulysse*, *licorne* ou autres *Tarzan* renvoient bien à des référents, mais à des référents existant dans un monde imaginaire (Kleiber 1997a : 11). La notion d'existence doit alors être adaptée à cet effet :

Existence is a tricky concept in any case, and we must allow for various kinds of existence pertaining to fictional and abstract referents (or, alternatively, show how these apparently diverse kinds of existence relate to the physical existence of spatiotemporally continuous and discrete objects) (Lyons 1977 : 183)

Le discours aurait donc le pouvoir, de ce point de vue, de construire des mondes alternatifs à celui dans lequel nous vivons, de référer à des mondes possibles où se rencontrent non seulement des référents fictifs, mais aussi des situations qui pourraient paraître insensées dans le monde réel et tangible :

[...] there is no single metaphysical, or conceptual, framework which underlies every kind of human discourse. Statements, or propositions, which might be held to be contradictory, or absurd, in a more or less scientific discussion of the physical world may be regarded as perfectly acceptable in a mythological or religious context, in poetry, in the narration of a dream, or in science fiction (Lyons 1977 : 167)

L'axiome d'existence des référents dont on parle, de même que la conformité de la langue au monde se voient donc préservés moyennant la reconnaissance de capacités cognitives *ad hoc* de la part des usagers de la langue (Kleiber 1997a : 11). Dans cette perspective, le monde réel que les usagers perçoivent reste cependant le monde privilégié qui sert de point de référence

pour la conception de leurs représentations<sup>7</sup>. Ce n'est, selon Kleiber, qu'à travers sa reconnaissance qu'il leur est possible d'envisager d'autres mondes possibles et de référer à des éléments fictifs, qui, après tout, sont constitués de « morceaux "réels" : animal, cheval, corne, etc. » (*ibid.* : 15, *cf.* aussi Lyons 1977 : 211). La solution de nos jours largement admise est ainsi que les référents dont on parle ont bien une existence, soit dans le monde qui nous entoure (situation habituelle), soit dans un autre monde possible. Cette conception des mondes possibles multiplie des *versions* de la réalité afin de préserver un principe de vérité (*cf. infra* §2.2), consistant à garantir l'existence des référents dans notre monde réel, ou, à défaut, dans un monde possible. Elle suppose que chaque référent a son corrélat linguistique préexistant au-dehors, à travers une sorte de bi-univocité entre réalité tangible (ou virtuelle) et langue.

Néanmoins, on peut objecter contre cette vision des choses que la notion de *mondes possibles* a été conçue pour les besoins de l'analyse logique et qu'elle « n'appartient pas au métalangage de la linguistique » (Gary-Prieur 1994 : 21). Si l'on se donne pour objectif la description scientifique des usages effectifs des locuteurs, il n'y a pas lieu de juger de l'adéquation d'un référent à un univers *ad hoc*. Il suffit de constater que le discours construit lui-même ses référents, aussi fictifs ou fantaisistes soient-ils, dans son univers propre, comme c'est le cas du fameux *Schmilblick* de Pierre Dac dans l'extrait ci-dessous :

- (6) C'est dans la nuit du 21 novembre au 18 juillet de la même année que les frères Fauderche ont jeté les bases de **cet extraordinaire appareil** dont la conception révolutionnaire risque de bouleverser toutes les lois communément admises tant dans le domaine de la physique nucléaire que dans celui de la gynécologie dans l'espace [...] **Le Schmilblick des frères Fauderche** est, il convient de le souligner, rigoureusement intégral, c'est-à-dire qu'il peut à la fois servir de Schmilblick d'intérieur, grâce à la taille réduite de ses gorgomoches, et de Schmilblick de campagne [...]. (extrait du sketch de Pierre Dac, 1950)

A notre sens, il n'est pas nécessaire de postuler l'existence d'un monde imaginaire car il n'y a pas de véracité des propos à vérifier : l'objet, dès lors qu'on en parle, existe tout simplement dans l'univers construit par le discours à l'œuvre (*cf.* Searle ci-dessus), comme n'importe quel objet auquel il est fait référence dans cette même situation de parole (*e.g.* les *frères Fauderche*, les *gorgomoches*). On peut ainsi en parler comme n'importe quel autre objet du discours et prédiquer des propriétés, aussi loufoques soient-elles, à son égard. L'existence des objets dont on parle est donc une existence posée dans le discours et celle-ci rend caduque la question de la conformité avec le réel ou l'imaginaire.

---

<sup>7</sup> *Cf.* Fauconnier (1984), pour qui « l'espace réel »  $E_0$  est celui par rapport auquel se situent tous les autres espaces mentaux.

## 2.2 Sens et référence

La problématique du rapport entre l'emploi d'une expression linguistique et sa référence soulève inévitablement la question du sens de ce signe. Depuis des siècles, les philosophes s'interrogent sur la manière dont la langue permet de désigner un « morceau » de la réalité. C'est généralement le *sens* qui est invoqué comme responsable de la bonne détermination référentielle. Le sens est donc étroitement lié à la notion de référence. Nous retraçons ici dans les grandes lignes, sur la base de travaux de spécialistes, les conceptions principales de cette relation au cours des siècles.

On considère souvent ce passage du *Peri Hermeneias* d'Aristote comme fondateur de la théorie du signe de la linguistique moderne (Panaccio 1996 : 6-7) :

Les sons émis par la voix sont les symboles des états de l'âme, et les mots écrits des symboles des mots émis par la voix. Et de même que l'écriture n'est pas la même chez tous les hommes, les mots parlés ne sont pas non plus les mêmes, bien que les états de l'âme dont ces expressions sont les signes immédiats soient identiques chez tous, comme sont identiques aussi les choses dont ces états sont les images (Aristote, *De l'interprétation* 16a2-8, trad. Tricot)

Aristote conçoit les « états de l'âme » (παθήματα τῆς ψυχῆς<sup>8</sup>) ou, pourrait-on dire, les impressions ressenties par les sujets, comme des images ressemblant (ὁμοιώματα<sup>9</sup>) aux choses du monde. Ils constituent en quelque sorte leur contrepartie cognitive et comme elles, se distinguent par leur stabilité et leur universalité, là où leurs symboles acoustiques ou graphiques (le code écrit étant voué à signifier l'oral) sont sujets à la variation interlinguistique et à l'arbitraire. Quoi qu'il en soit, on perçoit bien l'idée d'une médiation conceptuelle<sup>10</sup>, cognitive, entre le signe et les objets du monde.

Parmi les travaux de l'Antiquité gréco-latine touchant au processus de signification, Rey (1973) note la conception particulièrement développée des Stoïciens, à l'origine de l'élaboration d'un véritable « appareil conceptuel » (p. 29). Il cite à cet égard le témoignage de Sextus Empiricus :

Les Stoïciens disent que trois choses sont liées : ce qui est signifié, ce qui signifie et l'objet. De ces choses, celle qui signifie c'est la parole par exemple « Dion » : ce qui est signifié c'est la chose même qui est révélée par elle et que nous saisissons comme durable par notre pensée, mais que les Barbares ne comprennent pas, bien qu'ils soient capables d'entendre le mot prononcé, alors que l'objet est ce qui existe à l'extérieur : par exemple, Dion en personne. (*Adversus Mathematicos*, VIII, 11-12)

Autrement dit, la parole signifie et révèle quelque chose de conventionnel dans la pensée, cette fois propre à la langue en question, qui s'oppose à l'objet même existant en dehors de la

---

<sup>8</sup> *pathémata tês psuchês*

<sup>9</sup> *homoiómata*

<sup>10</sup> Selon Panaccio (1996 : 7), les « états de l'âme » comprennent vraisemblablement les concepts (νοήματα).

langue. On peut donc entrevoir à travers ces propos les prémisses d'une théorie du signe qui sera remise au goût du jour des siècles plus tard (cf. ci-après le triangle sémiotique).

A partir des questions sémantiques soulevées notamment par les philosophes de l'Antiquité, les scolastiques du Moyen Age continuent de s'interroger sur le processus de signification. Une question est notamment de savoir si les mots signifient des concepts ou des choses. Fondée sur la conception transitive d'Aristote, la position courante au Moyen Age est de considérer que les termes signifient directement des concepts, et indirectement des choses. Le premier processus est appelé *significatio*, tandis que le second relève de la *suppositio* : si chaque occurrence d'un mot *signifie* le même concept, elle *suppose* un individu différent à chaque fois (Read 2011). Les termes reçoivent toutefois des conceptions diverses selon les auteurs médiévaux<sup>11</sup>. Ce qui semble néanmoins faire consensus est la primauté de la *significatio* sur la *suppositio* : on n'accède à la *suppositio* qu'en vertu de sa *significatio*, qui détermine les objets auxquels il est fait (ou possible de faire) référence.

Cette distinction sémantique perdure et se voit systématisée dans la *Logique de Port-Royal* d'Arnauld & Nicole (1662=1874 : 54 sqq.), ouvrage qui a exercé une influence considérable sur les travaux en logique jusqu'à la fin du XIX<sup>e</sup> siècle. L'unité sémantique de base de la théorie d'Arnauld & Nicole est l'*idée*, par laquelle est signifié le contenu objectif de la pensée (Buroker 1993 : 457). La sémantique de Port-Royal établit des relations entre les mots, les idées et les choses de manière transitive, à l'instar de la « sémiotique » aristotélicienne : les mots employés pour exprimer des idées signifient également les choses signifiées par les idées (*ibid.* : 464). Cela dit, les auteurs de la *Logique* ne manquent pas de relever que la correspondance entre les mots et les idées est imparfaite : il n'y a pas de garantie que la structure de la langue parlée reflète adéquatement la structure logique des idées (Arnauld & Nicole 1662=1874 : 68-87). Dans leur classification des *idées*, Arnauld & Nicole opposent les idées singulières aux idées générales. Les premières servent à représenter un seul individu (via un « nom propre », e.g. *Socrate, soleil*), alors que les secondes se montrent capables de représenter plusieurs choses (via un nom commun ou un adjectif, e.g. *homme, humain, humanité*). Les idées générales peuvent ainsi représenter une classe d'individus possédant l'attribut exprimé, autrement dit son *étendue*. C'est au cœur de cette problématique que les auteurs établissent la distinction entre *compréhension* et *étendue* (ou *intension* et *extension*<sup>12</sup>), préfigurée à travers les concepts médiévaux de *significatio* et *suppositio*. La *compréhension* d'une idée se définit comme « les attributs qu'elle enferme en soi, et qu'on ne peut lui ôter sans la détruire », son *étendue* comme « les sujets à qui cette idée convient ; ce qu'on appelle aussi les inférieurs d'un terme général, qui, à leur égard, est appelé supérieur, comme l'idée

---

<sup>11</sup> Pierre d'Espagne distingue par exemple la *suppositio naturelle* de la *suppositio accidentelle*, la première correspondant à l'ensemble d'objets auxquels s'applique potentiellement le mot, la seconde équivalant à la référence effective d'un terme dans une proposition particulière. Chez Guillaume d'Ockham, c'est la *significatio* qui recouvre l'ensemble des choses pouvant être caractérisées par le mot (Read 2011).

<sup>12</sup> Les appellations *intension* et *extension* (e.g. Carnap 1947) l'emportent aujourd'hui en sémantique sur celles de *compréhension* et *étendue*. Par ailleurs, les concepts recouvrent *grosso modo* ceux que Mill (1843=1988) nomme *connotation* (à ne pas confondre avec sa signification usuelle de description subjective) et *dénotation*.

du triangle en général s'étend à toutes les espèces diverses de triangle » (*ibid.* : 54). Il découle de cette théorie sémantique la fameuse « loi de Port-Royal » (Aroux 1992) selon laquelle compréhension et étendue varient en proportion inverse : l'ajout d'attributs entraîne la restriction de l'étendue de l'idée, de même que la réduction d'attributs en augmente l'étendue. Ainsi, c'est la compréhension qui détermine l'étendue d'une idée.

Père de la logique moderne avec sa conception du calcul des prédicats, Gottlob Frege contribue également à la philosophie du langage et à la sémantique par la définition de notions dont on se sert amplement dans les théories actuelles du sens ainsi que par leur inscription dans une sémantique vériconditionnelle. C'est dans une tentative de résolution du problème des énoncés d'identité non tautologiques qu'il fait intervenir la notion de *sens* (*Sinn*), par opposition à celle de *dénotation* (*Bedeutung*) (1892=1971).

Sa contribution s'inscrit dans un projet scientifique dont l'objectif est avant tout l'élaboration d'une notation formelle des énoncés mathématiques, dont on rendait compte jusqu'alors par le recours à la langue naturelle : Frege s'efforce ainsi de symboliser, au moyen d'un système logique formel appelé *Begriffsschrift*, les tournures linguistiques usuelles des mathématiques, jugées imprécises<sup>13</sup>. Le logicien distingue trois niveaux de la structure sémantique, à savoir les signes, leur dénotation et leur sens :

[...] il est naturel d'associer à un signe (nom, groupe de mots, caractères), outre ce qu'il désigne et qu'on pourrait appeler sa dénotation, ce que je voudrais appeler le sens du signe, où est contenu le mode de donation de l'objet.  
(1892=1971 : 103)

Frege assimile l'objet désigné par un signe à sa dénotation et considère le mode de donation comme une composante du sens. Il illustre ces distinctions par le célèbre exemple des noms propres *étoile du matin* et *étoile du soir*, qui possèdent la même dénotation, en l'occurrence, la planète *Vénus* en tant qu'objet désigné, mais dont le sens diffère par le mode de donation : autrement dit, le même objet n'est pas présenté de la même manière selon l'expression utilisée ; chacune implique des conditions d'emploi différentes<sup>14</sup>. Contrairement au sens d'un signe, une dénotation n'est, selon Frege, pas toujours associée à ce dernier. Il prend à cet effet l'exemple de l'expression *le corps céleste le plus éloigné de la terre*, qui, malgré l'expression d'un sens, ne possède pas de dénotation (p. 104). La dénotation, au sens de Frege, est donc un concept fondé sur l'existence avérée des objets du monde qui nous entoure (*cf. supra* §2.1) : « [l]a dénotation d'un nom propre est l'objet même que nous désignons par ce nom » (p. 106). Cette quête de l'authenticité se voit exprimée à travers la question de la dénotation des propositions : « c'est donc la recherche et le désir de la vérité qui nous poussent à passer du sens à la dénotation » (p. 109). Pour Frege, la dénotation d'une proposition consiste en sa valeur de vérité, c'est-à-dire, le fait qu'elle soit vraie ou fausse (p. 110). Cette conception

---

<sup>13</sup> « Presque toujours le langage ne donne pas, sinon allusivement, les rapports logiques ; il les laisse deviner sans les exprimer proprement. » (Frege 1892=1971 : 65)

<sup>14</sup> Il en va de même pour les notations algébriques «  $2^4$  » et  $4 \times 4$  », qui « sont des noms propres du même nombre mais n'ont pas le même sens » (*ibid.* : 89).

faisant de la vérité un enjeu de la théorie de la signification va servir de point de départ au développement d'une sémantique dite vériconditionnelle initiée par Tarski (1935), dans un effort de précision de la notion de vérité. C'est dans cette perspective que vont s'interpréter les « conditions de satisfaction » auxquelles doit répondre le référent : « The condition that the referent must satisfy the description has commonly been interpreted by philosophers to imply that the description must be true of the referent » (Lyons 1977 : 181). Comme nous l'avons déjà évoqué (*supra* §2.1, *cf.* aussi *infra* §6.3.1), nous prenons nos distances avec une approche évaluant le degré de conformité d'un contenu au monde.

Parallèlement, une perspective oppositive et « négative » de la signification apparaît avec la théorie du signe linguistique développée par Ferdinand de Saussure au début du XX<sup>e</sup> siècle. Selon le linguiste genevois, un signe se définit par l'association d'une image acoustique (signifiant) et d'un concept (signifié), solidaires comme le recto et le verso d'une feuille (1916=2008). Le contenu invoqué est appréhendé en termes de « valeur », qui ne se conçoit que par opposition à d'autres valeurs : c'est uniquement la nature différentielle des signes qui procure à chacun une existence au sein d'un tel système (2002 : 28). Quant au référent, qui relève de l'activité de parole et de l'extralinguistique, il n'est pas fondamentalement apparié au signe : « le signe linguistique unit non une chose et un nom, mais un concept et une image acoustique » (1916=2008 : 98). Les héritiers de la sémantique structurale s'efforcent, en théorie, de maintenir cette limite :

Les mots de la langue ne sont pas en relation immédiate avec les choses. Ils ont un signifié, mais n'ont pas de référence. Le mot *maison*, en tant que mot de la langue, ne dit pas une maison. Le mot *cheval* ne dit pas un cheval ; il n'est en relation immédiate avec aucun être concret ; il exprime seulement un ensemble de propriétés, et la question de savoir s'il existe dans l'univers des êtres auxquels appartiennent ces propriétés n'a aucune pertinence pour le lexique. (Le Guern 2003 : 32)

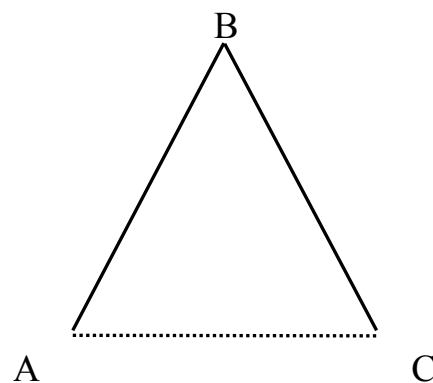
Néanmoins, dans les faits, décrire le signifié en faisant abstraction du référent n'est pas une mince affaire. Il s'agit de mettre au jour les traits distinctifs (ou *sèmes*, *cf.* Greimas 1966 : 22) d'un signe relativement à un ensemble de signes, autrement dit, d'identifier des éléments abstraits du signifié dont la composition (le *sémème*) s'opposerait à celle d'autres signes voisins. Cependant, l'immixtion du référent est difficile à éviter et il n'est pas rare que l'analyse componentielle (ou *sémique*) se confonde paradoxalement avec l'observation des référents des lexèmes, c'est-à-dire avec l'énumération des propriétés à partir de la réalité extralinguistique. Ainsi en va-t-il de la fameuse analyse de Pottier (1963) du champ sémantique des noms de sièges (*chaise, fauteuil, tabouret, canapé, pouf*), qui aboutit à l'émergence de traits tels que *avec dossier, avec bras, sur pied*, etc., dont la pertinence ne relève vraisemblablement pas de la langue, mais, comme le relève Moussy (1991 : 65), de critères perceptibles en situation.

Il est clair que l'existence avérée et tangible d'objets « correspondant » aux lexèmes choisis (ici des types de sièges) pour l'analyse tend à court-circuiter l'analyse oppositive des signifiés

eux-mêmes. Afin d'éviter ce genre de transposition référentielle, Le Guern (2003 : 37) suggère carrément de travailler sur le sens figuré des mots<sup>15</sup>. Saussure lui-même (1916=2008 : 160) proposait d'étudier l'opposition entre des « synonymes » – du genre *redouter*, *craindre*, *avoir peur* – à partir desquels il est plus ardu d'identifier des corrélats tangibles.

Mais de manière générale, l'intervention du référent dans le processus de signification figure désormais en bonne place dans les schématisations traditionnelles du signe linguistique postérieures à Saussure. Celles-ci prennent régulièrement la forme d'un triangle (dit *sémiotique*) : leurs sommets incarnent généralement A) le signe du point de vue de sa forme, B) sa composante conceptuelle et C) l'objet extra-linguistique, ainsi qu'illustré ci-dessous, par la figure adaptée d'Ogden & Richards (1923 :11) :

**Figure 1** : Triangle sémiotique d'après Ogden & Richards (1923)



La relation entre A et C, autrement dit la relation de référence entre une expression linguistique et l'objet désigné, est généralement considérée comme indirecte (rendue par les pointillés), tandis que les relations AB et BC sont présentées comme plus fondamentales du processus de signification (*cf.* conception transitive *supra*). D'un auteur à l'autre, les conceptions, extensions du phénomène et terminologies peuvent sensiblement varier. Ces variations importent cependant peu pour notre propos<sup>16</sup>. Ce schéma, que l'on pourrait tenir pour traditionnel du processus de signification, représente bien la relation triadique déjà examinée par les philosophes au cours de l'histoire et que résume, de manière concise, la maxime scolastique *vox significat [rem] mediantibus conceptibus* (Lyons 1977 : 96)

\*

<sup>15</sup> « Dans l'emploi d'un mot au sens propre, il n'est guère possible de distinguer les éléments de signification qui relèvent strictement de la linguistique – les composantes du signifié lexical – et ceux qui sont liés à la connaissance de l'objet extralinguistique. Déterminer le signifié du mot *chat*, en faisant abstraction de l'animal concret, n'est pas possible à partir des emplois où le mot *chat* désigne un chat. L'autonomie du signifié par rapport à la référence n'est perceptible que dans les emplois où *chat* sert à désigner autre chose qu'un chat. » (Le Guern 2003 : 37)

<sup>16</sup> Pour une synthèse, voir Lyons (1977 : 95-99).

La question du sens et de la référence, ainsi que de leurs notions voisines, est aujourd'hui très marquée par ces influences antérieures à fondement logico-philosophique. Afin de mettre un peu d'ordre parmi ces termes diversement employés par les philosophes (Lyons 1977 : 174), nous proposons ci-dessous un récapitulatif des notions les plus répandues de nos jours et de leur(s) acception(s) en sémantique référentielle, à l'égard desquelles, pour certaines, nous avons déjà pris quelque distance méthodologique que nous expliciterons de manière détaillée dans le bilan qui clôt cette section (§2.4).

Le *sens* d'un signe est considéré comme un contenu mental, qui est diversement appréhendé en termes de *concept*, *signifié*, *intension*, *compréhension* ou encore *référence (sic) virtuelle* (cf. Milner 1982 : 10). Dans la plupart de ses acceptions, il sert à déterminer l'objet qui « tombe » sous lui, il indique son « mode de donation » (Frege 1892=1971). Ce contenu est souvent envisagé en termes de conditions, par exemple chez Milner (*ibid.*), qui le définit comme « un ensemble de conditions que doit satisfaire un segment de réalité pour pouvoir être la référence d'une séquence où interviendrait crucialement l'unité lexicale en cause ».

La *dénotation* est la capacité pour un lexème de délimiter un certain nombre d'objets qui satisfont à son sens. Elle détermine l'*extension*, l'*étendue*, ou encore le *denotatum* du lexème (Lyons 1977 : 207), c'est-à-dire l'ensemble des objets auxquels le lexème peut s'appliquer correctement<sup>17</sup>. Chez Frege, la notion de *dénotation* (traduction de *Bedeutung*) s'applique non pas aux unités lexicales, mais aux noms propres, autrement dit aux désignateurs, et consiste en l'objet ainsi désigné.

La *référence* d'un signe consiste en la relation de désignation effective entre un signe linguistique et un objet extra-linguistique. C'est une notion dépendante de l'acte d'énonciation, contrairement à celles de sens ou de dénotation (du point de vue lexical) (Lyons 1977 : 176). Dès lors, la référence n'est pertinente que dans des contextes d'actualisation de la parole, c'est-à-dire qu'elle ne s'applique qu'à des occurrences d'expressions utilisées en contexte. Par opposition à la référence virtuelle, Milner (1982 : 10) nomme *référence actuelle* ce « segment de réalité associé à une séquence ». On dote généralement l'acte de référence d'une visée *identificatoire* d'un objet de la réalité (e.g. Lyons 1977 : 177, Charolles 2002 : 9).

Le *référent* est donc l'objet extralinguistique qui correspond – selon les écoles, potentiellement ou effectivement – à une expression référentielle. Pour certains (entre autres Frege, cf. *supra* et Russel (1919 : chap. XVI)), seuls peuvent être appelés référents les objets appartenant au monde environnant. A l'inverse, certains considèrent que l'acte de référence présuppose l'existence du référent (cf. Searle, *supra*) : « [d]ès lors que quelqu'un parle de quelque chose cela suffit pour que celles et ceux à qui il s'adresse soient mis en demeure de supposer que cette chose existe [...] » (Charolles 2002 : 32). Pour contourner la difficulté

---

<sup>17</sup> « We will use the term denotatum for the class of objects, properties, etc., to which the expression *correctly* applies. » (Lyons 1977: 207, nous soulignons)

posée par les expressions sans corrélats extra-linguistiques, l'alternative adoptée est la notion de *possible* (*êtres, univers* ou *mondes*), rendant compte de l'existence du référent dans « au moins une version du monde » (*ibid.* : 36).

### 2.3 Référence et prédication

Depuis les apports de la logique frégréenne, on oppose généralement les *expressions référentielles*, destinées à *référer* à (ou identifier) ce dont on parle, aux *expressions prédicatives*, qui servent à *assigner des propriétés* particulières à cet objet identifié (Lyons 1977 : 23). On dit des secondes qu'elles *prédiquent* sur un référent ou que par leur biais des propriétés sont prédiquées à propos d'un référent :

(7) Chopin était polonais (Charolles 2002 : 23)

Dans l'exemple (7), le locuteur réfère à un individu particulier au moyen du nom propre *Chopin*, auquel il attribue la propriété d'être polonais. On dit du SN *Chopin* qu'il sert de support à la prédication (ou de thème au propos) (*ibid.*).

Bien que Frege ait formalisé cette bipartition, c'est en fait à Aristote (*Organon*) qu'on doit la distinction entre sujet et prédicat : le sujet et le prédicat sont les deux *termes* (< du latin *terminus*, *i.e.* les éléments terminaux de l'analyse logique, par opposition à l'acception linguistique) qui entrent en relation pour former une proposition. A travers la notion de prédicat, Aristote vise à établir une relation par rapport à laquelle se situe le sujet. Une proposition est envisagée, dans cette perspective, de manière essentiellement binaire : les termes y représentent, respectivement, 'le sujet' et 'ce qu'on dit du sujet'.

Dans la logique classique néanmoins, les expressions se répartissent plutôt selon leur individualité ou leur généralité (*e.g.* Arnauld & Nicole 1662=1874 : ch. VI, voir *supra* §2.2), *i.e.* les noms propres exprimant une idée singulière *vs* les termes communs marquant une idée générale à laquelle correspond un ensemble d'individus. La nouveauté de Frege est de formaliser le prédicat comme un élément non saturé, nécessitant un objet qui « tombe sous le concept ».

La distinction entre sujet et prédicat est régulièrement transposée au plan syntaxique sur celle entre SN sujet et SV, dont l'association constitue une proposition simple (Lyons 1977 : 430). Cette analyse bi-partite de la phrase est profondément ancrée dans la grammaire traditionnelle et elle est même considérée comme un phénomène universel :

There must be something to talk about and something must be said about this subject of discourse once it is selected [...]. The subject of discourse is a noun. As the most common subject of discourse is either a person or a thing, the noun clusters about concrete concepts of that order. As the thing predicated of a subject is generally an activity in the widest sense of the word [...] the verb clusters about concepts of activity. No language wholly fails to distinguish

noun and verb though in particular cases the nature of the distinction may be an elusive one. (Sapir 1921: 117)<sup>18</sup>

Il n'est cependant pas rare de rencontrer des faits échappant à cette distinction. On peut prendre à témoin les énoncés dits « impersonnels » :

(8) j'aimerais bien aller en Ecosse *mais bon il pleut toujours* (oral, ofrom)

Il est inconcevable de sélectionner un objet auquel appliquer le prédicat 'pleuvoir' : un tel énoncé exprime tout simplement un procès sans agent. A l'inverse, un SN peut constituer à lui seul une énonciation<sup>19</sup>, sans faire l'objet d'une quelconque prédication :

(9) - Attention ! *Vos gants* ! s'écriait Mme Tison. Trop tard. Le rouge à lèvres avait laissé sa vilaine traînée... (Garat, A.-M., *Pense à demain*, 2010 < Frantext)

Dans l'exemple ci-dessus, une locutrice dirige l'attention de son allocutaire et la met en garde contre une situation implicite dans laquelle est impliqué le référent (ses gants) : elle invite à adopter une attitude particulière à l'égard de ce dernier – explicitée via l'interjection *Attention !* – sans pour autant lui octroyer une propriété particulière. Nous aurons l'occasion de revenir *infra* (§3.1) sur d'autres problèmes que soulève cette approche binaire.

Là où Aristote se concentre sur ce qui est prédiqué du sujet, peu importe le nombre d'objets impliqués, la logique moderne conçoit les prédicats comme mettant en relation un certain nombre d'individus :

(10) Chopin a composé des valse (Charolles 2002 : 24)

Dans cet énoncé, on peut considérer, outre l'attribution à 'Chopin' de la propriété de 'composer des valse', que l'expression prédicative *a composé* met en relation un individu ('Chopin') avec 'un ensemble de valse'. Le calcul logique des prédicats élaboré par Frege vise à formaliser en ces termes la structure interne des propositions simples. Dans la tradition logique atomiste et empiriste, le calcul des prédicats est considéré comme représentant correctement la forme logique sous-jacente des énoncés linguistiques par la mise en correspondance de celle-ci avec la structure des états-de-choses dans le monde réel (Lyons 1977 : 147-148). Depuis la naissance du calcul des prédicats, on distingue – quantifieurs mis à part – les *noms des prédicats*, les premiers référant à des individus, les seconds se présentant comme des opérateurs au moyen desquels on peut construire une proposition à partir des noms. Une proposition simple se traduit par une fonction appliquée à son ou ses argument(s), les noms servant d'arguments et le prédicat de foncteur. Par exemple, on peut rendre par la notation  $G(j)$  l'énoncé *John est grand*, où  $j$  représente l'individu 'John' et  $G$  le prédicat 'grand'. Selon le nombre d'arguments que requiert un prédicat, celui-ci est qualifié d'unaire

---

<sup>18</sup> Lyons critique l'imprécision terminologique de cet extrait où *sujet* est utilisé aussi bien pour désigner le référent que l'expression référentielle et où fait défaut une distinction entre les notions de *nom* et d'*expression nominale*, de même qu'entre celles de *verbe* et d'*expression verbale* (Lyons 1977 : 429).

<sup>19</sup> On parle dans ce cas de *nominativus pendens* ou de *hanging topic* (Groupe de Fribourg 2012 : 173).

(ou monadique), de binaire (ou dyadique), de ternaire (triadique), etc. Ainsi, la proposition 'x aime y' se notera  $A(x, y)$  (*ibid.* : 149). Au plan syntaxique, cette vision qui considère le verbe comme le noyau autour duquel gravitent les autres constituants se retrouve par exemple dans la théorie de la valence de Tesnière (1959).

Quelle que soit l'approche choisie, il n'en demeure pas moins que la catégorie nominale est clairement associée à la fonction référentielle, tandis que la catégorie verbale est liée à la fonction prédicative. Il a néanmoins été remarqué que certains SN ne manifestaient pas dans tous les contextes une fonction référentielle. Lyons (1977) illustre cela à partir de l'énoncé suivant, qui peut s'analyser de la même manière dans sa version française :

(11) Giscard d'Estaing is the President of France (p. 185)

Dans une première interprétation de l'énoncé, on peut attribuer à l'individu 'Giscard d'Estaing' la propriété 'président de la France'<sup>20</sup>. Le SN *the President of France* n'a pas pour vocation de référer à un individu, mais bien de prédiquer quelque chose de l'individu nommé Giscard d'Estaing, d'où une interprétation prédicative pour ce SN. A ce titre, ce prédicat peut être remplacé par un autre de même sens, par exemple le SV *gouverne la France* (Apothélos 1995a : 26). Mais l'énoncé (11) peut être interprété différemment : la copule *be* permet une lecture équative, établissant une relation d'identité entre deux individus, le SN *the President of France* induisant dès lors une interprétation dite référentielle<sup>21</sup>. Dans ce cas, les SN de l'énoncé sont interchangeables et le déterminant défini est obligatoire en français, tandis que ce n'est pas le cas dans la lecture prédicative (*ibid.*). La distinction entre emplois « attributif » et « référentiel » a fait l'objet de nombreuses discussions et critiques (entre autres Kleiber 1981, Récanati 1983, Galmiche 1983) qui n'aboutissent pas à un consensus mais qui ont le mérite de montrer que les SN sont susceptibles de questionner la relation entre rôles référentiel et prédicatif. En tous les cas, il en résulte que la frontière n'est pas directement transposable à partir de la syntaxe, chaque occurrence en contexte requérant un traitement particulier.

## 2.4 Bilan

Il ressort de cet exposé des notions fondamentales en matière de référence un certain nombre de postulats sous-jacents qui restent solidement ancrés dans la tradition sémantico-référentielle telle qu'elle est pratiquée de nos jours :

- a) Les expressions référentielles ont des *corrélats extralinguistiques* réels (ou virtuels) *préexistant* dans le monde ou dans l'une de ses versions possibles.

---

<sup>20</sup> Donnellan (1966) appelle ces emplois « attributifs ».

<sup>21</sup> La première interprétation pourrait se dégager par exemple de la réponse à la question : *Quelle fonction assume Giscard d'Estaing ?*; la seconde à celle-ci : *Qui est le président de la France ?*

- b) Pour qu'un acte de référence soit réussi, les référents doivent *vérifier* le sens de l'expression (le sens étant appréhendé en termes de *conditions de satisfaction*), autrement dit, les attributs exprimés doivent être jugés *vrais* du référent.
- c) Les expressions référentielles sont vouées à *identifier* un référent bien *déterminé*.

Néanmoins, un certain nombre de faits linguistiques posent des difficultés à l'égard de ces postulats, parmi lesquels ces deux exemples attestés tirés de l'écrit :

- (12) En revanche, ce que vous pouvez faire pour limiter la consommation journalière (et **cette recommandation** concerne l'ensemble de la famille), c'est préparer les desserts avec des édulcorants. (M. Montignac, *Je mange donc je maigris !*, p. 186)
- (13) Constantinople, 1919. Cris, vociférations. Effluves de tabac turc, vapeurs d'alcool. Castagne. Un bar un peu glauque, des marins français d'un côté, des Anglais de l'autre. De temps en temps, **ça** explose. Bagarre générale. (Verlant, *Gainsbourg*, p. 11 (incipit))

L'exemple (12) illustre un procédé de référence à l'acte d'énonciation lui-même (l'acte de 'recommander'), en cours de réalisation, en même temps qu'au contenu de cet acte (ce qui est recommandé) dont la caractérisation est en suspens au moment de la désignation. En effet, l'information pertinente de l'objet visé n'est délivrée que via le second mouvement de la construction pseudo-clivée, à savoir 'préparer les desserts avec des édulcorants'<sup>22</sup>. Cet exemple démontre bien la nature construite de l'objet considéré, qui est le fruit du discours en cours produit par le locuteur dans un contexte particulier (conseil diététique). Il paraît dès lors peu vraisemblable de considérer ce référent comme préexistant au discours, dans un monde réel ou alternatif.

L'exemple (13) contient pour sa part un désignateur vague, le pronom démonstratif *ça*. Le référent en question émane du contexte, à savoir une scène d'ambiance à construire au moyen d'une énumération de SN. Ce « quelque chose » qui explose, à interpréter à partir de l'atmosphère décrite par l'auteur, est volontairement laissé dans le flou, sous-déterminé (cf. Corblin 1991)<sup>23</sup>.

Ces deux exemples mettent ainsi en évidence l'usage de désignateurs, en l'occurrence des démonstratifs dédiés à la référence indexicale, qui s'accommodent mal aux principes dégagés *supra* : d'abord, les objets auxquels il est fait référence au moyen des expressions ne préexistent pas, dans un quelconque monde, au discours, puisque c'est celui-ci même qui les crée (*contra a*). Dès lors, l'évaluation de la conformité de la description employée par rapport

<sup>22</sup> On peut analyser ce procédé référentiel comme une ana-cataphore. Pour l'examen d'ana-cataphores de ce type dans les parenthèses, voir Johnsen (2008) et (2014).

<sup>23</sup> On pourrait arguer que l'emploi se rapproche d'une forme impersonnelle et qu'il est voué à exprimer l'émergence d'une explosion, à la manière d'un présentatif comme *il y a une explosion*. Pour un approfondissement des emplois « impersonnels » de *ça*, voir *infra* (Ch.IV §2.1.2).

à un objet du monde devient caduque (*contra b*)<sup>24</sup>. Ensuite, il faut admettre l'existence de références à caractère sous-déterminé, comme l'illustre (13), auxquelles il est difficile d'assigner une identité, des attributs distinctifs ou des contours bien dessinés (*contra c*).

Ces deux exemples ne sont cependant pas singuliers. Au contraire, les discours attestent régulièrement des occurrences de ce type, si bien que l'on gagnerait en généralité à traiter tous les référents comme des objets de nature cognitive construits par le discours (plutôt qu'à multiplier les univers potentiels pour chaque nouvelle occurrence). C'est dans ce sens qu'un certain nombre de chercheurs ont réorienté la problématique de la référence vers la fin du XX<sup>e</sup> siècle. La fin de ce chapitre (§6) sera consacrée à ces questions. Avant cela, nous nous proposons d'examiner diverses tentatives de classement sémantique des référents.

---

<sup>24</sup> Par rapport à la notion de vérité, on peut ajouter à ces exemples d'autres procédés référentiels qui n'y satisfont pas, sans pour autant que soit mise en échec l'identification référentielle. Lyons (1977 : 182) donne l'exemple d'un locuteur qui prend par erreur le professeur de linguistique pour le facteur et qui y réfère par une expression correspondante, du type « le facteur ». La référence peut tout à fait aboutir, que l'allocutaire partage cette méprise, ou qu'il soit conscient de l'inadvertance de son interlocuteur. De même, l'emploi de descriptions ironiques, dont le locuteur connaît la « fausseté », ne posent généralement pas d'obstacle à la référence (*ibid.*).



### 3. Typologie des référents

Dans la description sémantique traditionnelle, les référents, en tant qu'objets appartenant au monde environnant ou à l'une de ses *versions*, relèvent de divers types que certains auteurs ont tenté de mettre en évidence à partir de critères s'appliquant essentiellement à la réalité tangible. Par ailleurs, une idée communément partagée<sup>25</sup>, que l'on trouve par exemple dans les ouvrages grand public<sup>26</sup>, est qu'aux parties du discours correspondent des types d'objets caractéristiques, par exemple : la désignation d'êtres, d'individus ou de choses pour les noms ; l'expression d'actions, d'états, de manières d'être ou de devenirs pour les verbes. Si ces projections ont été révisées aussi bien au vu de faits typologiques (*e.g.* Croft 1991<sup>27</sup>) que pour une langue particulière (*e.g.* Flaux & Van de Velde 2000) (*e.g.* l'existence de nombreux noms exprimant des quantités, des qualités, des procès, etc.), il n'en reste pas moins qu'elles sont souvent implicitement ou explicitement admises en grammaire ou en sémantique. L'une des typologies de référence en la matière est celle de Lyons (1977 : 438-452), ce dernier précisant toutefois qu'elles ne valent que pour une sous-classe de chaque catégorie syntaxique, autrement dit, pour les représentants considérés comme typiques de la catégorie.

#### 3.1 Les ordres de Lyons (1977)

Afin de décrire la structure de la langue et les relations qu'entretiennent syntaxe, sémantique et référence, Lyons établit une typologie référentielle présentée comme intuitive (fondée sur un « naive realism », p. 442) à partir de ce que l'on peut observer dans le monde : il distingue à cet effet les entités de premier ordre, à savoir, les objets physiques discrets, tels que les personnes, les animaux et autres objets discrets inanimés. Au sein de ce premier ordre, il rappelle la différence entre les personnes, qui occupent une place privilégiée, et les autres types d'objets, différence qui se voit lexicalisée ou grammaticalisée dans la grande majorité des langues (*ibid.*). Néanmoins, les propriétés communes des objets du premier ordre sont jugées constantes du point de vue perceptif : ceux-ci s'inscrivent en effet dans un cadre spatio-temporel donné et sont publiquement observables (*cf.* les *basic particulars* de Strawson 1959). Dans le cadre d'un discours, on peut y référer et leur attribuer des propriétés, actions qu'il est possible de représenter formellement en termes de calcul des prédicats du premier ordre (Lyons 1977 : 443). Lyons oppose à cette catégorie, supposée consensuelle, celles, plus complexes, des entités de deuxième et troisième ordres. Parmi les entités de deuxième ordre, il fait figurer les événements, processus, états-de-choses, etc., eux aussi situés dans l'espace-temps, en ce qu'ils sont dits pouvoir *se produire* (par opposition à la capacité d'exister des éléments du 1<sup>er</sup> ordre). Bien qu'on puisse y référer, dans certaines langues, à la manière des

---

<sup>25</sup> Cf. la citation de Sapir (1921) *supra* §2.3.

<sup>26</sup> Voir le *Petit Robert*, le *Dictionnaire de l'Académie* ou encore le *Larousse* sous les entrées respectives de *nom* ou *verbe*.

<sup>27</sup> Croft (1991) nuance ces correspondances traditionnelles et propose de les envisager comme des « corrélations non marquées » (ou « prototypes »). Pour lui, les fonctions syntaxiques représentent des manières de « conceptualiser » l'information communiquée. Ainsi, face aux limites des rôles sémantiques usuels (agent, instrument, etc.), il envisage le schéma actantiel d'un verbe comme un réseau de relations causales déterminant le choix des types d'arguments du verbe.

individus (par exemple au moyen de nominalisations, telle qu'une *explosion*, une *réunion*, etc.), elles représentent plutôt des constructions perceptibles et conceptuelles : tandis qu'un même objet physique discret ne peut se trouver en plusieurs endroits au même moment – ce qui suppose pour celui-ci un principe de continuité spatio-temporelle – une entité de deuxième ordre telle qu'un événement peut se produire au même moment dans plusieurs endroits différents. La distinction entre une même situation et le même type de situation, sa contrepartie générique, est donc moins claire que pour un objet physique. Ces critères semblent faire consensus, selon Lyons, « within the metaphysical framework of naive realism » (p. 444). Enfin, le troisième ordre comprend les entités abstraites représentées entre autres par les propositions, situées hors de l'espace-temps (*ibid.*). Celles-ci se distinguent à leur tour des entités de deuxième ordre par leur caractère non observable et par le fait qu'on ne peut pas dire d'elles qu'elles *se produisent* ou *se passent*, ni dans l'espace ni dans le temps. Lyons les caractérise ainsi :

Third-order entities are such that 'true', rather than 'real', is more naturally predicated of them ; they can be asserted or denied, remembered or forgotten ; they can be reasons, but not causes ; and so on. In short, they are entities of the kind that may function as the objects of such so-called propositional attitudes as belief, expectation and judgement (p. 445).

Selon les propos de Lyons, on peut référer aux entités de troisième ordre notamment par le biais de nominalisations lexicales (par exemple en français, un *jugement*, un *raisonnement*, etc.), cependant distinctes de celles utilisées pour le deuxième ordre. Au sein de cette typologie, Lyons propose de considérer les entités du premier ordre comme entités par excellence, par opposition à celles des deuxième et troisième ordres, dont la conception se fait à partir des premières et dont les possibilités de référence dépendent de la structure de chaque langue (*ibid.*). Pour en revenir aux relations entretenues entre syntaxe et sémantique, Lyons renomme ce qu'on appelle généralement *noms concrets* par *noms de premier ordre*, qu'il considère être les noms les plus typiques (e.g. *garçon*, *chat*, *table*, etc.). Les noms de deuxième et troisième ordres sont souvent reconnaissables par leur caractère dérivé (*arrivée*, *étonnement*, etc.), bien que certains lexèmes simples puissent également être typiques de ces ordres, comme, respectivement, *état*, *raison* ou *idée* (p. 446).

Lyons (1977) se dit conscient de la nature naïve de sa typologie vouée à une première approximation. Celle-ci n'a d'ailleurs échappé aux critiques (Apothéloz 1995b, Rastier 2004<sup>28</sup>). Depuis, de nombreux chercheurs se sont efforcés d'examiner plus en détail la problématique, par exemple en révisant la typologie des rôles sémantiques traditionnels (Croft 1991), la catégorie nominale (e.g. Flaux *et al.* 1996, Flaux & Van de Velde 2002, Huyghe 2015<sup>29</sup>) les référents de type événementiel (Van de Velde 2006) ou autres objets « abstraits » (Asher 1993, voir *infra* §3.2). Il n'en demeure pas moins que la typologie intuitive de Lyons

---

<sup>28</sup> « Fait d'affirmations non argumentées, l'exposé en reste au bon sens, c'est-à-dire au préjugé hélas commun que toutes les langues représentent les mêmes entités. Pour décrire leur sens, il faut donc décrire ces entités, que l'on discrimine en fonction de propriétés spatio-temporelles. » (p. 20)

<sup>29</sup> Présentation d'un numéro consacré entièrement aux *types nominaux*.

sert de fondement à de nombreux ouvrages en linguistique<sup>30</sup>, ou à des projets d'envergure comme *WordNet*<sup>31</sup> et *EuroWordNet*<sup>32</sup>, bases de données lexicales respectivement pour l'anglais et multilingue (une dizaine de langues européennes) utilisant des ontologies qui projettent sur les parties du discours des types univoques selon les ordres ci-dessus, dans le but de « construire une représentation conceptuelle des référents » (Rastier 2004 : 11).

Mais le problème que pose ce genre d'ontologie est que tout un ensemble de faits linguistiques résiste à ces catégorisations :

- (14) **L'intolérance des tolérants** *existe*, de même que la rage des modérés (Hugo, *Les Travailleurs de la mer*, 1866, p. 118 <TLFi<sup>33</sup>)
- (15) Et rappelle-toi que nous sommes ta famille, à présent, **ta vraie famille** (Queffélec, *Les Noces barbares*, 1988 [1985], p. 201 <TLFi)

L'exemple (14) montre que la propriété d'exister peut être octroyée à des objets autres que des individus discrets et perceptivement stabilisés (1<sup>er</sup> ordre). De la même manière, difficile de considérer que la famille évoquée en (15) constitue un objet abstrait indépendant de l'espace-temps, comme l'induirait l'adjectif *vraie*. On voit par ailleurs que la notion de *vrai/faux* en langue ne s'applique pas exclusivement à des propositions, au sens philosophique du terme ; en d'autres mots, elle ne situe pas le contenu d'un énoncé par rapport à son adéquation au monde, mais elle prédique un degré de conformité par rapport à une norme préconçue (Berrendonner 1985), en l'occurrence, le concept de famille tel qu'il devrait être pour le locuteur. La prise en compte de tels énoncés nous pousse à examiner davantage les indices linguistiques reflétant la manière dont les objets-de-discours sont catégorisés par les locuteurs, plutôt que de tenir pour acquis certaines catégories établies sur la base de l'observation du monde environnant.

### 3.2 Deuxième et troisième ordres : tentative de classement

Selon Lyons, les entités des deuxième et troisième ordres posent davantage de problèmes de traitement que celles du premier ordre. En effet, toute une littérature s'est interrogée sur la question du statut des objets de ce type. Vendler (1967, 1970) distinguait déjà, notamment à partir de la manipulation de prédicats sélectionnels (*containers*) sur des *sentential nominals*<sup>34</sup>, les événements des faits et des propositions : les *tight containers* (e.g. *to occur*, *to take place*, *to begin*, *to last*, etc.) se combinent avec des arguments de type événement, les *loose containers* (*to be unlikely/certain*, *to surprise*, *to cause*, etc.) avec des faits (Vendler 1967), enfin certains *containers* (*to believe*, *to deny*, etc.) sont spécifiques aux propositions (Vendler

<sup>30</sup> Voir par exemple Moeschler & Reboul (1994 : 36) ou encore Dik (1997), qui ajoute un ordre zéro pour les prédicats et un quatrième ordre pour les actes de langage.

<sup>31</sup> <https://wordnet.princeton.edu/>

<sup>32</sup> <https://www.illc.uva.nl/EuroWordNet/>

<sup>33</sup> *Trésor de la langue française informatisé*, <http://atilf.atilf.fr>

<sup>34</sup> C'est-à-dire des expressions « phrastiques » pouvant fonctionner comme arguments de verbes. Ainsi en va-t-il de *that John died*, *John dying* ou *John's death*.

1970). L'idée est que si deux types de *nominals* manifestent une distribution différente à l'égard des *containers*, ils réfèrent à des types sémantiques distincts.

Asher (1993) se propose d'affiner la typologie de Vendler, en la situant notamment sur une échelle d'immanence mondaine (*world immanence spectrum*), fondée sur des critères métaphysiques : les entités nommées *eventualities* (événements, états, activités, accomplissements, achèvements) se présentent à l'extrémité du pôle « concret » en vertu de leurs propriétés spatio-temporelles, de leurs effets causaux et de leur contingence (*i.e.* le fait d'être susceptible de se produire). A l'opposé, les entités abstraites de type *propositions* ne manifestent pas ces caractéristiques : elles existent indépendamment de l'espace-temps dans tous les mondes possibles et ne sont pas dotées d'un potentiel causal. Les entités « factuelles » (faits, possibilités, situations, états de choses) se situent pour leur part entre les deux pôles d'immanence, manifestant des propriétés tantôt propres aux événements, tantôt propres aux propositions (Asher 1993 : 32). Asher se fonde également sur des critères linguistiques pour l'anglais tels que l'aspect (*perfect vs imperfect*) et la formes des expressions (constructions en *that...*, en *-ing*, nominalisations, etc.), mais surtout, à la manière de Vendler (*cf. supra*), sur les contextes compatibles avec les objets abstraits (p. 16-22). Ci-dessous, la distribution du *container believe*, qui sélectionne a priori un objet de type propositionnel, illustre une distinction sémantique entre les arguments du verbe dans chacun des exemples :

(16) Sam believed that Fred hit Mary. (p. 22)

(17) \*Sam believed Fred's hitting Mary (*ibid.*)

Selon Asher, la construction verbale en *that* est compatible avec un verbe d'attitude « propositionnelle », contrairement aux « derived nominals » en *-ing*, « that intuitively denote events » (*ibid.*). L'auteur observe notamment l'usage des anaphores, à ses yeux particulièrement aptes à refléter de telles distinctions sémantiques :

(18) ?[The destruction of the city]<sub>i</sub> took several hours. Fred hadn't believed it<sub>i</sub> when he had first heard about it<sub>i</sub>. (p. 36)

(19) [The destruction of the city]<sub>i</sub> took several hours. Fred hadn't believed that it<sub>i</sub> could happen. (*ibid.*)

L'anomalie ressentie en (18) résulte aux yeux d'Asher d'un « clash » de contraintes, le verbe *believe* requérant un objet propositionnel tandis que la nominalisation en *-ing* renvoie forcément à un événement. Si l'on rétablit un *container* compatible avec un événement comme en (19), l'anomalie disparaît. Néanmoins, l'auteur n'envisage pas des alternatives d'interprétation pourtant ici possibles, comme par exemple, l'éventualité d'une référence au contenu global de l'énoncé précédent. Toute une série de travaux recourt à la typologie d'Asher et l'applique à des données de langues diverses, entre autres Gundel *et al.* (2003)

pour l'anglais<sup>35</sup>, Consten *et al.* (2007 : 93) pour l'allemand<sup>36</sup> et Amsili *et al.* (2005) pour le français<sup>37</sup>.

Amsili *et al.* (2005) mentionnent, parmi les *containers* (ou conteneurs) d'événements en français, des verbes comme *arriver, se produire, avoir lieu, se passer, assister à, être témoin de, manquer, rater*, où l'événement s'exprime au moyen d'un argument (pro)nominal sujet ou objet<sup>38</sup> :

(20) La chute de Marie / ça s'est produit(e) alors que le directeur arrivait. (Amsili *et al.* 2005 : 18)

(21) Tout le labo a assisté à la chute de Marie / à ça. (*ibid.*)

Les propositions, quant à elles, se manifestent essentiellement sous la forme d'une subordonnée complétive d'un verbe d'attitude « propositionnelle » du type *croire, penser, prétendre*, etc. :

(22) Jean croit que Marie est tombée. (*ibid.* : 20)

Quant aux faits<sup>39</sup>, ils se distinguent notamment des propositions par la reconnaissance de leur authenticité. Dès lors, ils sont souvent exprimés via des compléments de verbes comme *savoir que, regretter que, se souvenir que, prouver que, il est vrai que*, etc. ou sujets de verbes du type *surprendre, énerver, décevoir*, etc.<sup>40</sup> :

(23) Léa sait que Marie est tombée. (*ibid.* : 22)

(24) Que Marie soit tombée a surpris tout le monde. (*ibid.*)

Dans certaines situations, les auteurs remarquent qu'un type d'objet donné peut se voir « coercé » (Pustejovsky 1995, Lauwers & Willems 2011). Le phénomène de *coercition* se

---

<sup>35</sup> Gundel *et al.* (2003) soutiennent, sur la base de données attestées et de jugements d'acceptabilité, que les propositions sont des types d'entités cognitivement moins accessibles que les événements et que cette situation se voit dans la distribution des pronoms non accentués (*it*) vs accentués (*this*), les premiers se montrant moins susceptibles de référer à une proposition que les seconds. Cela présuppose, on le voit, que les objets introduits soient clairement précatégorisés.

<sup>36</sup> Les auteurs relèvent des recatégorisations anaphoriques d'un type à un autre, néanmoins uniquement possibles dans une direction de l'échelle d'immanence d'Asher, à savoir du concret vers l'abstrait, ce dont les auteurs rendent compte par le concept d'*ontology changing complexation* : *The Americans tried to invade the building but were forced back by shots from the top floor* (=événements). *This proves* (=fait) *that the situation in Bagdad isn't under control yet.* (< *ibid.*, p. 88)

<sup>37</sup> Il s'agit d'une tentative de représentation formelle du processus anaphorique impliquant les objets abstraits dans le cadre de la DRT (Kamp & Reyle 1993)

<sup>38</sup> Les auteurs notent en outre, à la suite de Davidson (1967), que les énoncés entiers « d'action » expriment des événements : « il faut admettre que toute phrase (d'action, tout du moins) fait intervenir un événement existentiellement quantifié ». L'exemple donné est : *Marie est tombée hier en sortant du labo* (p. 17-18).

<sup>39</sup> Les auteurs ajoutent la catégorie des *situations*, qui sont en fait des événements perçus (p. 18-19).

<sup>40</sup> Selon les auteurs, les assertions isolées expriment des faits : « Toute phrase assertive isolée est pragmatiquement présentée comme vraie » (p. 22). Ils donnent comme exemple : *Marie est tombée*. Il est étonnant de constater que cet énoncé exprime pour eux un fait alors que l'exemple cité dans la note 38 exprime à leurs yeux un événement !

manifeste lorsque les propriétés sémantiques (lexicales, aspectuelles, etc.) d'une unité entrent en contradiction avec les traits inhérents de l'élément effectivement sélectionné. Il en résulte une modification (par *accommodation* ou encore *forçage*) de la nature sémantique sur ce dernier. Dans l'approche en question, c'est le *container* qui impose un changement de type conforme pour l'interprétation de son argument :

(25) Pierre a été surpris par l'arrivée de Jean (Amsili *et al.* 2005. : 26)

Dans cet exemple, le conteneur *surprendre* est supposé contraindre l'expression de l'agent (*l'arrivée de Jean*) à être interprétée comme un fait, « le fait que Jean soit arrivé » (*ibid.*) plutôt que ce qu'il devrait théoriquement représenter à l'origine, à savoir un événement.

A nos yeux, les traits sélectionnels imputés aux *containers* en question semblent trop rigides pour décrire les faits réellement mis en œuvre par les usagers en contexte. En effet, la manipulation des *containers* retenus se montre peu compatible avec l'idée de diversité des emplois, mais aussi avec la polysémie inhérente des unités lexicales : la plupart possèdent différentes acceptions et constructions conventionnelles, d'ailleurs répertoriées dans les usuels lexicographiques « grand public ». Certaines significations sont ainsi tout à fait compatibles avec des expressions référant à des entités réputées de « 1<sup>er</sup> ordre », comme en témoignent ces exemples donnés dans les dictionnaires<sup>41</sup> :

(26) Le lendemain de son mariage, **il** dut descendre dans le salon à l'heure du déjeuner et *se produire* devant une douzaine de personnes (Feuillet, *Mariage monde*, 1875, p.91 < TLFi)

(27) [...] il est prudent comme **un gendarme qui** veut *surprendre* un braconnier (Murger, *Nuits hiver*, 1861, p. 77 < TLFi)

Ces acceptions étant visiblement usuelles pour les usagers de la langue, il n'y a pas de raison de postuler un phénomène de « coercition » à partir d'une signification « première », pour expliquer la polysémie à l'œuvre : le référent de *il* n'est pas reformaté comme un événement, ni celui du SN *un gendarme* comme un fait. Aussi sommes-nous d'avis que les significations attribuées aux *containers* concernés sont moins spécifiées que le revendiquent les adeptes de la méthode.

Par conséquent, la capacité de ces derniers à discriminer des sous-types de référents « abstraits » (des ordres « supérieurs ») nous paraît discutable. Ici également, les *containers* ne nous semblent pas à même de révéler des types univoques :

(28) Notre mère, qui avait **raté sa vocation de surveillante pour centrale de femmes**, se chargea de veiller à sa plus stricte application (Bazin, H., *Vipère*, 1948).

---

<sup>41</sup> Ces emplois ne sont pas propres au style littéraire : **Les artistes qui se produisent** dans la rue devant une foule, ne cesseront jamais de m'étonner. (blog, <http://bloguart.com>) ; Il y a sûrement d'autres moyens, et surtout plus légaux, pour se détendre au volant. Ce samedi matin, **la police a surpris** un automobiliste une main sur le volant et l'autre tenant un joint (Tribune de Genève, 20.07.2004).

Il ne nous paraît pas très intuitif de catégoriser la « vocation » d'une personne comme un « événement ». Même genre de réaction ci-après, où le type « factuel » imposé par le *container* à l'objet du SN *le sommeil* n'est pas très convaincant<sup>42</sup> :

(29) Je fis en me couchant d'autres réflexions qui me parurent conduire à pouvoir expliquer tout ce qui m'était arrivé par des moyens naturels. **Le sommeil** me *surprit* au milieu de ces raisonnements. (Potocki, *Manuscrit trouvé à Saragosse*, 1815 < TLFi)

En fait, nous avons montré dans un travail antérieur à partir de données en français parlé spontané (Johnsen 2010) que les locuteurs, dans leurs actes de référence, ne semblent pas établir de telles distinctions entre les objets. En français parlé, ce sont surtout les pronoms sous-spécifiés et non-catégorisants *ce/ça* qui renvoient aux objets qui ne possèdent pas d'attribut de dénomination propre :

(30) à cinq heures de l'après-midi, faut te mettre dans la tente parce que là, tu as les moustiques qui viennent partout. Donc **ça** c'est l'horreur. (oral, pfc < Johnsen 2010 : 153)

Au vu de (30), on constate tout d'abord que la méthode des *containers* a ses limites car le prédicat effectivement utilisé (*être l'horreur*) ne s'apparente pas à l'un des verbes proposés par ses promoteurs. En effet, comment déterminer le type d'objet auquel est appliqué un prédicat « ordinaire » ? On peut tenter de se reporter à l'expression qui a permis d'introduire l'objet, à savoir la CV *tu as les moustiques qui viennent de partout* : mais comment juger si l'invasion de moustiques invoquée est perçue comme un événement, un fait ou une proposition ? Quant au prédicat *est l'horreur*, il n'apparaît pas non plus vraiment déterminant... Tout bien considéré, il apparaît superflu de procéder à une telle interprétation analytique. A nos yeux, la productivité des formes sous-spécifiées *ce/ça* vient en partie du non-marquage du trait d'individuation, qui suggère au contraire une certaine indétermination catégorielle (*ibid.*). Par ailleurs, d'après les données typologiques de Corbett (1991), aucune langue ne coderait des oppositions sémantiques de ce genre au moyen de pronoms (Fraurud 1992 : 31, Cornish 1999 : 82).

Il faut donc reconnaître que l'objectif des typologies proposées *supra* est de dégager des critères propres au domaine de la métaphysique, mais que ceux-ci, au-delà des emplois considérés comme « typiques », sont difficilement applicables aux phénomènes référentiels à l'œuvre dans les usages véritables des locuteurs. Du reste, on peut relever le paradoxe de la

---

<sup>42</sup> Amsili *et al.* (2005 : 26) relèvent eux-mêmes un cas équivoque, sans toutefois se risquer à se positionner à son égard : *Vous croyez donc à l'accident* ? (Simenon, TLFi). Le *container croire* fait du SN *l'accident*, en principe événementiel, un complément « apparemment propositionnel » : « Il reste que la question de savoir s'il s'agit d'un argument propositionnel reste ouverte » (*ibid.*)

démarche qui consiste à poser des contraintes et établir des catégories, pour en avouer après coup les défauts de prédictibilité<sup>43</sup> :

It [the anaphora test] shows that there is an unexpected fluidity to the typing in the typology of abstract objects and suggests that at a more fundamental level, many distinct types of natural language metaphysics may be very similar, if not identical. (Asher 1993 : 40)

### 3.3 Objets « indiscrets »

A l'opposé de ces modèles logico-sémantiques, un certain nombre de travaux (Berrendonner 1990a, 1994, Apothéloz 1995b, Apothéloz & Béguelin 1995, Corminboeuf 2011, Johnsen 2013 et Berrendonner 2014) mettent en évidence la nature malléable et hybride des référents du discours à travers l'usage des expressions référentielles. Berrendonner (1994) appelle narquoisement *objet indiscret*, par opposition à une vision discrète et rigide du monde, le résultat d'une « indifférenciation » référentielle, autrement dit, la référence à « un tout hybride, polymorphe et logiquement sous-spécifié » (p. 217). On peut relever différentes situations impliquant ce type d'objets. L'exemple ci-dessous illustre le cas d'une expression activant un objet « attracteur » (Berrendonner 1990a), c'est-à-dire susceptible d'évoquer toute une constellation d'objets qui lui sont associés :

(31) Manifestation à Bourg : **le lynx** descend dans la rue (presse, *Le Progrès* < *ibid.* : 154)

Le SN *le lynx* désigne dans cet exemple non pas l'animal mais, par métonymie, les éleveurs furieux des ravages causés par le lynx. Ce dernier représente la cause de la protestation par les manifestants. Mais on aurait tort de réduire la métonymie à un simple déplacement référentiel. L'emploi du SN *le lynx* permet au contraire de condenser et d'évoquer tout un « feuilleton » dont l'animal est le sujet : des discours antérieurs à caractère axiologique, un ensemble d'événements ou de positions le concernant (par ex. sa protection, sa réintroduction, la chasse, etc.) et bien entendu, ses opposants. En tant qu'attracteur, il fonctionne comme « représentant pour tout un micro-secteur de connaissances » (*ibid.* : 155).

Apothéloz (1995b) se penche sur le cas de nominalisations lexicales, réputées pour leur polysémie, susceptibles d'évoquer soit un objet de type procès, soit l'un de ses ingrédients :

(32) Il leur fallait manger maintenant dans des gargotes, pour quelques sous, parmi la crasse et la vulgarité, une nourriture qu'il ne supportait pas. Pour *ces repas*, il allait chercher Tonka, ponctuellement, comme s'il s'agissait d'un devoir. (R. Musil, *Trois femmes* < *ibid.* : 151)

Dans ce cas, le SN *ces repas* peut désigner aussi bien l'action de se nourrir que les mets évoqués. Mais la polysémie du nom ne donne pas forcément lieu à une réelle ambiguïté d'interprétation. Lorsque le contexte n'est pas discriminant, Apothéloz invoque la possibilité

---

<sup>43</sup> Cf. aussi : « Notons que ceux-ci [les noms d'entités abstraites] sont d'ailleurs rarement utilisés avec un sens conforme à leur type ontologique, si ce n'est dans le discours philosophique ou (méta-)linguistique ». (Amsili *et al.* 2005 : 26)

de « discerner soit un procès, soit l'un de ces ingrédients, soit encore une indifférenciation des deux » (p. 169). Cette « sous-spécification » met ainsi à mal la rigidité des échelles logico-sémantiques traditionnelles. L'existence de ce genre d'*amalgames cognitifs* reflète « une certaine latitude d'approximation dans la désignation » (Apothéloz & Béguelin 1995 : 258). Les référents évoqués dans les discours ressemblent de la sorte à des « agrégats polymorphes et traitables par les usagers comme de véritables *objets gigognes* » (*ibid.*).

Parmi la diversité des objets indiscrets, on peut encore relever l'existence de *dualités*, examinées par Berrendonner (1994) et (2014) ainsi que par Corminboeuf (2011). Il s'agit d'objets indiscrets qui se présentent comme « à la fois un et deux » (Berrendonner 2014 : 179).

(33) **La famille royale** passa un merveilleux été dans **leur** château persuadé(e/s) que rien ne pourrait **leur** arriver (oral, radio < *ibid.*)

Dans l'exemple ci-dessus, on identifie un individu collectif ('la famille royale) à la classe de ses membres. Le caractère syntaxiquement lié du déterminant possessif dans *leur château*, malgré la marque de 6<sup>e</sup> personne, conduit à une interprétation coréférentielle plutôt qu'associative. Certains types de dualités semblent privilégiés dans ce genre d'indifférenciation : outre la relation *collectif-classe*, on peut mentionner les relations *lieu-occupants* (34), *objet-dénomination* (35), *classe-type*<sup>44</sup> (36), particulièrement visibles à travers l'emploi de pronoms clitiques, relatifs ou de déterminants possessifs (Berrendonner 2014<sup>45</sup>) :

(34) Il n'y a que **Zürich** à accuser un fort retard, résultat de **leur** manie de se cramponner désespérément à un réseau de tramways d'une lenteur désespérante (presse < *ibid.* : 172)

(35) Le prototype de l'anaphorique, c'est **le pronom\ le pronom/ qui** étymologiquement signifie mis à la place du nom\ (oral, confrencier < *ibid.* : 174)

(36) Le programme **des maîtres** devra être équilibré dans la semaine, ceci pour **lui** permettre des moments de préparation de **ses** cours. (circulaire < *ibid.* : 175)

Les indices en présence suggèrent de confondre les objets plutôt que de les envisager de manière distincte. Enfin, en regard de ces situations d'amalgame, relevons encore la situation inverse, sans doute moins routinisée, de fractionnement d'un même objet (Corminboeuf 2011) :

(37) Dès qu'on essaie de me ranger dans des cases, je suis trop *nombreuses*, on fait des crises de claustrophobie... (Motin, BD < *ibid.* : 476)

---

<sup>44</sup> Sur la notion de *type*, voir *infra* §4.2.

<sup>45</sup> L'opération de dualité via des pronoms sera plus amplement étudiée *infra* (Ch. II §4.2.2).

Le prédicat collectif *nombreux*, ici appliqué à la locutrice, autrement dit un individu, amène à concevoir celle-ci de manière fragmentée, c'est-à-dire comme décomposée en de multiples facettes.

Les faits d'« indiscretion » présentés ici remettent clairement en question les typologies logiques existantes de même que la conception leibnizienne de l'identité qu'elles adoptent pour le traitement des référents du discours (Berrendonner 2014 : 169), à savoir que « deux objets sont identiques (= le même) s'ils possèdent exactement les mêmes propriétés, et sont nécessairement distincts s'ils diffèrent par au moins une propriété ». Les données observées montrent que les locuteurs disposent d'une marge de manœuvre qui leur permet de construire en discours des entités hybrides, fictives, sous-déterminées, etc. que les modèles référentialistes ne permettent pas de prévoir.

### 3.4 *Bilan*

Dans les tentatives de classification logique proposées, on constate que les distinctions établies se fondent sur des critères propres à la réalité environnante ou à une forme de métaphysique : propriétés perceptuelles, ancrage spatio-temporel (stabilité, contingence), relation de causalité, etc. Ces démarches semblent reposer sur un principe qui vient s'ajouter aux trois postulats de la *doxa* sémantique cités *supra* au §2.4 a)-b)-c) :

d) La structure du monde prédétermine la structure de la langue.

L'approche en question s'efforce de transposer les propriétés et le fonctionnement de notre univers sur la langue et à en mesurer la conformité. La démarche que nous proposons dans ce travail est résolument inverse : elle consiste à observer ce que les usages linguistiques nous révèlent sur la manière dont les locuteurs se représentent le monde (*cf. infra* §6.3.1). Ainsi, certaines distinctions sémantiques se reflètent à travers les faits attestés en discours : on remarque en effet que certains objets sont désignés, en contexte, au moyen d'un pronom personnel de 3<sup>e</sup> personne de type *il/elle*, alors que d'autres appellent de préférence un rappel au moyen de *ce/ça* :

(38) en Suisse si on croise euh le regard d'un homme euh qui par hasard euh | nous sourit parce que voilà **il** est sympathique puis qu'on sourit | \_ | euh c'est tout | \_ | on **ça ça** peut se très bien s'arrêter là euh | \_ | on traverse le passage piéton et puis **c'**est bon (ofrom)

Cette opposition montre que les référents respectifs sont envisagés par les locuteurs sous des formats différents. Dans (38), l'emploi du clitique *il* indique que le locuteur réfère à un *individu* présenté comme porteur d'un nom qui le catégorise, en l'occurrence celui d'« homme » antérieurement mentionné, dont le clitique adopte par défaut le genre<sup>46</sup>. En revanche, les démonstratifs *ça* puis *c'* ne présupposent pas la référence à des objets de type individué, dotés

---

<sup>46</sup> L'étiquette lexicale supposée par le genre du clitique n'est pas toujours mentionnée dans le discours, comme on le verra à maintes reprises dans la suite de ce travail.

d'une étiquette nominale usuelle (Corblin 1991, Kleiber 1994a, Carlier 1996, Johnsen 2010). En effet, le contexte conduit à concevoir l'existence de référents aux contours vagues ; l'interlocuteur va interpréter la référence du pronom *ça* comme la situation type décrite, à savoir, le contact établi entre deux personnes à travers leur sourire. Quant au *c'* de la séquence *c'est bon*, qu'on peut paraphraser par *c'est tout* ou *ça va*, il évoque l'épisode contextuel décrit, auquel le locuteur nie toute suite éventuelle au moyen du prédicat.

Cet exemple attesté illustre ainsi la distinction entre individus nommables et objets non nommables (*grosso modo* la différence entre le 1<sup>er</sup> ordre et les ordres supérieurs de Lyons 1977) – « nommable » au sens où le locuteur choisit de présenter le référent comme porteur ou non d'une dénomination. Certains auteurs relèvent en effet l'absence d'une dénomination courante pour certains objets :

Les actions, événements, etc., ne disposent pas de noms en propre. Ils n'appartiennent pas à une classe référentielle dont les individus portent le même nom. Opposés aux entités classifiées, qui sont en quelque sorte des « choses nommées », ils ne sont que des choses. (Kleiber 1994a : 24)

Ainsi, les pronoms *ce* ou *ça* (*cf. infra* 2<sup>e</sup> partie §2.1.2) seraient particulièrement idoine pour ce genre de référence. A travers l'exemple (38), nous soutenons l'idée que les faits linguistiques permettent de mettre en évidence la manière dont les locuteurs catégorisent et conçoivent les objets auxquels ils font référence (*cf. infra* §6.3 l'approche « fribourgeoise »). Par conséquent, en lieu et place des distinctions « métaphysiques » comme les « ordres » de Lyons, à fondement logico-philosophique et conçues sur la base d'une observation « naïve » de la réalité, nous privilégions une approche partant des usages. Apothéloz (1995b : 160) critique à cet égard la rigidité de ces modèles *a priori* qui ramènent les catégories référentielles à celles de la tradition logique et philosophique, dont l'objectif scientifique est distinct, et qui montrent leurs limites face à des faits empiriques mettant en jeu des phénomènes référentiels inclassables en ces termes. Dans le cadre des opérations de référence en discours, la langue est à remettre au centre de l'attention en vue d'étudier le « filtrage cognitif » qu'opèrent ses usagers pour parler du monde (Rousseau 1996). Ainsi, malgré les limites d'une conception linguistique de l'opposition concret-abstrait (*e.g.* Flaux *et al.* 1996, Wilmet 1996<sup>47</sup>), celle-ci sert de fondement à la plupart des typologies sémantiques (entre autres Lyons 1977, Asher 1993, Flaux & Van de Velde 2002, Amsili *et al.* 2005, Consten *et al.* 2007). Or, elle est depuis longtemps remise en cause en linguistique :

Un des critères les plus usuels est le caractère « concret » ou « abstrait » du sens, l'évolution étant supposée se faire du « concret » à l'« abstrait ». Nous n'insisterons pas sur l'ambiguïté de ces termes hérités d'une philosophie désuète<sup>48</sup>. (Benveniste 1966 : 298)

---

<sup>47</sup> Wilmet relève pas moins de sept acceptions différentes, à orientations psychologique (=extrait), ontologique (= immatériel), discursive (=générique), mentaliste (= conceptuel), sémantique (= réduit), morphologique (= dérivé), qui ne vont pas sans poser de problèmes au sein de chaque approche.

<sup>48</sup> A propos du développement de valeurs sémantiques pour un mot.

A nos yeux, les critères retenus à partir de l'observation du monde sont peu appropriés à la description des phénomènes de référence. En effet, on en revient au vieux et vaste débat sur l'arbitraire du signe, qui remet en cause la conception de la langue comme simple nomenclature de la réalité, calquant ses oppositions sur celle-ci : chaque langue possède son propre système de signifiés. Le fait que les langues ne « découpent » pas le monde chacune de la même manière montre ainsi une certaine autonomie de leur part par rapport à celui-ci (cf. les différences bien connues dans la désignation des diverses nuances du spectre des couleurs dans des langues distinctes ou le système des noms de parenté dans différentes langues). En témoigne l'impression fréquente que telle notion dans sa langue maternelle ne se traduit pas aisément dans une autre langue, parce qu'elle contient des traits dont l'autre langue ne rend pas compte<sup>49</sup>, ces contrastes reposant notamment sur des représentations socio-culturelles distinctes.

Or, en établissant une « grammaire du réel » (Apothéloz & Béguelin 1995), c'est-à-dire en cherchant dans le monde les propriétés que doit satisfaire un référent pour porter tel nom, la sémantique admet le principe de nomenclature du réel pourtant depuis longtemps remis en cause. Plutôt que de rechercher les distinctions sémantiques dans le fonctionnement du monde, il nous paraît plus pertinent de dégager celles-ci à partir des manifestations d'une langue. Souvent cantonnée à l'étude du « système », la sémantique, *a fortiori* la sémantique référentielle, aurait à notre sens beaucoup à gagner de l'observation des faits en contexte, pour cerner l'invariant des expressions effectivement utilisées, dans une visée de généralisation. Par définition, l'examen des procédés référentiels dépasse forcément le domaine du « système » pour gagner celui de « l'usage du système ». En effet, dès lors que l'on s'intéresse à l'acte de référence, il s'avère essentiel de tenir compte de la « variété des conditions » (*ibid.*) dans lesquelles les sujets parlants agissent, domaine traditionnellement relégué à l'analyse pragmatique. De ce point de vue, les interlocuteurs ne sont pas réductibles à de « simples instances d'enregistrement et de reproduction » de la relation mot-chose, ils poursuivent des objectifs communicationnels variés qui interviennent dans leur manière de désigner les référents, comme on le verra *infra* (§6.3.6).

Un dernier aspect de ces typologies prête encore à discussion, à savoir la notion polysémique de *proposition*. Nous avons déjà attiré l'attention ailleurs (Johnsen 2010) sur les méprises qui résultent de cette polysémie. En effet, la notion est utilisée aussi bien en logique qu'en linguistique. Et en linguistique, elle s'emploie en syntaxe comme en sémantique, avec des sens différents, respectivement l'association d'un sujet et d'un syntagme verbal, et le contenu de cette même forme syntaxique. A l'origine, la notion est issue de la tradition logique avant de se répandre dans le domaine de la langue :

---

<sup>49</sup> On pense par exemple au nom *tartine* en français, dont l'anglais ne rend compte du concept qu'à travers des périphrases (*bread with jam/butter*, etc.) et l'allemand que par composition (*Butterbrot*, *Marmeladenbrot*, etc.). Le Guern (2003 : 28-29) donne de nombreux exemples dont celui de *libellule*, traduit en anglais par *dragonfly* qui contient un sème fantastique qu'on ne retrouve pas en français ou ceux des équivalents de *gazelle* et *antilope* en punu, une langue du Gabon, comprenant respectivement dans cette langue les sèmes [intelligent] ou [stupide].

Le terme de *proposition* remonte aux grammaires logiques où il désignait toute construction minimale porteuse d'un jugement : l'association d'un sujet (ce dont on dit quelque chose) et d'un prédicat (ce que l'on dit du sujet). [...] Progressivement, la notion de *proposition* s'est confondue avec celle de *phrase* pour désigner l'unité syntaxique et prédicative combinant un sujet grammatical et un groupe verbal. (Riegel *et al.* 2009 : 784-785)

Dans son sens logique, la proposition intervient à des niveaux d'analyse différents. Pour le logicien, la proposition, son objet d'analyse, est ce qui est exprimé par un énoncé déclaratif en vue d'un commentaire sur le monde et qu'on peut juger en termes de valeur de vérité. On peut dire dans ce cas d'une proposition qu'elle est vraie ou fausse conformément au modèle du monde dans lequel elle s'interprète. Le concept de proposition est donc dans ce sens un outil métalinguistique. Ainsi, au niveau de sa dénotation, une proposition logique peut renvoyer, selon les typologies *supra* (§3.2) tantôt à un « événement », tantôt à un « fait », tantôt à une « situation » ou tantôt... à une « proposition » (!) lorsqu'elle complète un verbe d'attitude par exemple. Dans cette dernière acception restreinte, une proposition désigne donc un sous-type d'objet dit « abstrait » (au sens d'Asher 1993) que dénote le plus souvent une *Que-P* qui complète un verbe d'attitude « propositionnelle » (verbes de croyance, de jugement, etc.)<sup>50</sup>, comme dans l'exemple déjà cité :

(39) Jean croit que Marie est tombée = (22)

Selon les typologies *supra*, on devrait considérer que l'énoncé complet ci-dessus consiste en une « proposition » au sens métalinguistique (car susceptible de supporter une valeur de vérité) dénotant un « fait » (comme tout énoncé isolé, *cf.* Amsili *et al.* 2005 : 22), à savoir le fait que Jean croit en une certaine... « proposition » (exprimée par *que Marie est tombée*). On ne saurait dénier ici la confusion engendrée...

Par ailleurs, il nous paraît peu vraisemblable et contre-intuitif de considérer que la séquence *Marie est tombée* puisse représenter, toujours d'après les typologies *supra*, trois types d'objets distincts dans les exemples (39) (en l'occurrence une « proposition »), (40) (un « fait ») ou (41) (un « événement ») :

(40) Léa sait que Marie est tombée = (23) (Amsili *et al.* 2005 : 22)

(41) Marie est tombée hier en sortant du labo (*ibid.* : 18)

On ne voit pas en quoi un procès, ici dénoté par *Marie est tombée*, relève d'une nature différente selon qu'il est l'objet d'une croyance (proposition), selon qu'il est explicitement l'objet d'un savoir (fait) ou indépendamment d'une quelconque modalité (événement<sup>51</sup>). A

---

<sup>50</sup> *Cf.* la critique de Lyons (1977 : 146) : « [...] it cannot simply be assumed that, because a sentence that operates as a clause within another more complex sentence sometimes expresses a proposition (and moreover the same proposition as it would express as an independent sentence), it always expresses a proposition when it operates as a clause in a more complex sentence. »

<sup>51</sup> Ou un fait ? *Cf.* la contradiction d'Amsili *et al.* (2005) à ce sujet que nous avons relevée dans la note 40 *supra*.

nos yeux, il évoque le même objet, lequel peut à son tour entrer dans la construction d'objets plus complexes (la croyance/connaissance du procès en question par quelqu'un, etc.).

En fait, la notion de proposition suscite une grande controverse au sein même de la tradition logico-philosophique (Lyons 1977 : 141)<sup>52</sup>. Etant donné sa circularité ainsi que l'embarras auquel elle donne lieu, nous préférons nous en passer, autant que faire se peut, dans le cadre de la présente étude.

---

<sup>52</sup> « The term 'proposition' is very troublesome. » (*ibid.* : 142)

#### 4. Les expressions référentielles : catégories majeures

L'acte de référence consistant à convoquer un référent dans le discours est rendu possible par l'usage d'expressions référentielles (*supra* §1). Afin d'avoir une vision d'ensemble des ressources linguistiques à la disposition des locuteurs, nous présentons ci-dessous une synthèse des débats en vigueur sur les catégories principalement concernées par les opérations de référence, parmi lesquelles les syntagmes nominaux (SN) et les pronoms personnels sont tenus pour représentatifs. Nous traitons des différentes catégories dans l'ordre suivant : d'abord les noms propres (Npr), qui, dépourvus de descripteur lexical, présentent un fonctionnement référentiel bien distinct des SN descriptifs, avec lesquels ils partagent toutefois la possibilité de se combiner avec toutes sortes de déterminants (§4.1). Nous distinguons ensuite les différentes sous-catégories de SN lexicaux, en fonction du déterminant impliqué car celui-ci renseigne, en tant qu'*actualisateur* (Bally 1932) ou *mot révélateur* du nom (Blanche-Benveniste & Chervel 1966 : 11), sur la manière dont est saisi un objet (son extensité, cf. *ibid.* : 12) et sur le type d'opération cognitive qu'il permet d'accomplir. Nous traitons ainsi d'abord du SN défini (§4.2), puis du SN démonstratif (§4.3). Nous faisons ensuite un détour par les pronoms personnels conjoints et disjoints (§4.4) qui contiennent l'instruction, comme les SN défini et démonstratif, de recruter un référent supposé accessible dans l'espace discursif commun. Ceci n'est pas le cas du SN indéfini (§4.5), que nous considérons en dernier, car il contribue à introduire un objet inédit dans cet espace. A ce titre, il se voit souvent exclu des expressions référentielles, conformément à l'approche philosophique voulant que référence implique identification (Gary-Prieur 2011 : 28), ou alors parce que du point de vue sémantique, il n'est pas considéré comme dépendant d'un contexte quelconque (*e.g.* Corblin 1987a : 14, Riegel *et al.* 2009 : 293). Nous partons au contraire du principe que dès lors qu'on envisage une opération de référence, toute expression référentielle, qu'elle réactive ou introduise un objet, nécessite la prise en compte d'un cadre de référence<sup>53</sup>. Précisons encore que nous laissons de côté un certain nombre d'expressions (les SN possessifs, les pronoms démonstratifs et indéfinis, etc.), car notre objectif est de fournir un bilan critique sur les expressions qui ont fait couler le plus d'encre dans les théories sur la référence, plutôt qu'un inventaire exhaustif des expressions référentielles. Pour terminer, nous discutons la pertinence de modèles à orientation cognitive (§4.6) qui classent les expressions référentielles sur des échelles établies sur les modalités de récupération du référent.

##### 4.1 Les noms propres

Comme la problématique de la référence, le nom propre (désormais Npr) n'a pas fait l'objet d'un intérêt marqué de la part des linguistes avant la deuxième moitié du XX<sup>e</sup> siècle, du fait

---

<sup>53</sup> Même si les modalités de recours au contexte diffèrent selon le type d'expression, la distinction commune (Milner 1982, Moeschler & Zufferey 2010 : 125) entre expressions référentielles autonomes (*le chien du voisin/Un chien/Charlie est dans la cuisine* < *ibid.* : 125) vs privées d'autonomie référentielle (*je/il est linguiste*) nous semble ainsi devoir être reconsidérée.

de sa marginalité par rapport à la conception du signe en vigueur, celle qui requiert la conception du signifié à travers une approche systématique et oppositive de la langue (Gary-Prieur 1994 : 3). Par contre, le Npr représente une notion fondamentale dans la tradition logique parce que c'est à travers lui que s'opère l'accès à l'individu singulier (cf. la notion d'*idée singulière* cf. Port-Royal *supra* §2.2). Dans ce sens logique, le Npr – parfois appelé *nom singulier* ou *individuel* (cf. Mill 1843=1988) – représente toute expression référentielle qui renvoie à un objet singulier (e.g. Frege 1892=1971), cf. pour rappel les exemples *Socrate*, *soleil*, par opposition au N *homme*, qui peut convenir quant à lui à un nombre indéfini de personnes. Dans cette perspective, ce n'est pas tant à l'emploi du Npr qu'on s'intéresse, mais plutôt au rapport du désignateur à l'individu, en tant qu'« opérateur d'individualisation » (Pariente 1973 cité par Gary-Prieur 1994 : 16). On comprend dès lors pourquoi les Npr « véritables » (Frege 1892=1971 : 103) tels que *Socrate*, *Aristote*, etc. ont servi de désignateurs typiques d'individus singuliers. C'est donc tout naturellement que l'analyse des logiciens repose sur l'observation de « propositions » composées d'un Npr de ce type, évidemment sans déterminant et en position référentielle (Gary-Prieur 1994 : 16).

Malgré les contributions en linguistique ces dernières décennies (e.g. Le Bihan 1974, Kleiber 1981, Gary-Prieur 1994, Jonasson 1994, Laurent 2016<sup>54</sup>), la tradition grammaticale s'inspire essentiellement de l'approche logico-philosophique<sup>55</sup>. Deux philosophes en particulier ont influencé la conception du Npr en linguistique, à savoir Mill (1843=1988) et Kripke (1972)<sup>56</sup>. Mill distingue, parmi les noms singuliers, les noms connotatifs (*le premier empereur de Rome*) des noms non connotatifs (*Jean, Londres, l'Angleterre*). Du point de vue grammatical, ces derniers coïncident avec ce que la tradition entend par nom propre. Chez Mill, la connotation représente l'ensemble des attributs impliqués par une expression et est assimilée à celle de signification :

[...] lorsque les noms fournissent quelque information sur les objets, c'est-à-dire, lorsqu'ils ont proprement une signification, cette signification n'est pas dans ce qu'ils dénotent, mais dans ce qu'ils connotent. Les seuls noms qui ne connotent rien sont les noms propres et ceux-ci n'ont, à strictement parler, aucune signification. (1843=1988 : 35)

La conclusion qu'on en tire à propos des Npr, c'est l'absence de signification qui les caractériserait et c'est généralement le seul aspect qu'on retient de la thèse de Mill, comme le déplore Gary-Prieur (1994 : 18-19).

De la même manière, Gary-Prieur regrette l'image simplificatrice qu'on dégage de la thèse de Kripke à propos des Npr et de leur rôle de *désignateur rigide*. Elle souligne que Kripke ne s'intéresse pas au Npr en soi, mais à la relation que celui-ci entretient avec un individu ainsi nommé, comme l'indique la notion de *naming* dans le titre de son ouvrage. D'après Kripke, il

<sup>54</sup> Il s'agit de la présentation d'un numéro entièrement dédié aux noms propres.

<sup>55</sup> Gary-Prieur (1994 : 14) remarque que les définitions de dictionnaire proposent une définition logiciste plutôt que linguistique.

<sup>56</sup> Pour cette synthèse des travaux en logique, nous nous fondons sur l'ouvrage de Gary-Prieur (1994).

existe une relation de causalité entre l'emploi d'un Npr et un acte antérieur, le « baptême initial » (= *naming*), présupposé dans tout emploi ultérieur du Npr (Gary-Prieur 1994 : 19). Outre cette théorie causale, on retient généralement de Kripke la notion de désignateur rigide qu'il attribue aux Npr, dont la propriété serait de « désigner le même objet dans tous les mondes possibles » (*ibid.* : 20), par opposition aux désignateurs accidentels, à savoir les expressions descriptives (*e.g. le président de la République*) dont la référence est susceptible de changer selon le monde considéré. Cette définition, du fait de la confusion autour de la notion de *mondes possibles* propre au domaine de la logique, a parfois été réinterprétée, à tort, comme signifiant qu'un Npr renvoie toujours au même référent (*ibid.* : 22). Or, l'invariance de la référence aux mondes stipulés n'est valable, selon Kripke, que pour un même énoncé susceptible d'être interprété dans un monde hypothétique<sup>57</sup>. Cette idée entraîne parfois un autre malentendu lié précisément au contexte d'énonciation. En effet, cette indifférence aux mondes possibles est parfois traitée comme une indépendance au contexte d'énonciation<sup>58</sup>. À l'inverse, Gary-Prieur (1994 : 25) montre bien que la notion de *naming* (baptême) chez Kripke est étroitement liée à l'énonciation : un Npr n'est pas rigidement lié à un référent déterminé, c'est chaque *occurrence* de Npr qui est rigidement liée à son porteur en vertu d'un acte initial.

En somme, les théories de Mill et Kripke mettent le doigt sur des aspects pertinents du Npr, qui sont toutefois à réévaluer dans une approche linguistique. Gary-Prieur (1994) propose à ce titre un examen particulièrement minutieux dont pourraient s'inspirer les grammaires qui récupèrent les aspects simplifiés de la logique et dans lesquelles le Npr est présenté, sur le plan formel, comme généralement dépourvu de déterminant mais pourvu, à l'écrit, d'une majuscule initiale<sup>59</sup> ; sur le plan sémantique, il y est décrit comme privé de signification et servant à identifier un individu particulier :

Le nom propre n'a pas de signification véritable, de définition ; il se rattache à ce qu'il désigne par un lien qui n'est pas sémantique, mais par une convention qui lui est particulière. [...] Les noms propres s'écrivent par une majuscule (§99, a), sont généralement invariables en nombre (§§523-524), se passent souvent de déterminants (§588). (Grevisse & Goosse 2011 : §461)

---

<sup>57</sup> Les deux exemples suivants montrent ainsi qu'un même Npr peut renvoyer à des individus différents : *Si Aristote n'avait pas été grec, la pensée occidentale aurait été différente ; Si Aristote continue à poursuivre le chat, je l'enferme dans sa niche.* (Gary-Prieur 1994 : 22) Mais Kripke entend par désignateur rigide qu'au sein de chacun de ces énoncés, le Npr renvoie à un seul référent, qu'il s'agisse du monde réel ou d'un monde contrefactuel. Dans chacun de ces exemples, le Npr renvoie ainsi au même porteur de ce Npr quel que soit le monde envisagé.

<sup>58</sup> Cf. l'ambiguïté des propos de Riegel *et al.* (2009 : 337) : « Les noms propres sont cognitivement stables, puisqu'ils désignent leur porteur, indépendamment des variations qu'il peut subir et des situations où il se trouve engagé ».

<sup>59</sup> Cette caractéristique n'est pas propre au Npr : la majuscule possède une valeur très polysémique, servant à la fois de marque segmentale, d'indicateur honorifique, d'indice d'interprétation figurée, de marque iconique d'un phénomène prosodique, etc. En outre, son emploi est très variable d'un locuteur à l'autre et selon les genres, comme on le voit à travers les libertés graphiques dans certains écrits « spontanés » (chat, SMS, courriels, etc.), y compris dans les occurrences de Npr (prénoms, noms de famille, lieux, etc.).

Du point de vue distributionnel, s'il est vrai que les Npr ont la capacité d'apparaître en position de sujet sans déterminant<sup>60</sup>, cela ne représente nullement une contrainte : certains Npr sont systématiquement précédés de déterminants définis, comme les noms de pays, régions ou de cours d'eau (*la Suisse, le Rhône*)<sup>61</sup> (Blanche-Benveniste & Chervel 1966), alors que d'autres s'emploient tantôt avec, tantôt sans déterminant<sup>62</sup>. En fait, à la suite de Kleiber (1981), Gary-Prieur souligne que la distribution des Npr ne se distingue pas de celle des N communs en général, car ceux-là admettent le même éventail de déterminants ou d'expansions du nom<sup>63</sup>, dont voici quelques illustrations :

- (42) *Qu'un Cercaire se tue face à la défaite, c'est compréhensible (La fée Carabine, Pennac < Gary-Prieur 1994 : 30)*
- (43) *Il ne reconnaissait plus son Bernard (Les Faux-Monnayeurs, Gide < ibid. : 33)*
- (44) *Marc disait vrai, elle en était persuadée. C'était bien le Garrett Jones qu'elle connaissait, vu à travers des lunettes très noires (La vérité sur Lorin Jones, Lurie < ibid. : 32)*

Si les Npr ne reposent sur aucun signifié lexical inscrit en langue pour orienter leur interprétation, c'est que celle-ci est essentiellement une affaire de contexte. En effet, l'emploi d'un Npr est fortement lié à des situations d'énonciation particulières : il implique un acte de baptême antérieur<sup>64</sup> ainsi que la reconnaissance tacite de cette convention dénominative par les interlocuteurs (Gary-Prieur 1994 : 28). A cette dépendance contextuelle s'ajoute l'arbitraire de l'acte de baptême (*ibid.* : 27), n'importe quelle entité étant susceptible de se voir attribuer n'importe quel Npr<sup>65</sup> et la dénomination pouvant être remise en cause à loisir de la part des membres de la communauté en question, composée d'au moins deux individus dont le locuteur. Blanche-Benveniste & Chervel (1966 : 27) invoquent d'ailleurs, pour marquer le caractère ancré dans la situation de parole, un « sème » [+locuteur] dans la signification d'un Npr.

---

<sup>60</sup> L'absence de déterminant devant les noms caractérise également d'autres contextes comme des locutions verbales du type *avoir honte/intérêt/peur*, des titres d'ouvrages, de presse, etc. ou encore des tournures syntaxiquement contraintes (*Rendez-vous fut pris*) ou propres à des styles littéraires (proverbes, énumérations), télégraphiques, etc. (*cf.* Blanche-Benveniste & Chervel 1966)

<sup>61</sup> Blanche-Benveniste & Chervel (1966 : 8) attribuent la présence du déterminant dans ce cas au trait [+limitrophe] propre à l'article défini. Les pays, cours d'eau et certaines régions seraient conçus comme des entités finies par opposition aux villes, îles, personnes, etc. qui seraient appréhendées comme des étendues non délimitables et saisies de manière ponctuelle (*cf.* l'usage de *à* pour les lieux).

<sup>62</sup> Par exemple : *Quelle classe, ce Jeremy Irons !* (Gary-Prieur 1994) *vs* *Jeremy Irons confirme ainsi un retour au premier plan pour 2017.* (presse <https://www.cineserie.com/news>).

<sup>63</sup> Pour un inventaire des contextes distributionnels et une analyse sémantique de leurs valeurs, voir Gary-Prieur (1994).

<sup>64</sup> Excepté pour les contextes strictement « dénominatifs » où l'acte de baptême est simultané à l'énonciation, associant tel Npr à tel individu, du type *Qui sait, en France, que le plus grand détective australien s'appelle Napoléon Bonaparte ?* (*ibid.* : 28).

<sup>65</sup> La prédilection « baptismale » envers certains types d'objets est de nature socio-culturelle, les usagers de la langue tendant à baptiser des objets avec lesquels ils entretiennent des relations privilégiées (humains, institutions, monuments, animaux domestiques, lieux, événements marquants ou autres objets familiers, etc.). Voir Charolles (2002 : 60).

Outre cet aspect par définition variable, une caractéristique proprement sémantique qu'on peut dégager des Npr consiste en ce que Kleiber (1981 : 342) appelle le *prédicat de dénomination* qu'il véhicule : l'emploi d'un Npr signifie que le référent désigné s'appelle /Npr/. Cette thèse du prédicat de dénomination est assez convaincante dans les cas où le référent est bien porteur du Npr, mais elle fonctionne cependant moins bien pour les emplois figurés du genre :

(45) *Georges Sand* est sur l'étagère de gauche (Fauconnier repris par Gary-Prieur 1994 : 35)

Dans ce cas, il faut distinguer le référent effectivement visé pour l'énonciation en cours (le livre) du référent initialement baptisé de la sorte (l'auteur nommée *Georges Sand*), qui intervient dans l'interprétation métonymique en question (*ibid.* : 35). La prise en compte du *référent initial* (*ibid.* : 29) est donc nécessaire à l'interprétation du Npr, qui se fonde sur la connaissance de celui-ci. Le référent initial représente ainsi l'individu auquel a été octroyé le Npr lors du baptême initial, identification qui intervient à titre de présupposé du Npr (*ibid.* : 28). Cet aménagement permet à Gary-Prieur d'expliquer aussi bien les cas où le référent désigné est le référent initial, les cas où le référent effectif est une construction discursive, fondée sur une relation métaphorique, métonymique ou encore une simple « image » (*ibid.* : 37) et les cas où le Npr fonctionne de manière prédicative. Dans ces cas, l'interprétation passe par la prise en compte de certaines propriétés du référent initial, sélectionnées sur la base du savoir partagé (comme pour les Npr « lexicalisés », e.g. *Don Juan, Hercule*, etc.) (Jonasson 1991 : 7) ou d'un simple univers de croyance (Gary-Prieur 1994 : 51) :

(46) Tout se passe comme si Saddam Hussein, *ce Faust moderne*, avait choisi la transgression comme mode de comportement. (*Journal de Genève* < *ibid.* : 47)

Dans cet exemple, il n'est même pas nécessaire de recourir aux connaissances encyclopédiques sur le référent initial (le personnage littéraire) pour interpréter le Npr : c'est le contexte lui-même qui fournit les propriétés présentées comme pertinentes par le locuteur, en l'occurrence le comportement transgressif implicitement attribué au référent initial (*Faust*) et explicitement reporté sur l'individu nommé *Saddam Hussein* (*ibid.* : 48).

En définitive, la thèse du référent initial de Gary-Prieur se révèle efficace pour expliquer les trois fonctionnements principaux du Npr (*ibid.* : 58-60), à savoir le fonctionnement *dénommatif* où s'opère véritablement le baptême du référent initial (*Ils ont eu une Cécile*) et dont l'interprétation ne repose que sur le prédicat de dénomination ; le fonctionnement *identifiant* (emploi typique du Npr) où l'acte de dénomination est présupposé et où le Npr désigne le référent initial (*Cécile dort*), à propos duquel on prédique quelque chose ; enfin l'interprétation *prédicative*, qui nécessite la prise en compte de certains attributs pertinents du référent initial pour les projeter sur le référent effectif (*cf. ex. (46)*). Cette analyse, qui tient compte dans le sens du Npr de sa relation avec un référent initial, contrairement à celui du N commun qui mobilise des compétences lexicales, représente à nos yeux une remise en cause

particulièrement judicieuse d'une approche réduisant le Npr à un désignateur rigide et sans signification.

## 4.2 Les SN définis

Par SN définis, nous retenons ici les expressions qui appartiennent au paradigme morphologique *le N*. Bien que les SN démonstratifs soient parfois rangés parmi les expressions définies (e.g. Riegel *et al.* 2009 : 285) en vertu du caractère identifiable de leur référent, nous les traiterons dans une section à part pour faire ressortir les différences de saisie référentielle liées à l'emploi des déterminants.

On reconnaît généralement que *le N* véhicule un présupposé d'existence et d'unicité (Ducrot 1972, Corblin 1987, Riegel *et al.* 2009 : 283) à l'égard de son référent, qu'on peut paraphraser par « il existe un et un seul x qui possède la propriété dénotée par N »<sup>66</sup>. La fonction référentielle du SN consiste à identifier un individu unique correspondant à la description de N :

(47) J'ai mangé *la pomme*. (Gary-Prieur 2011 : 29)

En (47), le SN présuppose qu'il y a un et un seul individu correspondant à la propriété 'pomme' à identifier au sein d'un cadre de référence à présumer, aussi envisagé ailleurs en termes de *domaine d'interprétation restreint* (Corblin 1987), d'*ensemble partagé* (Hawkins 1978) ou de *circonstances d'évaluation* (Kleiber 1990c : 257)<sup>67</sup>. Il est évident que le cadre référentiel pour lequel vaut l'unicité est à réévaluer pour chaque occurrence de SN défini. Nous considérons donc ce domaine d'interprétation comme une notion flexible, car il peut se réduire à un contexte restreint comme à l'univers de discours (cf. les *types* et les référents *quelconques* ci-dessous) selon l'énonciation considérée.

Il arrive assez régulièrement que le présupposé d'existence ne soit pas satisfait du point de vue de l'interlocuteur : on peut dans ce cas invoquer la notion de *coup de force présuppositionnel* (Ducrot 1972 : 51) de la part du locuteur, qui consiste à présupposer l'existence effective d'éléments qui ne figurent en fait pas encore dans le savoir partagé :

(48) *La marquise* sortit à cinq heures (Titre et *incipit* d'un roman de C. Mauriac<sup>68</sup>)

Sur la base d'un consensus feint, le coup de force oblige l'interprète à remédier immédiatement au déficit informationnel par l'importation, après coup, des éléments

---

<sup>66</sup> Cette condition d'unicité est parfois envisagée en termes de totalité (Kleiber 1981, Gary-Prieur 2011 : 28), ou comme la capacité à isoler un et un seul référent des autres par l'intermédiaire de N (Corblin 1987a : 104), afin d'étendre la description au pluriel, aux SN à N [+massif] ou aux emplois génériques.

<sup>67</sup> Kleiber adopte une perspective vériconditionnelle, puisqu'il s'agit des « circonstances dans lesquelles une description trouve sa vérité » (*ibid.*). Pour notre part, notre objectif n'étant pas de vérifier à quelles conditions une proposition est vraie (cf. une approche philosophique), mais de déterminer quel est l'ancrage contextuel d'un énoncé, les notions de domaine d'interprétation ou de cadre référentiel nous paraissent plus judicieuses.

<sup>68</sup> Mauriac fait référence ici à un exemple d'*incipit* imaginé et condamné par Paul Valéry, aux dires d'André Breton dans son *Manifeste du surréalisme*.

manquants impliqués dans les représentations partagées. Ce procédé n'est pas rare en discours et peut intervenir avec d'autres types d'expressions présupposantes à des fins variées comme les pronoms de 3<sup>e</sup> personne (*cf. infra* §4.2), les adjectifs à valeur anaphorique à l'instar de *autre, premier*, etc. (Berrendonner & [Reichler-]Béguelin 1996). On comprend dès lors les rendements des SN définis dans les débuts de roman comme ci-dessus, plaçant le lecteur *in medias res*, c'est-à-dire au milieu d'une scène en cours<sup>69</sup> (Gary-Prieur 2011 : 25).

Le contenu du N se révèle déterminant pour l'identification du référent visé, l'opposant ainsi à d'autres référents distincts par leur signalement. Blanche-Benveniste & Chervel (1966 : 9) invoquent à cet égard la notion de *saisie externe* qu'opèrent les SN définis, par opposition à la *saisie interne* propre aux indéfinis (*cf. infra* §4.5) :

(49) J'ai vu un camion et une voiture. *La voiture* roulait vite. (*ibid.*)

L'emploi du déterminant défini marque le contraste entre deux signifiés (ici le signifié de *camion* en opposition avec celui de *voiture*) : « *le* dit les limites externes de la notion » (*ibid.* : 9). À l'inverse, le défini est non marqué quant à l'aspect continu ou discontinu (*ibid.*), comme le montre sa compatibilité en anaphore aussi bien avec un référent préalablement introduit comme comptable (*cf.* (49) ci-dessus) qu'avec un massif :

(50) J'ai acheté du vin et de la bière. *Le vin* était cher.

La notion de contraste notionnel de Blanche-Benveniste & Chervel est à considérer au niveau sémantique : c'est à travers la nature oppositive du signifié envisagé par rapport aux autres signifiés du système que le référent manifeste son unicité. Cette saisie externe n'implique pas forcément la présence effective ou explicite de référents entrant en contraste avec celui visé. On voit bien en (47) et (48) que les SN renvoient respectivement à l'unique objet correspondant à la description du signifié (par opposition à d'autres) dans le cadre référentiel considéré, sans l'intervention d'objets d'une autre sorte.

On sait que les SN définis sont susceptibles d'être interprétés de manière « générique », d'une autre façon toutefois que les indéfinis (*cf. supra* §4.5) :

(51) *L'homme* a transformé le monde. (< Corblin 1987a : 103)

Dans les grammaires, on assimile souvent cet emploi à celui au pluriel désignant la classe extensionnelle des individus correspondant à N ( $\approx$  *les hommes ont transformé le monde*). Berrendonner (2002a : 41) montre cependant que les deux types de SN ne supportent pas les mêmes prédicats, notamment les prédicats de dénombrement :

(52) Les pandas sont nombreux. (< *ibid.*)

(53) \* Le panda est nombreux. (< *ibid.*)

---

<sup>69</sup> Voir à cet égard l'ouvrage de Gollut & Zufferey (2000) sur les modalités de construction de l'univers discursif dans les incipits de la *Comédie humaine*.

Corblin (1987a : 95) propose d'y voir une différence entre la *classe* (52) (« collection d'analogues discernables ») et l'*espèce* (53), sans toutefois fournir une définition précise de cette dernière, supposée être suffisamment intuitive et présentée comme proche de la notion philosophique de *kind* de Kripke (1972). En outre, pour justifier la condition d'unicité dans ces cas, Corblin (1987a : 104) attribue à *le N* la capacité d'isoler l'espèce N des autres espèces « indépendamment de tout contexte immédiat d'usage ». Cette formulation nous paraît discutable et même entrer en contradiction avec des exemples fournis par l'auteur lui-même :

(54) Au moyen âge, il y avait bien sûr des enfants, mais *l'enfant* n'existait pas. (p. 96)

Il y a bel et bien ici un cadre de référence restreint par le locuteur, explicitement délivré par le complément circonstanciel *au moyen âge*, au sein duquel s'interprète la prédication 'l'enfant n'existait pas'. Nous considérerons donc que les SN définis nécessitent pour leur interprétation, comme toute expression référentielle, un cadre de référence (explicite ou implicite) pouvant dans certains cas s'étendre à l'univers en général. Mais cela ne résout pas la question de la nature sémantique des SN définis « génériques » singuliers. D'autres hypothèses ont été avancées, comme celle de Martin (1986), qui ramène leur contenu à une intension :

*Le* est intensionnel, c'est-à-dire que dans *le chat*, il renvoie à l'intension de *chat*, c'est-à-dire à l'ensemble des propriétés qui font qu'un chat est un chat. *Le chat* réfère génériquement à la « /chatitude/ », à ce que le locuteur considère comme caractéristique du chat. (p. 190)

L'une des limites de cette approche intégrant des éléments abstraits dans le processus interprétatif est qu'elle s'adapte mal aux énoncés impliquant des prédicats événementiels ou d'espèce, comme le signale Kleiber (1990e) :

(55) *Le castor* a été introduit en Alsace par les autonomistes en 1936. (*ibid* : 47)

(56) *Le lynx* est en voie de disparition. (*ibid.*)

Dans les exemples ci-dessus, ce n'est évidemment pas le concept de castor qui a été introduit, ni celui de lynx qui disparaît... L'hypothèse est donc à abandonner. Kleiber propose pour sa part d'apparenter les SN définis génériques singuliers à des SN massifs, même lorsque le N est sémantiquement comptable, le déterminant *le* entraînant un procédé de massification : « La massification opérée consiste en une présentation homogène des occurrences rassemblées par le substantif comptable » (1990e : 158). Pour Berrendonner (2002) néanmoins, cette conception pose elle aussi des difficultés, illustrées par l'échec d'opérations de partition ou de totalisation en principe applicables aux continuums :

(57) \*Les zoos détiennent plus de la moitié du panda (*vs* La vallée de la Jogne produit plus de la moitié du gruyère). (< *ibid.* : 42)

(58) \*Tout le panda vit en Chine (*vs* La grande distribution rafle tout le gruyère). (*ibid.*)

Selon Berrendonner, les SN définis singuliers à valeur dite générique renvoient à ce qu'il appelle des *types*. Les *types* représentent à ses yeux des objets situés dans le versant « intensionnel » de l'univers de discours, contrairement aux objets spécifiques s'inscrivant dans le versant « extensionnel » de nos représentations et en relation étroite avec l'énonciation en cours. Les types ne sont pas des ensembles de propriétés, ou concepts (*cf.* Martin 1986), mais bien des objets comptables, ainsi que l'illustre cette énumération de types, rassemblés sous un SN introducteur de classe :

- (59) Les mots dont on se sert pour exprimer les pensées sont le Substantif, l'Adjectif, l'Article, le Pronom, le Verbe, la Préposition, l'Adverbe, la Conjonction, la Particule ou l'Interjection (Wailly < Berrendonner 2002 : 47)

Les types servent ainsi d'*instruments de catégorisation* pour les objets particuliers de l'univers extensionnel, de *normes idéales*, proches des idées platoniciennes, comme l'incarnent les expressions du genre : *Le N par excellence, le parfait N* (*ibid.* : 49) ou ci-dessous :

- (60) [à propos de sondages dont la diffusion publique est interdite] L'Elysée les aura, les partis politiques les auront, les journalistes les auront, mais l'électeur, lui, *l'électeur majuscule, l'électeur en soi*, ne les aura pas. (radio < *ibid.*)

En somme, la notion de *type* s'intègre totalement à la description générale du SN défini, s'adaptant aux notions de présupposé d'existence et d'unicité, de même qu'à celle de cadre de référence (*i.e.* le versant « intensionnel » de l'univers du discours) et de saisie externe (son signifié l'oppose aux signifiés d'autres *types*).

Il est à noter, pour finir, que les SN définis n'ont pas l'exclusivité de la référence aux *types*, comme en témoigne cet emploi :

- (61) Les vrais pandas sont en voie de disparition. Les paysans grignotent les terres sur lesquelles *cet animal timide* trouve sa nourriture (< *ibid.* : 46)

Le référent est d'abord envisagé, via le SN défini pluriel, comme la classe extensionnelle des individus *pandas*, puis reconfiguré en tant que *type* au moyen d'un SN démonstratif et d'un déterminant possessif. Cet exemple nous offre l'occasion de nous tourner à présent vers le cas des SN démonstratifs.

### 4.3 Les SN démonstratifs

Le SN *ce N* identifie et inscrit « un élément privilégié » dans « la sphère du locuteur » (Blanche-Benveniste & Chervel 1966 : 10). Cette localisation au cœur de l'énonciation rattache le SN démonstratif aux marqueurs dits *token-réflexifs*, c'est-à-dire qui « renvoient effectivement à un objet dont l'identification est liée aux circonstances d'énonciation de leur occurrence » (Kleiber 1984 : 66) ou plus généralement aux marqueurs *déictiques* (De Mulder 2000, Charolles 2002, [Reichler-]Béguelin 1995a, Gary-Prieur 2011) (parfois aussi appelés

*indexicaux* en particulier en philosophie du langage, cf. Levinson 2004 : 97). Ainsi, l'interprétation des SN démonstratifs repose essentiellement sur le « contexte d'énonciation immédiat » (Béguelin 1998 : 96), sans toutefois révéler *ipso facto* la localisation du référent :

Ce sont des signaux qui attirent l'attention de l'interlocuteur sur l'existence d'un référent à identifier dans la situation d'énonciation de l'occurrence sans indiquer par eux-mêmes quel est, pour la situation donnée, ce référent (Kleiber 1984 : 68)

En effet, le renvoi aux circonstances d'énonciation peut impliquer des sources d'interprétation de diverses natures. Un SN démonstratif permet par exemple au locuteur de diriger l'attention de l'allocutaire sur un objet rendu perceptivement saillant par l'occurrence du démonstratif, éventuellement accompagné d'un geste (ou autre indice paraverbal) :

(62) Passe-moi *ce livre*, s'il-te-plaît. (Gary-Prieur 2011 : 71)

Il peut également servir à désigner un objet tout récemment introduit via le contexte linguistique :

(63) Stéphane Mallarmé a renouvelé la poésie du XIX<sup>e</sup> siècle ; *ce poète* a eu de nombreux disciples, dont Paul Valéry. (Riegel *et al.* 2009 : 1038)

(64) Il y a de temps en temps un œuf.

On tire *cet œuf* d'un sac, comme un numéro de loterie, et on le met à la coque, le malheureux ! C'est un véritable crime, un coquicide, car il y a toujours un petit poulet dedans. (Vallès, *L'enfant* < [Reichler-]Béguelin 1995a : 66)

La désignation peut en outre opérer sur un signe en tant que tel plutôt que sur l'objet qu'il désigne (procédé nommé *deixis textuelle* < Lyons 1977) :

(65) Laboratoire ? « *Ce terme* insinue que nous testons un produit sur un cobaye », rétorque Emilie Vincent. (presse, *La Liberté*, 08.11.05)

Parfois, le SN démonstratif renvoie à un objet temporairement sous-déterminé, dont la caractérisation est imminente (cf. la notion de *cataphore infra* §5.2) :

(66) Suis bien *ce conseil* : ne bois que de l'eau (Riegel *et al.* 2009 : 1030)

Bref, le démonstratif est particulièrement adapté aux situations référentielles impliquant une forme de proximité avec l'énonciation. Un tel processus identificatoire distingue le SN démonstratif du SN défini, dont le N intervient de manière décisive pour le repérage référentiel établi sur la base d'un contraste notionnel (cf. saisie externe). Par opposition, on attribue souvent une valeur partitive ou un contraste interne au SN démonstratif (Blanche-Benveniste & Chervel 1966, Kleiber 1984, Corblin 1987, Charolles 2002, Gary-Prieur 2011) :

(67) Tu as oublié *ce carreau* (à une personne qui vient de laver les vitres) (< Charolles 2002 : 111)

Le SN *ce carreau* met en évidence un carreau par opposition à d'autres carreaux. Selon Kleiber (1984 : 76), le SN démonstratif représente l'abréviation d'une structure classificatoire sous-jacente du référent glossée *c'est un N/du N*, autrement dit, il relève d'une quantification partitive. Si l'interprétation contrastive ou partitive est fréquente, on ne peut cependant la ramener à une propriété du déterminant démonstratif (Hawkins 1978, Béguelin 1998). L'exemple ci-dessous résiste en effet à cette analyse :

(68) Gutenberg inventa l'imprimerie et *cette imprimerie* allait bouleverser la culture (Corblin 1987 : 204)

La technique de l'imprimerie ne peut être envisagée comme appartenant à une classe d'imprimeries'. Le démonstratif consiste ici à attirer l'attention sur l'élément tout juste évoqué – sur le mode du connu (l'art en question étant supposé appartenir à notre savoir encyclopédique) –, en le thématissant par progression linéaire, type de progression dont les rendements argumentatifs ont été mis en évidence par Combettes (1983). On doit donc admettre, à travers ce genre d'exemples, que le contraste interne ne représente pas une condition d'emploi du SN démonstratif, qui peut servir à désigner un objet conçu comme unique dans le domaine considéré.

Quoi qu'il en soit, le fait que le N choisi par le locuteur ne constitue pas le point de départ de l'identification référentielle, comme dans le cas du SN défini, offre davantage de liberté au locuteur pour procéder à des (re)categorisations lexicales (Kleiber 1984, Corblin 1987a, [Reichler-]Béguelin 1995a, 1998, Charolles 2002, Gary-Prieur 2011), qui répondent à des stratégies discursives variées (Apothéloz 1995a, [Reichler-]Béguelin 1995a, 1998<sup>70</sup>, Cornish à par.) :

(69) L'affiche de Benetton exhibant un nouveau-né sanguinolent, interdite dans plusieurs pays européens, ne sera placardée ni en ville de Fribourg ni, vraisemblablement, dans le district de la Gruyère. En Suisse, le combat contre *cette* « utilisation abusive du corps humain à des fins commerciales » avait été engagé la semaine dernière par le Vaudois Félix Gluz [...] (*L'Express* < Apothéloz 1995a : 73)

Le procédé de catégorisation est par exemple l'occasion de condenser par nominalisation (ici sur une base implicite) des éléments (objets et relations entre eux) introduits via le contenu d'une (ou plusieurs) constructions verbales antérieures, en même temps que de fournir une perspective subjective sur la base de contenus implicites (*ibid.* : 37). Outre l'acte de référence résomptive proprement dit et la prédication subjective opérée par la recatégorisation lexicale, Apothéloz souligne dans l'exemple ci-dessus le caractère éminemment polyphonique du démonstratif, marqué par l'usage des guillemets, signalant à la fois une « dénomination empruntée » et une distanciation du journaliste vis-à-vis du jugement de valeur manifesté à travers ces propos rapportés.

---

<sup>70</sup> Ces deux articles dressent un inventaire des rendements discursifs des SN démonstratifs, le premier à partir de sources diverses, le second dans les *Fables* de La Fontaine.

Plus généralement, on reconnaît la capacité des démonstratifs à relayer des changements de point de vue, et à promouvoir ainsi une orientation inédite du référent (Marandin 1986, Corblin 1987a, Apothéloz 1995a, Reichler-]Béguelin 1995a, 1998, Cornish à par.) :

- (70) [il est question de grossesses qui se passent mal] La femme se trouve soudain « vidée », face à quelqu'un de très exigeant, totalement égoïste, parfois pas très joli, et elle ne se sent pas prête à affronter *cet étranger*. (*Optima* 1989 < [Reichler-]Béguelin 1995a : 75)

Avec d'autres indices (guillemets, adverbes d'intensité, etc.), le SN démonstratif livre le point de vue d'une mère déconcertée vis-à-vis de son nourrisson : l'effet de proximité se traduit par une « prégnance cognitive » (Béguelin 1998) donnant accès aux représentations de la femme en question et susceptible d'exercer une réaction d'empathie sur le lecteur. Ce procédé<sup>71</sup> a été appelé *deixis am Phantasma* (Bühler 1934=2009), *deixis empathique* (Lyons 1977), *deixis mémorielle* (Fraser & Joly 1980) ou encore *pensée indexicale* (Kleiber 1990d). Le choix du N, à travers des recatégorisations implicites, permet donc de manier des effets de discours subtils. Certains associent à l'usage du démonstratif l'expression d'une rupture avec les circonstances d'énonciation antérieures (De Mulder 1998, Cornish 1999 : 54, Charolles 2002 : 119-121). Au vu d'exemples comme (63), (64) ou (68) qui, à notre sens, marquent plutôt la continuité à travers différents types de progression thématique, nous sommes plus nuancée sur ce point, et considérons les potentiels effets de rupture comme le résultat, en discours, des phénomènes de recatégorisation. Ci-dessous, le scripteur joue ainsi sur la rupture d'isotopie propre à la métaphore en jeu :

- (71) Les larges fenêtres d'une clinique californienne offrent aux malades depuis leurs chambres une vue champêtre sur un magnifique pré planté de splendides arbres. *Ce tapis vert* accueille de gracieux flamants. (*Optima* 1991 < [Reichler-]Béguelin 1995 : 76)

La métaphore se dote ici d'un rendement argumentatif visant à connoter positivement la clinique en question (*ibid.*). On peut ainsi rattacher ces nombreuses exploitations discursives à l'ancrage énonciatif du démonstratif.

#### **4.4 Les pronoms personnels**

Nous nous limitons ici au cas des pronoms *personnels*, les autres types de pronoms (indéfinis, démonstratifs, relatifs, possessifs, interrogatifs) ayant chacun des spécificités qui dépassent le propos de cette synthèse critique. Par ailleurs, certains pronoms feront l'objet d'un examen approfondi dans les chapitres suivants (en particulier *il(s)*, *ça*, *on*), raison pour laquelle nous n'entrons pas dans tous les détails ici. L'enjeu de cette section est simplement de situer le fonctionnement des pronoms personnels par rapport aux autres catégories référentielles majeures.

---

<sup>71</sup> Pour une discussion sur la relation entre démonstratifs et point de vue, voir Kleiber (2003) et Cornish (à par.).

Parmi les pronoms personnels, on distingue les pronoms *conjoint*s (ou clitiques) (*je, tu, il/elle/on, nous, vous ils/elles*) et *disjoint*s (ou toniques/accentués, cf. la série *Moi*<sup>72</sup>, *Toi, Lui/Elle, Nous, Vous, Eux/Elles*). Cette distinction sera plus amplement discutée *infra* (Ch.II §1) mais il nous paraît nécessaire de l'évoquer brièvement, car elle induit certaines différences interprétatives. Les pronoms disjoints partagent la distribution des SN, tandis que les pronoms conjoints sont contraints d'apparaître toujours dans l'environnement immédiat d'un verbe. Les deux catégories se déclinent en personne (1<sup>ère</sup>, 2<sup>e</sup>, 3<sup>e</sup>, 4<sup>e</sup>, etc.), en genre (masc., fém., ou non-marqué) et en nombre (sg., pl.), mais contrairement à celle des pronoms disjoints, la forme des pronoms conjoints a la particularité de varier selon la fonction grammaticale (*je/me, tu/te, il/le/lui, etc.*), reflétant un marquage casuel.

Comme les Npr, les pronoms – conjoints et disjoints – se distinguent par leur défaut de contenu lexical. Leur interprétation est donc sensible au contexte d'une manière différente. En outre, la dépendance au contexte se distingue néanmoins selon le trait de personne. En effet, les 1<sup>ère</sup> et 2<sup>e</sup> personnes renvoient à proprement parler aux « personnes » de l'énonciation (Benveniste 1966) tandis que la 3<sup>e</sup> personne renvoie à un référent tiers, exclu de l'interlocution : « *je* signifie "la personne qui énonce la présente instance de discours contenant *je*" » (*ibid* : 252) et *tu* représente « l'"individu allocuté dans la présente instance de discours contenant l'instance linguistique *tu*" » (*ibid* : 253). On peut assimiler les 1<sup>ère</sup> et 2<sup>e</sup> personnes à des formes *token-réflexives*, mais à la différence des démonstratifs qui ne livrent pas par eux-mêmes les « coordonnées » de leur référent (cf. *supra* §4.3), elles sont supposées indiquer réflexivement l'identité du référent à trouver. A ce titre, on parle de symboles indexicaux *transparent*s ou *complets* (Kleiber 1984 : 67), de *déictiques purs* (Kerbrat-Orecchioni 1980), *directs* (Vuillaume 1993) ou encore *primaires* (Cornish 1999) « parce que leur sens est tel qu'il détermine a priori le type de référent dénoté » (Kleiber 1984 : 67). Béguelin (2002) relativise néanmoins la prétendue transparence des pronoms au regard de phénomènes d'hétérogénéité énonciative subtils, sollicitant des compétences inférentielles de la part de l'interprète :

- (72) -Vous dites que, contrairement aux idées reçues, l'individualisme, le repli sur soi ne sont pas une tendance de fond.  
 - Je pense à une dame qu'on a interviewée : elle avait donné son numéro à SOS amitié et puis avait mis son téléphone sur liste rouge. C'est très significatif. Il y a le besoin de rester chez *moi* dans ma bulle, et celui de partager avec les autres. (presse < Béguelin 2002 )

Dans cet exemple, le pronom *moi* ne renvoie évidemment pas au locuteur de l'énonciation courante désigné par le *je* introductif, mais il s'inscrit dans une « énonciation jouée, fictive, peut-être cursive » (*ibid.*). Des discours directs peuvent ainsi tout à fait surgir dans une énonciation hôte sans démarcation typographique :

---

<sup>72</sup> Nous marquons leur nature disjointe par la majuscule, afin de distinguer certaines formes conjointes homonymes.

(73) Paris, 1960. Tandis qu'Israël se prépare au procès Eichmann, un jeune étudiant en droit rencontre sur l'impériale d'un bus une jeune étudiante allemande. Coup de foudre réciproque, mais il faut que *je te* dise : *tu* es une enfant de l'Allemagne amnésique et *moi je* suis juif, et *mon* père a été déporté. Regarde ces photos de morts-vivants, écoute les récits des camps. Voilà, maintenant *tu* es une Allemande qui sait. À *toi* de choisir. C'est ainsi, ou à peu près, qu'inspirée et soutenue par Serge Klarsfeld, qui l'épousera dans six mois, Beate se lance dans la longue traque que l'on sait. (presse < Béguelin 2002 )

Les indices de 1<sup>ère</sup> et 2<sup>e</sup> personne ne désignent ici ni directement ni simplement le narrateur ou le lecteur mais invitent à envisager une énonciation fictive et parallèle à celle en cours et à en conjecturer l'identité des participants (*ibid.*). Il s'agit donc pour l'interprète de prendre en considération cette dimension parfois implicitement hétérogène de l'énonciation dans les discours afin de mener à bien l'identification, pas si transparente, des référents visés.

A propos des formes dites « plurielles » *nous* et *vous*, on peut invoquer la notion « personne amplifiée » de Benveniste (1966), *nous* ne représentant pas véritablement une multiplication de *je* mais plutôt une amplification pouvant demeurer diffuse, permettant une généralisation de la description des emplois variés de *nous* (modestie, majesté, *je+x*, etc.). Il en va de même pour *vous*, désignant un *tu* « élargi » dont les emplois discursifs peuvent être divers (politesse, *tu+x*, etc.) (*cf. infra* § 2.1).

Comme nous l'avons dit plus haut, les pronoms de 3<sup>ème</sup> personne ne servent pas à référer aux participants de l'énonciation, mais à des individus externes. En ce sens, Benveniste (1966) les appelle *non-personnes*. Les objets auxquels ils font référence sont présentés comme individués et appartenant à une catégorie<sup>73</sup>, ce dont la marque de genre est l'indice (*cf. infra* Ch.II §2.4). Le genre peut être interprété comme la trace d'une dénomination implicite du référent :

(74) On a adopté un bébé. *Elle* s'appelle Lily. (tv, série *Modern family*, saison 1)

Dans l'exemple ci-dessus, le trait de genre (fém.) indique un changement d'attribut de dénomination de l'individu introduit, au détriment d'un maintien de l'étiquette préalablement utilisée (*bébé* n. m.) (Berrendonner & [Reichler-]Béguelin 1995). Au vu de leur économie lexicale, les grammaires présentent les pronoms de 3<sup>e</sup> personne comme un moyen d'éviter la verbalisation d'une information évidente, dont la récupération se base sur le contexte. Par contexte, on peut entendre le contexte linguistique précédent, qui introduit verbalement le référent visé dans le discours, mais aussi toute forme d'input au discours (percepts situationnels, inférences, etc.) :

---

<sup>73</sup> Cette catégorisation n'étant pas obligatoirement explicitement introduite. L'attribut de dénomination peut demeurer implicite (voir *infra* Ch.II §2.4).

(75) *Il* va venir tout de suite. [un censeur d'une école s'adressant à la mère d'un élève qui attend devant le bureau du proviseur] (Kleiber 1990a : 33)

La condition généralement invoquée pour un rappel via un pronom de 3<sup>e</sup> personne est la saillance préalable du référent : le pronom doit assurer la continuité référentielle d'un objet déjà placé au centre de l'attention (*cf. infra* Ch.II §3.3), d'où des sanctions normatives, en contexte scolaire par exemple, à l'égard d'enchaînements de ce type :

(76) Je traversai le mur et me retrouvai à l'intérieur. *Elle* était simple (Elle = [la pièce, la chambre], copie d'élève, [Reichler-]Béguelin 1993a : 328)

On retrouve néanmoins des procédés comparables chez des écrivains notoires, remettant du coup en question ces contraintes d'emploi (voir aussi *infra* Ch.II §4.2.2), comme ci-dessous dans la célèbre scène de *L'Éducation sentimentale* évoquant la première rencontre entre Frédéric et Mme Arnoux (personnage encore non introduit) :

(77) Frédéric, pour rejoindre sa place, poussa la grille des Premières, dérangea deux chasseurs avec leurs chiens.

Ce fut comme une apparition :

*Elle* était assise, au milieu du banc, toute seule ; ou du moins il ne distingua personne, dans l'éblouissement que lui envoyèrent ses yeux (*L'Éducation sentimentale*, Flaubert)

On rétorquera peut-être qu'une apparition implique sémantiquement un actant et que celui-ci figure dès lors déjà dans les représentations des interlocuteurs. Cela dit, ce n'est à notre sens pas (seulement) le critère de saillance cognitive qui motive le choix du pronom (noter également l'emploi de la majuscule<sup>74</sup>) : derrière le choix de taire le nom ou d'éviter un descripteur lexical, on peut observer la volonté de créer de véritables effets de points de vue : le lecteur capte et adopte ainsi le regard subjugué de Frédéric pour cette femme dont il ignore l'identité et dont il est déjà en train de tomber amoureux. On remarquera le va-et-vient entre les focalisations zéro et interne de ce passage, où le point de vue du narrateur réapparaît pour rectifier, peut-être avec ironie, la perception de Frédéric (« toute seule ; ou du moins il ne distingua personne »)<sup>75</sup>.

Une conséquence de la contrainte de continuité référentielle, si l'on y adhère, est que « [l]e pronom [...] n'a pas le pouvoir d'exprimer un changement de catégorie : il maintient le référent tel qu'il était dans la situation saillante de départ » (Charolles 2002 : 223). Là aussi, on trouve des contre-exemples qui illustrent des phénomènes de recatégorisation référentielle implicite à travers l'usage des pronoms :

(78) [à propos d'une roulotte « take-away »] En ce jeudi de juillet, la clientèle défile – *ils* sont parfois une douzaine à faire la queue – et les mets proposés jouent l'éclectisme, entre l'Europe et l'Asie. (presse, *La Liberté*, 24.07.12)

<sup>74</sup> *Cf.* l'édition l'édition Charpentier de 1891.

<sup>75</sup> Voir Béguelin (2013) pour une analyse de l'emploi des pronoms de 3<sup>e</sup> personne chez Flaubert.

Cet exemple illustre le recalibrage, via le clitique *ils*, d'un individu collectif ('la clientèle') en la classe de ses membres. L'exemple (74) *supra* montre également que l'on peut modifier la catégorisation lexicale sous-jacente d'un référent ('un bébé') pour induire une interprétation différente du référent (e.g. son sexe) via l'usage d'un pronom. Ces questions seront longuement discutées *infra* (Ch.II §4.2).

Concernant les pronoms disjoints de 3<sup>e</sup> personne, on trouve des formulations restrictives à l'égard de certaines caractéristiques :

Ces formes se différencient des clitiqes ou « vrais pronoms » *elles/ils/le/la/les* en ceci qu'elles sont réservées aux animés (*Lui, il ne vaut pas cher* ne peut se dire que d'une personne ou d'un animal), qu'elles alternent, pour les inanimés, avec *celui-ci/là* et qu'elles exigent un signe d'ostension (dans les emplois en situation immédiate). (Charolles 2002 : 227)

Les deux extraits suivants mettant en jeu des renvois à des inanimés, puisés dans des genres d'écrit considérés comme « soignés », suffiront à récuser ces propos :

(79) Sans filtre, les mégots de Gitanes s'étaient quant à *eux* depuis longtemps déjà volatilisés, dans cette ultime mince colonne de fumée qu'ils exhalaient en se tortillant lorsqu'ils achèvent de se consumer. (Benoziglio, J.-L., *La voix des mauvais jours et des chagrins rentrés*, 2004 < Frantext)

(80) Donc, ne regrettez pas, mon fils, d'être semblable au vulgaire. Cela que vous dédaignez, en vous, mènera peut-être vos rêveries à la réalisation... et voici : votre goût même pour l'histoire, ce goût qui vous donne la première place entre vos condisciples, en cette matière du moins, c'est *lui* qui vous révèle votre force véritable. (Adam, P., *L'Enfant d'Austerlitz*, 1902 < Frantext)

Il faut cependant reconnaître avec Charolles (*ibid.*) que la forme accentuée des pronoms leur permet d'accomplir un véritable geste d'ostension, consistant à confirmer l'identité d'un référent, comme ci-dessus. En cela, les pronoms disjoints se rapprochent des expressions token-réflexives, par le fait que leur occurrence déclenche un recours immédiat aux circonstances d'énonciation pour l'identification du référent, que celui provienne du contexte linguistique (ci-dessus) ou de l'environnement perceptif (ci-dessous) :

(81) Qui va-t-on désigner ? *Moi/ Lui /Nous* (Riegel et al. 2009 : 371)

La confirmation d'identité associée à leur emploi peut être exploitée à des fins de contraste. En effet, l'emploi du pronom *lui* en (80) permet d'opposer 'le goût pour l'histoire' à tous les autres référents potentiels ('et rien d'autre'). Ci-dessous, le contraste est encore plus explicite :

(82) Nos pensées, disait-il, sont de petits couteaux que nous tâchons d'insérer entre les blocs informes des choses. Nous les savons, *eux*, si fragiles, et *elles*, si pesantes ! (Delattre, L., *Carnets d'un médecin de village* < Frantext)

Cependant, tous les emplois des pronoms disjoints ne mettent pas en jeu un accent d'insistance ou une valeur de confirmation, de contraste. Dans certains contextes, les pronoms disjoints se présentent comme suppléants des pronoms conjoints, là où ceux-ci ne peuvent apparaître en raison de leur statut syntaxique (Nølke 1997 : 60). C'est par exemple le cas dans des positions circonstancielles après une préposition (contrairement aux positions argumentales, où le pronom conjoint est tout désigné) :

- (83) - Paula est obligée de tapiner pour vivre. J'ai peur pour *elle*, de ce qu'on pourrait lui faire : vous comprenez, ou vous aussi vous êtes en pierre ? (Férey, C., *Mapuche*, 2012 < Frantext)

Dans l'exemple ci-dessus, le pronom disjoints *elle* n'est apparemment pas voué à insister d'une manière particulière sur l'identité du référent en question, ni à marquer un contraste spécifique, mais sert simplement à identifier la bénéficiaire du sentiment de peur. D'ailleurs, la construction qui suit rappelle cette bénéficiaire sous une forme conjointe (*lui*), dont la forme dative est disponible pour le verbe en question. On peut relever d'autres constructions syntaxiques favorables aux pronoms disjoints parce que peu accessibles aux pronoms conjoints<sup>76</sup> :

- (84) Le moment venu, je regarderai comme un devoir de ne rien vous cacher. Mais attendons Harrendt ; c'est un homme fort savant, et *lui* et *moi* serons probablement du même avis. (Dumas, A., *En Russie*, 1859 < Frantext)

Dans cet exemple, il n'y a pas de raison d'invoquer un surmarquage de l'identité référentielle ou un effet de contraste<sup>77</sup>.

#### 4.5 Les SN indéfinis

Nous traitons le cas des SN indéfinis en dernier car il se distingue des autres expressions par l'opération radicalement différente en jeu, à savoir l'introduction d'un objet inédit dans le domaine d'interprétation considéré. Cette habileté à introduire des objets nouveaux explique, selon Charolles (2002 : 153), la disposition des SN indéfinis à occuper des positions rhématiques, par exemple dans des constructions existentielles comme ci-dessous :

- (85) *Il y a un alligator* sous mon lit (titre d'un livre pour enfants)

L'interprétation des SN indéfinis n'est pas pour autant affranchie d'un cadre de référence. Dans l'exemple ci-dessus, l'énoncé s'interprète dans un cadre explicitement introduit par le circonstant *sous mon lit*, à situer lui-même à partir d'un contexte d'énonciation plus général impliquant un locuteur fictif, inféré à partir du genre textuel et d'éventuels autres indices à disposition. Charolles (*ibid.* : 154) observe que dans les *incipits* de romans, les SN indéfinis

---

<sup>76</sup> Voir cependant des cas de coordination de pronoms conjoints relevés par Zumwald Küster (2014).

<sup>77</sup> Ou alors, si surmarquage il y a, il résulte plutôt du détail des individus composant l'ensemble, au détriment de l'emploi d'un *nous*.

sont souvent précédés de compléments circonstanciels servant d’ancrage au récit et à ses personnages :

(86) *En 1800, vers la fin du mois d’octobre, un étranger, accompagné d’une femme et d’une petite fille, arriva devant les Tuileries à Paris [...] (La Vendetta, Balzac < ibid.)*

Parfois, l’interprétation d’un SN indéfini dépend même d’un contexte très restreint et peut se rapprocher d’un fonctionnement en anaphore associative (cf. *infra* Ch.II §4.1), où le référent désigné par le SN indéfini est à interpréter comme une partie d’un tout préalablement introduit (Charolles & Choi-Jonin 1995) :

(87) Les policiers inspectèrent la voiture. *Une roue* était pleine de boue. (Kleiber 2001a)

Il s’agit en effet d’interpréter la roue en question comme une roue *de la voiture* ou une *des roues de la voiture*.

Du point de vue sémantique, les linguistes s’accordent pour attribuer au déterminant *un* le trait [+dénombrable] (aussi appelé ailleurs discret, discontinu ou comptable) (e.g. Blanche-Benveniste & Chervel 1966, Gary-Prieur 2011, Riegel *et al.* 2009), par opposition au partitif *du*, qui présente le référent de N comme [+continu] (ou [+massif]) :

(88) Nous avons mangé un poisson. (Gary-Prieur 2011 : 21)

(89) Nous avons mangé du poisson. (*ibid.*)

Certains N, comme *poisson* ci-dessus, sont non-marqués quant au trait de continuité et ouverts aux deux emplois selon l’interprétation désirée par le locuteur, tandis que d’autres N privilégient un certain type de déterminant en fonction de leur sens lexical. C’est le cas, par exemple, du N *lycéen*, qui, doté d’un sème comptable, est généralement voué à désigner un individu humain, s’associant de la sorte à un déterminant indéfini. Néanmoins, il n’est pas rare en discours de « contrevenir » au trait de discontinuité du N pour créer un effet de massification :

(90) L’autobus, de section en section, dégorgeant *du lycéen* et *de la dactylo*, atteignant les confins de Neuilly-Plaisance (Bazin, H., *Madame Ex*)

A l’inverse, il est possible de construire des subdivisions référentielles au sein d’un concept N généralement perçu sans délimitation :

(91) Je me suis caché dans un angle de porte. L’Anglais était couché sur le ventre. *Un sang épais* coulait sur le sol. (Chalandon, S., *Mon traître*, 2007, < Frantext)

Le SN *un sang épais* renvoie à un sous-type de sang, ici explicitement caractérisé par l’épithète *épais*.

Une autre caractéristique du SN indéfini consiste dans le fait qu'il présuppose l'existence d'une classe de N : « derrière chacun des éléments de la série, c'est la série des occurrences qui se profile : *Un chat* s'oppose à tous les autres chats en disant le premier de la série, orientée arbitrairement par le discours à partir de lui » (Blanche-Benveniste & Chervel 1966 : 9-10). Autrement dit, *un N* opère une saisie interne au sein de la classe extensionnelle de N (*ibid.*), que Corblin (1987) reformule comme un procédé d'extraction d'un individu à partir de la classe. Cette opération d'extraction est parfois interprétée à tort comme impliquant la présence effective, dans la situation d'énonciation, de plusieurs objets de la même catégorie (*e.g.* Gary-Prieur 2011 : 29). Mais l'appartenance de l'exemplaire à une classe n'est qu'une information présupposée :

(92) J'ai mangé une pomme. (*ibid.*)

Pour interpréter cet énoncé, il n'est pas nécessaire d'envisager les autres référents du même type dans la situation décrite : l'usage de l'indéfini conduit simplement à assumer l'existence d'une classe notionnelle correspondante, dont le locuteur introduit un spécimen dans le discours.

Notons encore que le SN indéfini s'illustre dans des interprétations « génériques ». Dans ce cas, le cadre de référence s'étend à l'ensemble de l'univers discursif :

(93) *Un chien* reconnaît toujours son maître. (Blanche-Benveniste & Chervel 1966)

(94) *Une société* repose sur des valeurs. (Corblin 1987 : 50)

(95) *Un roman policier*, ça se lit en trois heures. (Gary-Prieur 2001 : 36)

Cette valeur générique se distingue par son interprétation de celle de l'emploi défini renvoyant à un *type* (*cf. supra* §4.2). Ici, l'indéfini met en jeu un élément perçu comme *quelconque* (Gary-Prieur 2011 : 35). Les caractéristiques du déterminant, à savoir ses traits [+comptable], [+nouveau] et son présupposé d'appartenance à une classe restent valables : en somme, dans ce type d'énoncé, le SN signale que le prédicat vaut pour n'importe quel référent prélevé de la classe présupposée<sup>78</sup>.

#### **4.6 Approche cognitive des expressions référentielles**

Dans une approche cognitive, des chercheurs ont proposé des modèles du fonctionnement des expressions référentielles fondés sur l'idée que celles-ci encodent dans leur signifié des

---

<sup>78</sup> Corblin (1987 : 50) décrit cet emploi en recourant à un procédé de multiplication du contexte propositionnel. Autrement dit, l'interprétation implique un prélèvement illimité d'un élément sur la classe, qui aboutit au parcours entier de celle-ci. L'auteur commente ainsi l'exemple (94) : « l'énoncé est vérifié pour un nombre non limité d'extractions d'un individu "société". Grossièrement : "prenez une société, elle repose sur des valeurs, prenez-en une autre (et autant de fois que vous voudrez), elle repose sur des valeurs" » (*ibid.* : 51). Même si c'est à l'idée de référent quelconque qu'aboutit l'interprétation, ce processus multiplicateur ne nous paraît pas indispensable pour la description.

instructions à l'intention de l'interprète en vue de la récupération du référent visé<sup>79</sup>. C'est à la suite de travaux invoquant un degré de *givenness* (littéralement 'caractère donné') ou de *familiarité* des entités discursives (e.g. Chafe 1976, Clark & Haviland 1977, Prince 1981), que les efforts ont porté sur les rapports entre différents types d'expressions référentielles et le statut cognitif des référents dans les représentations mentales des interlocuteurs. A cet égard, deux modèles sont amplement exploités dans les travaux en linguistique, à savoir celui d'Ariel (1988, 1990, 2001) et celui de Gundel *et al.* (1993). Aux yeux d'Ariel, certaines expressions référentielles signalent un degré d'accessibilité référentielle relatif, alors que pour Gundel *et al.*, les formes utilisées manifestent un statut cognitif particulier défini de manière qualitative.

#### 4.6.1 La théorie de l'Accessibilité

Selon Ariel, certaines expressions référentielles encodent un degré d'accessibilité, autrement dit, le fait qu'on a plus ou moins facilement accès à une représentation mentale du référent. C'est donc en fonction de l'accessibilité qu'un référent est supposé avoir pour l'interlocuteur qu'un locuteur sélectionne un type d'expression *ad hoc*. Ariel relève quatre critères susceptibles d'affecter l'accessibilité d'un référent :

- la distance entre un antécédent et un anaphorique
- le nombre de concurrents pour le statut d'antécédent
- le degré de topicalité du référent
- le rôle de l'unité considérée (intégration de l'antécédent dans le même cadre, point de vue, segment, paragraphe, etc. que l'anaphorique)

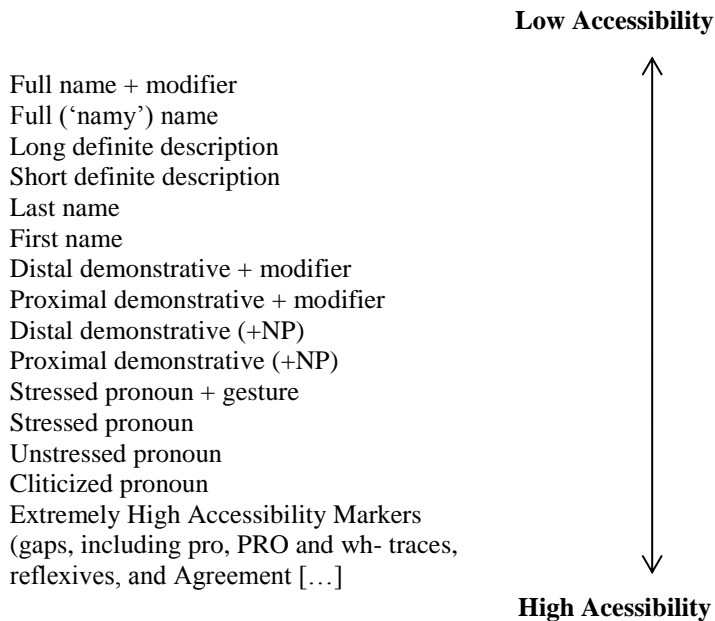
Demol (2010 : 130) relève la subtilité de l'interaction de ces facteurs, qui n'ont pas l'obligation d'être activés simultanément : « le concept d'accessibilité se présente par conséquent comme un concept fort complexe, de sorte qu'il ne suffit pas d'étudier un seul facteur pour en savoir plus sur le fonctionnement de telle ou telle expression référentielle ». Néanmoins, au vu de la difficulté de mesurer ces différents facteurs et leur interaction, Ariel (1988 : 70 ; 1990 : 31) retient essentiellement le critère de la distance pour ses relevés chiffrés – soutenant que les autres se révèlent généralement convergents – en y ajoutant l'un ou l'autre paramètre selon la pertinence du phénomène examiné (la saillance pour les marqueurs de haute accessibilité, l'unité pour les marqueurs d'accessibilité faible). Ses observations se fondent en grande partie sur des données attestées en anglais et en hébreu (1988, 1990), issues de sources diverses (romans, presse, textes scientifiques, interviews, conversation familière). Sur la base de ses résultats, Ariel (1990 : 73) propose une *Échelle d'Accessibilité*, situant à

---

<sup>79</sup> « [R]eferring expressions are no more than guidelines for retrievals » (Ariel 1988 : 68) ; « different determiners and pronominal forms conventionally signal different cognitive statuses (information about location in memory and attention state), thereby enabling the addressee to restrict the set of possible referents » (Gundel *et al.* 1993: 274-275).

l'une des extrémités les noms propres modifiés (e.g. *Joan Smith, the president*) comme les marqueurs d'accessibilité la plus faible, à l'opposé des ellipses, sujets nuls, formes *wh-* ou marques d'accord, qui indiquent l'accessibilité la plus élevée :

**Figure 2 :** Échelle d'*Accessibilité* d'après Ariel (1990)



Ariel considère cette échelle comme universelle : bien que chaque langue ne dispose pas de toutes les ressources linguistiques mentionnées, l'ordre de l'échelle resterait un critère constant à travers les langues. On y retrouve les expressions discutées dans l'inventaire *supra*, à l'exception des SN indéfinis du fait qu'ils ne renvoient pas à un référent déjà accessible dans la mémoire de l'interlocuteur. Les pronoms personnels se voient ainsi affectés d'une accessibilité élevée, autrement dit, ils sont considérés comme spécialisés dans la référence aux entités bien accessibles en mémoire ; les expressions démonstratives, de leur côté, sont dotées d'une accessibilité intermédiaire ; quant aux descriptions définies et aux noms propres, ils sont voués à récupérer des référents considérés comme moins accessibles. Ariel (1988 : 82) se réfère à la théorie de la *Pertinence* (Sperber & Wilson 1986) pour invoquer le fait qu'un marqueur d'accessibilité faible devra comporter, de manière compensatoire, plus d'informations lexicales que les formes situées sur le pôle inverse marquant une haute accessibilité : son coût de traitement élevé se verra donc compensé par des rendements proportionnellement pertinents.

#### 4.6.2 La Hiérarchie du Donné

Le modèle de Gundel *et al.* (1993) (*Givenness Hierarchy*) se distingue de celui d'Ariel par le fait que les différents statuts cognitifs sont définis qualitativement et sont envisagés dans une relation d'implication les uns par rapport aux autres : « each status entails (and is therefore included by) all lower statuses, but not vice versa » (p. 76). Ainsi, le statut *in focus*, pour un

référent, implique que celui-ci satisfait également les autres. Ces statuts sont présentés du plus restrictif (« en focus ») au moins restrictif (« identifiable quant au type d'entité ») :

**Figure 3** : La *Hierarchie du Donné* d'après Gundel *et al.* (1993)<sup>80</sup>

in focus	>	activated	>	familiar	>	uniquely identifiable	>	referential	>	type identifiable
{it}		{this that this N}		{that N}		{the N}		{indefinite <i>this</i> N}		{a N}
{il, ça}		{cela, ceci, ce N-ci, celui-là, celui-ci}		{ce N-là}		{le/la N}		{ce N(-ci) indéfini}		{un/une N}

A noter que contrairement à l'échelle d'Ariel, celle-ci n'intègre pas les Npr. Par contre, elle tient compte des SN indéfinis. Selon les auteurs, un référent *en focus* est situé dans le centre d'attention des interlocuteurs (*current center of attention*, p. 279), au sein de la mémoire à court terme. C'est le statut requis pour l'emploi des pronoms non accentués (« unstressed pronominals ») ou des anaphores zéro en anglais. Le statut *activé* indique que le référent à rechercher se situe dans la mémoire à court terme, mais pas forcément dans le centre d'attention des interlocuteurs (*I couldn't sleep last night. That kept me awake*<sup>81</sup>). Un référent *familier* signifie que l'allocutaire possède une représentation particulière du référent au moins dans la mémoire à long terme (*I couldn't sleep last night. That dog (next door) kept me awake*). Dans le cas du statut *uniquement identifiable*, l'allocutaire doit pouvoir accéder à une représentation unique du référent par le biais du N lexical, que celle-ci existe préalablement en mémoire ou non (*I couldn't sleep last night. The dog (next door) kept me awake*). Le statut *référentiel* implique pour l'allocutaire l'identification d'un référent correspondant à N, et signale que le locuteur s'attend à ce que l'allocutaire construise une représentation particulière du référent visé (*I couldn't sleep last night. This dog (next door) kept me awake*) Enfin, l'allocutaire accède à la représentation d'un référent *identifiable quant au type d'entité* par le biais de sa connaissance des éléments lexicaux, sans qu'une représentation particulière soit préalablement connue (*I couldn't sleep last night. A dog (next door) kept me awake*). La relation d'implication unidirectionnelle signifie qu'une entité en focus est également activée, familière, uniquement identifiable, référentielle, de même qu'identifiable quant au type. Dès lors, une expression associée à l'un de ces statuts impliqués peut en théorie être utilisée pour référer à un objet en focus : par exemple, un SN défini pourrait être employé dans cette situation. D'après les auteurs, les diverses possibilités se voient néanmoins pragmatiquement

<sup>80</sup> Cette figure est adaptée en français dans Gundel *et al.* (2000)

<sup>81</sup> Les exemples voués à illustrer la distribution des formes en fonction des différents statuts sont pour la plupart forgés. Néanmoins, aux dires des auteurs, la *hiérarchie* a été testée et comparée entre plusieurs langues (l'anglais, le mandarin, le chinois, le japonais, l'espagnol et le russe) sur un corpus diversifié (romans, nouvelles, presse écrite et orale, conversations), avec des similitudes mais aussi certaines divergences, par exemple dans l'emploi respectif des SN démonstratifs ou indéfinis. Il faut cependant reconnaître avec Demol (2010 : 136-137) que la méthodologie suivie n'est pas rigoureusement décrite : les critères relevés de manière non systématique reposent sur la syntaxe, la distance entre les mentions, des jugements de pertinence ou la prise en compte du savoir partagé (Gundel *et al.* 1993 : 291).

restreintes par la maxime de quantité de Grice (1975), requérant d'une part d'être aussi informatif que nécessaire, d'autre part d'éviter de donner plus d'informations qu'il ne faut.

### 4.6.3 Bilan

Au-delà des divergences terminologiques, méthodologiques ou théoriques de ces modèles, on peut néanmoins relever leur vocation unanime à doter les expressions référentielles d'instructions univoques pour la récupération référentielle. En cela, ces approches se fondent essentiellement sur l'idée d'une adéquation des formes aux intérêts de l'allocutaire, en reléguant au second plan d'autres aspects qui pourraient guider le choix du locuteur dans la désignation opérée. Par exemple, au vu de données comme (63), (64), (68), (69), (70), (71), où le référent rappelé a été récemment introduit, il nous semble contre-intuitif de justifier l'emploi du démonstratif par le caractère moyennement accessible ou préalablement en retrait du référent. De la même manière, ni l'accessibilité ni le statut cognitif ne parviennent à expliquer l'emploi d'un pronom personnel comme (77), même en faisant intervenir des lois pragmatiques (*e.g.* maxime de quantité). Les usages qui s'en écartent ne sont pas totalement négligés<sup>82</sup> par les auteurs, mais ils sont traités comme des exceptions – ou « violations », selon les termes des auteurs (Ariel 1990 : 200, Gundel *et al.* 2000) – faisant intervenir des phénomènes d'implications contextuelles ou d'accommodation (*cf. infra* Ch.II §4.2.1).

D'après Apothéloz (1995a : 73-74), l'emploi d'une expression référentielle implique deux processus : l'acte de référence proprement dit et l'acte de dénomination. Dans les théories dominantes, on part du principe que le dernier opère au service du premier, autrement dit, qu'on dénomme pour identifier. Mais c'est oublier que la dénomination peut opérer en vue d'autres objectifs (par exemple un apport significatif d'information via un SN lexical à des fins argumentatives, ou à l'inverse une sous-spécification à des fins de cryptage, de polyphonie, etc.). Or, l'appariement d'une forme linguistique à une modalité cognitive, comme le postulent les échelles, contribue à notre sens à banaliser le procédé référentiel, en marginalisant les facteurs contextuels en jeu (rapport des interlocuteurs, intérêts respectifs, genre de discours, etc.). Aussi, sans nier totalement la pertinence de ces échelles, nous préférons considérer que les critères cognitifs invoqués constituent un paramètre au même titre que d'autres dans le recrutement d'une expression référentielle.

Pour décrire plus scrupuleusement ces opérations, il est nécessaire de faire à présent le point sur les notions largement exploitées d'*anaphore* et de *deixis*, qui concernent la manière dont le référent visé est porté à la connaissance des participants de l'interaction, mais qui comportent, selon les conceptions, un certain nombre d'inconvénients théoriques.

---

<sup>82</sup> Voir par exemple le chapitre 8 d'Ariel (1990 : 198 *sqq.*) intitulé *Special uses of accessibility markers* ou Gundel *et al.* (2000 : 92-94).



## 5. Modes de désignation référentielle : *anaphore* et *deixis*

Parmi les opérations qui servent à récupérer un objet à partir de l'espace commun aux interlocuteurs, il est d'usage en linguistique textuelle d'opposer deux modes de référence, selon la manière dont on accède au référent. La conception traditionnelle peut être qualifiée de textualiste, étant donné qu'elle fonde la distinction sur l'accès au référent via le texte *vs* la situation extralinguistique. Une conception concurrente, qu'on peut appeler cognitive, invoque de son côté un critère de nouveauté ou de maintien du référent dans le savoir partagé, en lieu et place du critère textuel/situationnel. On peut illustrer ces deux conceptions à travers les exemples suivants repris de Kleiber (1992 : 614) :

(96) Paul a enlevé son chapeau. *Il* avait trop chaud. (p. 614)

(97) *Ce* chien cherche son maître (avec geste d'ostension). (*ibid.*)

Selon la conception textualiste, on explique généralement l'interprétation du pronom de 3<sup>e</sup> personne en (96) par le renvoi à un antécédent dans l'espace textuel en amont, en l'occurrence le nom propre *Paul* ; en (97) au contraire, le SN démonstratif s'interprète à partir de l'environnement extralinguistique concomitant à l'énonciation. D'après l'approche dite cognitive au contraire, l'emploi de *il* en (96) se justifie par le maintien du référent déjà établi dans le savoir partagé, tandis qu'en (97), l'emploi du démonstratif est favorisé par la volonté d'installer un référent au sein de l'attention des interlocuteurs.

Avant de présenter plus en détail ces deux approches ainsi que les problèmes qu'elles soulèvent, nous nous proposons de remonter brièvement aux sources des notions d'anaphore et de deixis.

### 5.1 Origine des notions d'anaphore et de deixis

D'après Bosch (1983), les premiers développements que l'on trouve sur la notion d'anaphore proviennent d'Apollonius Dyscole, grammairien grec d'Alexandrie du 2<sup>e</sup> siècle après J.-C. Dans son traité sur les pronoms (*Peri Antonymias*), il mentionne le fonctionnement soit déictique soit anaphorique des pronoms (personnels et possessifs) et rattache la deixis à la *prôté gnôsis* (connaissance première) et l'anaphore à la *deutéra gnôsis* (connaissance seconde), la deixis constituant la référence à des objets nouveaux dans le discours, l'anaphore celle à des objets déjà présents dans le discours (Bosch 1983 : 5-6). Cette distinction d'Apollonius aurait été redécouverte après des siècles par Windisch (1869) dans le cadre d'une étude sur les pronoms relatifs et c'est cette conception, devenue courante parmi les spécialistes des langues indo-européennes contemporaines, qui prédominait encore au début du 20<sup>e</sup> siècle<sup>83</sup> (*ibid* : 8). Selon Bosch (*ibid* : 5-6), c'est une formulation « malheureuse » de Windisch qui est responsable de l'interprétation actuelle assimilant la deixis à la référence situationnelle et l'anaphore à la référence intra-textuelle. L'approche d'Apollonius et de

---

<sup>83</sup> Voir par exemple Brugmann (1904) sur les pronoms démonstratifs.

Windisch serait en fait proche de celle en termes de *donné* et *nouveau* dans les théories cognitives plus récentes (e.g. Chafe 1976, Clark & Haviland 1977, Prince 1981, cf. *supra* §4.6). Cela dit, les notions d'anaphore et de deixis n'étaient pas particulièrement centrales, ni chez les grammairiens de l'Antiquité, ni chez les indo-européanistes, et leur servaient plutôt d'instruments auxiliaires de description voués à des objectifs de recherche particuliers (Bosch 1983 : 9).

## 5.2 Conception traditionnelle de l'opposition anaphore vs deixis

Bühler (1934=2009) se profile comme un acteur important de la conception théorique du couple *deixis-anaphore*. Parmi les procédés de pointage (*Zeigen*), il distingue le pointage d'objets (*sachliches Zeigen*) ou *deixis* (du grec 'montrer'), du pointage syntaxique (*syntaktisches Zeigen*) ou *anaphore*. Le premier procédé sert à référer directement aux objets du monde réel (ou imaginaire = *deixis am Phantasma*) ; le second est utilisé pour pointer sur des éléments du contexte linguistique, c'est-à-dire les expressions mêmes du discours, ou les choses telles qu'elles ont été caractérisées par les interlocuteurs dans le discours (p. 563). Il est à noter que pour Bühler, l'anaphore est une « deixis réflexive » (p. 561) où le contexte linguistique devient le champ déictique (*Zeigfeld*) et où les pointeurs « prennent en charge une fonction déictique qui assure un service interne » (p. 563, nous soulignons) en pointant en avant ou en arrière sur la chaîne du discours. Il y a là une distinction nette de traitement entre le processus de renvoi à un objet extérieur au discours ou à un objet mentionné. L'opposition entre référence extra-linguistique et intra-linguistique est ainsi bien établie : on fait désormais comme si la désignation d'un objet de l'environnement extérieur était d'un autre ordre que la désignation d'un objet du discours.

Cependant, la relation inclusive proposée par Bühler, faisant de l'anaphore un sous-type de deixis, est rarement mise au premier plan et les schémas classiques tendent à induire une interprétation équipollente des concepts. Le lien de parenté est surtout mis en évidence dans le processus d'émergence de l'anaphore : Lyons (1977 : 667-671) montre que les emplois anaphoriques des démonstratifs proximaux et distaux dépendent de leur composante premièrement déictique. De même, Halliday & Hasan (1976 : 32) et Fraser & Joly (1980 : 25) reconnaissent la primauté de la référence extra-textuelle (appelée *exophore*) sur la référence intra-textuelle (appelée *endophore*) : « exophore et endophore sont organiquement liées, celle-ci étant issue de celle-là » (*ibid.* : 25). D'ailleurs, pour ces derniers, le phénomène de deixis constitue un système de repérage spatio-temporel qui recouvre aussi bien l'exophore que l'endophore.

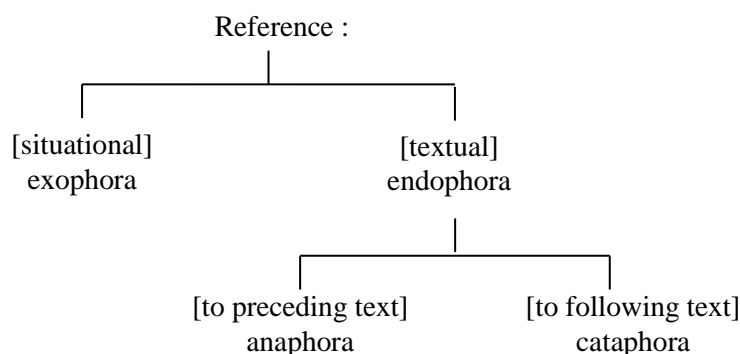
Il n'en reste pas moins que d'après leurs schémas comme d'après certains de leurs propos, les deux modes de référence (selon la localisation du référent) sont constamment mis dos à dos, plutôt que présentés dans une relation inclusive :

[I]nguistiquement, un objet peut avoir deux lieux d'existence : hors du discours ou en discours. On est du même coup amené à distinguer deux types

de référence : celle qui renvoie à un objet dont le lieu d'existence est extra-discursif, et que nous nommons *exophore*, et celle qui renvoie à un objet dont le lieu d'existence est intra-discursif et à laquelle nous donnons le nom d'*endophore*. (*ibid.* : 24)

L'un des schémas représentatifs à cet égard est celui d'Halliday & Hasan (1976 : 33), issu de leur ouvrage consacré à la problématique de la cohésion textuelle :

**Figure 4** : Schéma des modalités référentielles d'après Halliday & Hasan (1976)



A noter que les auteurs emploient ici le nom d'*anaphore* au sens strict pour désigner un sous-type d'endophore, celui où « l'antécédent », autrement dit l'expression à laquelle il convient de se reporter pour l'interprétation, est situé en amont dans le texte. Le terme d'*anaphore* s'oppose à celui de *cataphore* où l'expression référentielle dépend d'un segment situé en aval, conformément à l'étymologie du mot<sup>84</sup>.

### 5.3 L'anaphore sous l'angle textualiste

Cette conception de l'anaphore perçue comme un phénomène de dépendance interprétative asymétrique entre deux expressions linguistiques a été largement diffusée en linguistique et en didactique des langues, en particulier à la suite de l'ouvrage d'Halliday & Hasan. Aux yeux de ces derniers (*ibid.* : 31), le procédé cohésif par excellence est « l'endophore », marquant la continuité référentielle au fil du texte construit. Cette approche, reposant cruciallement sur une opération de repérage segmental dans la chaîne textuelle, ne va pas sans soulever bon nombre d'objections, comme on le verra *infra* (§5.4). Avant de les formuler, il est cependant nécessaire d'examiner les tenants et aboutissants de l'objet en question.

Pour décrire le lien qui unit un élément anaphorique à une expression à récupérer en amont dans l'environnement textuel, il est d'usage de dire qu'il *réfère* à son *antécédent* (Lyons 1977 : 660). Maillard (1974) recourt d'ailleurs au terme de « référé » pour désigner l'*antécédent*<sup>85</sup>. Cet emploi du verbe *référer* correspond en fait à celui du verbe latin '*referre*' dont il provient, lui-même traduit du grec '*anapherein*', composé de '*ana-*' (en haut, en

<sup>84</sup> Voir aussi chez Maillard, où la notion de *diaphore* – la référence au contexte linguistique – recouvre les cas de l'anaphore et de la cataphore selon la localisation respective de ce qu'il appelle le « référé » (Maillard 1974).

<sup>85</sup> Appelé ailleurs *source sémantique* (Tesnière 1959 : 85-87). Apothéloz (1995a : 310) relève également les noms d'*interprétant* et de *contrôleur*.

arrière ou de nouveau) et '*pherein*' (porter), littéralement 'reporter plus haut' ou 'répéter'<sup>86</sup>. Dans la perspective adoptée, l'anaphorique renvoie donc à un antécédent présent dans le contexte linguistique en amont (ou en aval pour son fonctionnement *cataphorique*). Ce sens du verbe *référer* est à distinguer de l'acception « philosophique » du verbe, renvoyant à l'acte de désignation d'un objet du monde (Lyons 1977 : 660). D'après cette seconde acception, le pronom ne réfère pas à son antécédent, mais au référent de cet antécédent. Cornish (2010 : 212-213) critique à cet égard l'emploi du verbe *référer* pour désigner la relation qui lie directement un anaphorique à son antécédent. Une variante du verbe *référer* au sens étymologique est le verbe *reprendre* ou sa nominalisation. On recourt ainsi couramment aux termes *reprise anaphorique*, dans les études linguistiques comme dans les grammaires, (voire de *substitution* ou *suppléance* dans le cas du pronom, cf. *infra* §3.1), etc. En tout cas, cette terminologie reflète bien le processus de repérage textuel invoqué :

un segment de discours est anaphorique lorsqu'il est nécessaire, pour lui donner une interprétation (même simplement littérale) de se reporter à un autre segment du même discours (Ducrot & Todorov 1979 : 358).

Les expressions anaphoriques se révèlent de la sorte sémantiquement dépendantes<sup>87</sup>, autrement dit, elles se présentent comme référentiellement incomplètes :

Les expressions anaphoriques se reconnaîtraient donc au fait qu'elles sont non saturées référentiellement ; isolées de l'environnement où elles recrutent une expression permettant de fixer leur référence, elles seraient ressenties comme incomplètes car impropres à constituer à elles seules un objet de discours autonome (Charolles 1991: 209).

De même, pour Tesnière (1959 : 86), un anaphorique est un mot préalablement vide, qui devient sémantiquement plein dès qu'il entre en connexion anaphorique avec un autre mot. Il use à cet égard d'une image éclairante :

on peut comparer avantageusement les anaphoriques à des ampoules électriques qui ne s'allument que lorsque le fil qui les alimente est mis en contact avec la source d'électricité. Il y a là en quelque sorte une prise de courant sémantique. (*ibid.* : 90)

On peut rapprocher de cette comparaison la notion de « chaînes de référence » (Chastain 1975). Corblin (1987c : 7) exploite cette notion pour « dépasser les contextes de simple succession de deux termes auxquels se limite le plus souvent le linguiste qui sort du domaine phrastique ». De même, Schnedecker (1997: 8) recourt à la notion pour désigner une succession d'expressions coréférentielles initiées par une forme lexicale afin d'étudier comment les différents maillons d'une chaîne contribuent au « continuum textuel ». L'objectif est d'examiner les différents types de chaînes et leurs composantes lexicales, leur rôle dans l'organisation textuelle, leur imbrication, leur sensibilité au genre discursif, ou encore à l'époque (Schnedecker & Landragin 2014).

---

<sup>86</sup> D'où le développement, en rhétorique, de son sens de répétition stylistique en début d'énoncé.

<sup>87</sup> Dans le même esprit, Kleiber (1994b : 47) évoque dans ce sens la notion de « désignateurs seconds » pour les pronoms anaphoriques.

En somme, la conception de l'anaphore comme la manifestation d'une dépendance interprétative entre un élément incomplet (l'anaphorique) envers une expression « donatrice » (l'antécédent) est celle en vigueur dans les usuels de grammaire ou de linguistique, mais aussi dans la plupart des modèles formels, computationnels et psycholinguistiques où le repérage segmental constitue un enjeu primordial<sup>88</sup>.

#### 5.4 *Limites d'une approche textualiste de l'anaphore*

Si l'approche de l'anaphore dans une perspective textuelle a l'avantage de considérer les rapports référentiels au-delà de l'enchaînement de deux énoncés concomitants, ouvrant ainsi l'analyse à l'empan d'un texte authentique et à sa cohésion, le processus interprétatif sous-jacent reste critiquable<sup>89</sup>. En effet, il implique une remontée, pour le destinataire, de la « chaîne de référence » jusqu'à la découverte de l'occurrence initiale (e.g. Brown & Yule 1983 : chap. 6). D'ailleurs, dans le cadre de l'oral, la trace auditive et sa disparition sont instantanées (Cornish 2010b) : il paraît dès lors invraisemblable d'invoquer une mémorisation des énoncés ou une identification des segments en cause (Valentin 1985). L'exemple suivant illustre la difficulté de remonter, même à l'écrit, à une expression initiale bien définie :

- (98) On lui avait offert un harmonica et il a commencé à jouer et à retranscrire tous les thèmes qu'il entendait à la radio. Il est sorti de l'hôpital passionné de jazz. Deux ou trois copains lui ont expliqué ce qu'étaient les grilles d'harmonie et il a appris le solfège comme ça. Alors, il a eu une idée : faire un grand orchestre en France, comme Woody Herman. Ce qui est extraordinaire, c'est qu'il a presque réussi. Il a embauché, au théâtre Mouffetard, vingt-cinq musiciens, tous excellents, et il a commencé à les faire répéter pour enregistrer un disque. Les répétitions étaient publiques et nous y sommes allés. Il y avait des musiciens qui sonnaient fantastiquement. Il y avait un arrangement qui s'appelait *Pithecanthropus Erectus*, mais je me trompe peut-être parce que je crois que Mingus a fait la même chose. Plus ça continuait, moins de musiciens venaient. En effet, c'étaient tous des premiers prix de conservatoire qui avaient été engagés pour répéter pour un disque qu'ils ne voyaient pas venir. Et je ne sais pas ce que c'est devenu, **ce truc**, mais en tout cas **cela** sonnait merveilleusement. (Perec, G., *Entretiens et conférences II [1979-1981]*, 2003, < Frantext)

Le référent visé successivement par *ce truc* et *cela* se construit à partir d'un nombre important d'objets et de relations entre eux. Il s'avère ainsi vain de cibler dans ce cas un élément initiateur.

---

<sup>88</sup> Ce repérage de l'antécédent s'explique par des objectifs spécifiques aux domaines, dont voici deux illustrations (respectivement en TAL et en psycholinguistique) : « L'anaphore est une relation linguistique entre deux entités textuelles définie lorsqu'une entité textuelle (l'anaphore) renvoie à une autre entité du texte (l'antécédent). Comme la présence d'anaphores dégrade considérablement les performances des systèmes de TAL, la question de leur résolution est étudiée depuis longtemps. » (Weissenbach & Nazarenko 2007: 146) ; « [...] third-person pronouns typically have explicit antecedents in the preceding text. [...] Our question is, does an Institutional *They* cause a processing problem ? It has been demonstrated that there are costs for missing antecedents [...] ». (Sanford *et al.* 2008: 372-373)

<sup>89</sup> Voir à cet égard les critiques de Valentin (1985) ou par exemple de Cornish (2010a, 2017).

On peut également mettre en cause la métaphore de la « chaîne », dans le sens où elle appuie l'idée d'immuabilité et d'homogénéité d'une séquence de « chaînons » ou « maillons » interdépendants et bien identifiés. Elle donne l'impression que les relations de coréférence sont contrôlées et déterminées à la base. Bien que cette notion puisse sembler fonctionnelle pour l'analyse de types de textes particuliers (Schneidecker 2014), elle nous paraît mal refléter les usages réels des locuteurs, *a fortiori* si l'on tient compte de l'oral spontané, où la gestion de la référence s'effectue sur le vif, avec toutes les approximations que cela peut comporter.

L'avènement des bases de données et l'intérêt pour les descriptions linguistiques en contexte, intégrant des aspects pragmatiques, communicatifs et interactionnels du discours, a contribué à remettre en cause cette approche essentiellement segmentale de l'anaphore (nominale ou pronominale). La notion d'antécédent, et surtout le rôle qu'on lui attribue, en prennent sérieusement pour leur grade au vu des situations suivantes :

- i) les cas de « discordances morpho-syntaxiques, sémantiques, et/ou référentielles » ([Reichler-]Béguelin 1988 : 18) :

(99) La couleur, c'est la liberté. **Elles** se mélangent, se superposent, s'opposent et se répondent. (*Marie Claire*, avril 1993, début de texte < Béguelin 1997a : 34)

Dans un exemple de ce type, en effet, le pronom affiche non seulement un « désaccord » morphologique avec « l'antécédent », mais témoigne également d'une modification implicite de perception du référent (d'abord appréhendé comme un type, puis comme la classe de ses occurrences). Le processus interprétatif ne peut donc se limiter à un simple phénomène de reprise.

- ii) Les cas où, tout comme en (98) *supra*, un antécédent est difficilement délimitable :

(100) Quelques camions faisaient des aller et retour entre la montagne et la ville. Mais **ces navettes** demeuraient évidemment dérisoires. (*Le Monde*, 9. 4. 1991 < [Reichler]-Béguelin 1995a)

Le choix lexical de l'anaphorique *ces navettes* repose, non seulement sur le SN *quelques camions*, mais aussi crucialement sur la prédication 'faire des aller et retour'. Considère-t-on le SN, le prédicat, ou alors l'énoncé entier comme « l'antécédent » ? La délimitation d'un segment responsable de l'interprétation demeure ainsi incertaine.

- iii) Les cas où un antécédent est absent :

(101) on a fait aussi un saut en parachute à San Diego | \_ | c'était assez euh | \_ | assez incroyable surtout que | \_ | ouais déjà on arrive là-bas euh | \_ | **ils** nous ont presque pas donné d'explications | \_ | et puis euh tout d'un coup **ils** nous mettent un | \_ | un espèce de gros euh | \_ | un espèce de gros baudrier avec deux grosses sangles qui passent sur les épaules (oral, ofrom)

Dans l'exemple (101) il n'y a tout bonnement aucun antécédent à rattacher à l'élément dit anaphorique. Face à ce genre de situations, la notion d'antécédent se montre inopérante ou nécessite des ajustements fondamentaux<sup>90</sup>.

Cornish (1999 : 41 *sqq.*) recourt pour sa part à la notion de *déclencheur d'antécédent* (*antecedent trigger*), qui désigne le support segmental, pour autant qu'il existe, qui a rendu disponible à un moment donné le référent visé par l'anaphorique<sup>91</sup>. Mais il montre, au-delà du rôle du déclencheur, que ce sont avant tout les propriétés sémantico-référentielles intrinsèques de l'expression anaphorique et son contexte prédicatif qui orientent au départ l'interprétation référentielle. La notion d'antécédent, au sens traditionnel, ne tient ainsi pas compte de l'aspect évolutif de l'objet-de-discours, entre le moment de son introduction et son rappel, l'évolution dynamique étant une caractéristique fondamentale des approches constructivistes de la référence et permettant d'expliquer les nombreux changements observés (entre autres Cornish *ibid.*, Apothéloz 1995b, Apothéloz & Béguelin 1995, Béguelin 1997a, 1997b, Berrendonner 1983, Berrendonner 1990a, Corminboeuf 2011, Groupe de Fribourg 2012, etc.). Les quelques exemples ci-dessus reflètent ainsi pleinement ce point de vue, où l'élément anaphorique n'a pas à être considéré comme un élément simplement subordonné à un autre segment. Au contraire, il est doté de caractéristiques sémantico-référentielles propres et susceptible de (re)configurer l'objet en contexte.

### 5.5 Conception cognitive de l'opposition anaphore vs deixis

Un autre problème soulevé à l'égard de la conception textualiste de l'anaphore et partant, de son opposition à la deixis comme référence extra-textuelle, est que la distinction ne se vérifie pas au niveau de la distribution des expressions référentielles : un grand nombre de formes (SN définis, démonstratifs, pronoms personnels et possessifs de 3<sup>e</sup> personne, etc.) endossent en discours les deux rôles, autrement dit, fonctionnent aussi bien « déictiquement » qu'« anaphoriquement » :

Non seulement les expressions recrutées comme anaphoriques ou déictiques par le critère de localisation ne sont pas homogènes, mais encore elles ne sont pas spécialisées pour l'une ou l'autre tâche. C'est dire que si l'on s'en tient au seul critère du lieu de résidence il n'y a aucune expression qui ne soit uniquement anaphorique et aucune expression qui ne soit uniquement déictique<sup>92</sup>. Article défini, adjectif et pronom démonstratif, adjectif possessif, pronom personnel de la troisième personne connaissent en effet les deux types d'emplois [...] (Kleiber 1992 : 616).

---

<sup>90</sup> Voir notamment [Reichler-]Béguelin (1995a) et Cornish (2010b) pour une critique des approches segmentales de l'anaphore.

<sup>91</sup> Quant à la notion d'*antécédent*, Cornish (1999 : 44 *sqq.*) la redéfinit au niveau du contenu (« discourse model description »). Il décrit sa fonction de la manière suivante : « providing a full interpretation for the semantically non-autonomous anaphor ».

<sup>92</sup> Il faut cependant reconnaître que les indices personnels de 1<sup>ère</sup> et 2<sup>e</sup> personnes font exception à ce double fonctionnement, renvoyant aux participants de l'énonciation, bien que l'identité de l'énonciateur doive parfois être inférée (*cf. supra* §4.4).

Un traitement éclaté des formes référentielles apparaît ainsi comme une analyse trop coûteuse à laquelle d’aucuns préfèrent renoncer. Le nœud du problème se situe dans le critère de localisation du référent (extra- ou intra-linguistique), manifestement non opératoire pour l’analyse. A celui-ci, certains chercheurs préfèrent un critère d’ordre cognitif : un retour aux sources, selon Bosch (1983), vers la conception d’Apollonius Dyscole d’après laquelle (*cf. supra* 5.1) la deixis est le mode d’accès à une connaissance nouvelle tandis que l’anaphore est le mode de maintien d’un référent déjà connu. Lyons (1977 : 673) s’inscrit dans cette approche :

[a]naphora presupposes that the referent should already have its place in the universe-of-discourse. Deixis does not; indeed deixis is one of the principal means open to us of putting entities into the universe-of-discourse so that we can refer to them subsequently.

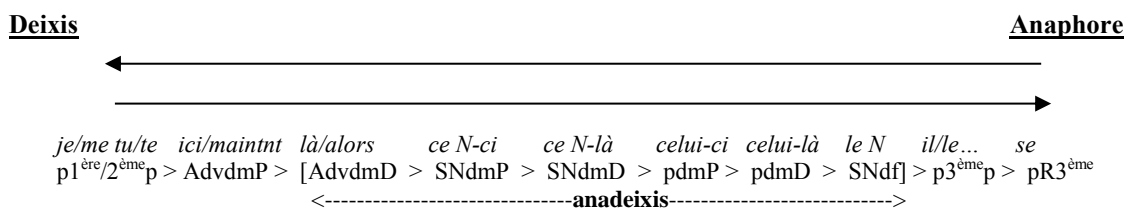
De même, Ehlich (1982) définit de la sorte les deux opérations référentielles :

The anaphoric procedure is a linguistic instrument for having the hearer continue (sustain) a previously established focus towards a specific item on which he had oriented his attention earlier. (p. 330)

The deictic procedure is a linguistic instrument for achieving focusing of the hearer’s attention towards a specific item which is part of the respective deictic space (*deiktischer Raum*). (p. 325)

Dans la foulée de ces travaux, Cornish (2010a) propose une échelle de *phoricité* ou d’*indexicalité* sur laquelle il place les différentes formes linguistiques selon leur degré d’anaphoricité ou de déicticité : certaines expressions sont vouées, respectivement, au maintien de l’attention sur un objet, d’autres au revirement de l’attention sur un élément référentiel nouveau :

**Figure 5** : Echelle de *phoricité* d’après Cornish (2010a)



Entre les deux pôles, certaines formes manifestent les deux propriétés (situation appelée *anadeixis*) dont l’effet est de placer au centre de l’attention un référent jusque-là en retrait. Cornish commente l’exemple suivant :

(102) Âgé de 19 ans, Adrien (Thomas Langmann), mêlé à un vol d’argent, s’enfuit de Bordeaux où **il** vit avec sa mère, et arrive à Paris chez son père, Clément (Jean-Pierre Léaud), qu’**il** n’a pas vu depuis quatre ans. **Ce père** est un être immature, qui se raccroche à sa jeunesse perdue en ayant une maîtresse de 19 ans, Louise (Judith Godrèche). **Elle** rêve de devenir comédienne ou, tout au moins, speakerine à la télévision. Clément ne perd pas une occasion de la rabaisser. Pour tenir, elle se drogue (...) (*Le Monde Supplément Radio –Télé*, 29-30 avril 2007 < *ibid.* : 123)

Dans cet exemple, l'usage de *il* est décrit comme typiquement anaphorique, car maintenant au centre de l'attention un référent bien établi, à savoir l'individu nommé *Adrien*. Le SN démonstratif *ce père*, d'après Cornish, est anadéictique en ce qu'il réfère à un individu moins imposant dans l'attention des interlocuteurs, bien que récemment introduit : le SN a ainsi pour fonction de récupérer ce référent en retrait, pour le placer au centre de l'attention, de même qu'il a une fonction désambiguïsatrice (deux référents étant en concurrence). A cet égard, l'auteur juge maladroit l'emploi du pronom *elle* plus bas pour désigner un référent au statut semblable, c'est-à-dire, qui vient d'être introduit mais encore non central dans l'attention des lecteurs. A ses yeux, une expression anadéictique telle que *celle-ci* ou *cette dernière* semblerait « plus naturel[le] ».

Comme nous l'avons déjà relevé *supra* (§4.6.3) à l'égard d'autres modèles fondés sur un critère cognitif<sup>93</sup>, il nous paraît cependant regrettable de traiter l'emploi d'une expression sur cette seule dimension, en l'occurrence les critères de continuité de l'attention ou de nouveauté, pour relever après coup l'inadéquation de données qui s'en écartent. A nos yeux, l'exemple (102) illustre davantage la diversité des paramètres qui peuvent favoriser l'emploi de telle ou telle expression. Ainsi, sans préjuger d'une quelconque adéquation, il nous paraît tout à fait intéressant de constater que deux référents peuvent précisément induire des moyens de récupération différents (*ce père* vs *elle*) alors qu'ils ont été introduits dans des conditions très semblables en (102), à savoir une activation récente parmi d'autres référents déjà établis, dans un type de *progression linéaire* (Combettes 1983), où le « rhème » d'un énoncé devient « thème » de l'énoncé successif, selon la terminologie de l'école de Prague. Il conviendrait en outre de voir si dans le genre discursif en question (synopsis de film), ce type de progression peut être mis au service d'effets de point de vue dans la présentation des personnages<sup>94</sup>.

Pour en revenir à l'extrait, tandis que lors du premier pointage (*ce père*), la concurrence des référents potentiels favorise sans doute l'emploi d'une forme lexicalement désambiguïsatrice, la forme *Elle* est jugée quant à elle suffisamment univoque pour atteindre la cible et même réaliser la « rupture de thème ». A noter par ailleurs qu'il n'est ici pas exclu que le pronom soit de type disjoint<sup>95</sup>. Quoiqu'il en soit, en mettant en jeu des référents au statut d'activation analogue mais rappelés par des moyens différents, cet extrait montre justement que les facteurs qui pèsent sur le choix de la forme ne peuvent être ramenés au seul critère cognitif.

D'après l'échelle proposée, en outre, on ne voit pas très bien en quoi les pronoms *je* ou *tu* introduisent des référents nouveaux dans le discours, étant donné qu'en général, ils renvoient au locuteur et à l'interlocuteur de l'énonciation en cours, donc à des référents ou rôles établis

---

<sup>93</sup> Entre autres spécificités théoriques et méthodologiques, les modèles d'Ariel (1988) et de Gundel *et al.* (1993) n'exploitent pas directement les notions d'*anaphore/deixis*.

<sup>94</sup> Béguelin (2013) montre dans une étude sur l'emploi des pronoms de 3<sup>e</sup> personne chez Flaubert que ce genre de rupture du thème par progression linéaire cher à l'auteur, par ailleurs critiqué par Proust, contribue à la création de divers effets discursifs (point de vue, maintien d'une équivoque, etc.).

<sup>95</sup> Comme dans cet exemple : *Lui rêve de liberté et de grands espaces...* (présentation d'une pièce de théâtre, <http://www.theatre-poudriere.ch/les-accueils/saison-2016-2017/a-venir>)

dans le savoir partagé par le fait de l'énonciation même. En cela, ils sont certes token-réflexifs, mais l'aspect « nouveau » nous paraît loin d'être évident dans le processus.

## 5.6 *Bilan*

L'opposition anaphore-deixis, on le voit, soulève de sérieux problèmes et est souvent source de confusion. Ni la conception traditionnelle, se fondant sur l'opposition texte/situation, ni l'approche cognitive, déplaçant l'axe sur le critère nouveauté/maintien du référent, ne parviennent à expliquer la distribution des expressions référentielles. A notre avis, la raison en est la prise en compte d'une seule dimension, là où de nombreux paramètres contextuels entrent en jeu. En somme, nous ne nions pas l'intervention de paramètres cognitifs, mais il nous semble que leur rôle doit être relativisé afin d'offrir une place plus importante à des facteurs plus conjoncturels de l'énonciation en cours. En appariant une procédure à une forme particulière, on donne à croire que c'est l'expression elle-même qui confère une certaine nature au référent<sup>96</sup> (inédit, donné, etc.), alors que c'est évidemment le contexte (*cf. infra* §6.3.3) dans son ensemble qui détermine la place de celui-ci dans le discours et son rapport aux autres objets. En accordant trop d'importance à un seul type de critère, on court ainsi le risque d'établir une norme à partir de laquelle on évalue les usages et celui de marginaliser ceux qui s'en écartent. Malgré nos réserves à l'égard de l'échelle proposée, nous nous sentons davantage en accord avec les observations de Cornish (à par.), sur l'intervention cruciale de paramètres liés à la subjectivité, l'intersubjectivité et la polyphonie dans le choix des expressions.

Par ailleurs, il nous semble que les propriétés communes à la deixis et à l'anaphore mériteraient une plus grande attention, plutôt qu'un traitement systématiquement oppositif. Dans tous les cas, on observe en effet un procédé de *pointage* vers un référent (*cf.* l'étymologie du terme *deixis*) dont, certes, le mode d'accès (situation immédiate, inférence, texte, etc.) peut varier. Mais si les mêmes expressions permettent l'accès à des référents issus de diverses sources, ne devrait-on pas conclure d'abord, au-delà des spécificités sémantiques illustrées *supra* (§4), à un processus fondamentalement commun ?

S'il fallait retenir un aspect de l'opposition *anaphore-deixis*, nous invoquerions la piste de la token-réflexivité proposée par Kleiber (1992 : 623). En effet, parmi les expressions donnant l'instruction de renvoyer à un objet appartenant au savoir partagé, nous avons relevé *supra* (§4) la nature token-réflexive qui caractérise certaines d'entre elles. Par exemple, un pronom comme *je*, à l'inverse de *il*, implique nécessairement la prise en compte de sa propre occurrence pour son interprétation (sans qu'il y ait pour autant de nouveauté à cet égard, *cf.* Cornish ci-dessus), le locuteur réflexivement désigné étant partie prenante de l'énonciation en cours. Pour le reste, nous prenons nos distances aussi bien vis-à-vis de l'approche selon la localisation du référent que de l'approche en termes de nouveauté/continuité, car nous

---

<sup>96</sup> Voir Kibrik (2011 : 389) pour une critique similaire de la circularité des certaines approches cognitives : « [...] they infer referent activation from the form of reference. »

considérons que chaque emploi est le résultat d'un calcul plurifactoriel tenant compte des stratégies et besoins des locuteurs selon divers objectifs communicationnels.



## 6. Vers un modèle constructiviste du discours

Dans les sections précédentes, nous avons présenté un certain nombre de fondements et d'outils ancrés dans la tradition sémantique qui traite des expressions référentielles. Il n'est peut-être pas inutile à ce stade de clarifier les préoccupations respectives des divers acteurs qui s'illustrent dans le domaine de la référence. Ce sont les philosophes, les premiers, qui se sont intéressés au lien entre langage et monde extérieur et à la capacité pour celui-là de représenter celui-ci (*cf. supra* §2). Dans leur perspective, une attention toute particulière est consacrée aux notions fondamentales de vérité et de validité des énoncés, de même qu'aux questions d'identité ou encore de croyance. Dans la mesure où le langage est voué à refléter le monde, il va de soi que la mesure de sa conformité et la quête de la vérité, en débusquant les obstacles, restent au cœur des préoccupations philosophiques.

En linguistique, la sémantique est fortement tributaire de ces idées, mais l'intérêt porte davantage sur les conditions d'emploi des expressions référentielles. On se demande par exemple quel est le sens véhiculé par telle ou telle expression par opposition à telle autre dans un même énoncé ou entre deux énoncés. On juge également la plausibilité de tel ou tel emploi dans un univers d'interprétation, en fonction de l'état du monde. Diverses manipulations d'énoncés sont soumises à l'intuition du linguiste afin de mettre au jour les diverses contraintes inscrites dans le signifié des formes étudiées.

Si l'on se donne pour objectif, comme dans ce travail, de décrire les opérations de référence en discours, il est nécessaire de se doter d'outils adéquats pour rendre compte, au-delà des contenus « littéraux » et de la fonction significative des expressions, des facteurs contextuels qui déterminent leur emploi dans une interaction. Nous partons ainsi de l'idée élémentaire que les participants d'une interaction verbale mettent en œuvre un discours dont la fonction est de créer un ensemble de représentations partagées (Groupe de Fribourg 2012 : 22-23). Dans le dernier quart du XX<sup>e</sup> siècle se sont développées, sur la référence et sur la nature du discours, des réflexions qui se démarquent des cadres de représentation logique du contenu des énoncés. Avant de présenter en détail le cadre d'analyse dans lequel nous nous situons (§6.3), revenons sur les origines de quelques notions fondamentales qui mettent en évidence la nature discursive des référents (§6.1) et plus généralement l'incidence cognitive de l'activité discursive (§6.2).

### 6.1 *Les référents discursifs*

On doit à Karttunen (1976) la notion de *discourse referent*. Par cette dénomination, l'auteur tient à se démarquer explicitement des approches philosophiques de la référence : « [w]e maintain that the problem of coreference within a discourse is a linguistic problem and can be studied independently of any general theory of extralinguistic reference » (1976 : 366). Ce travail s'inscrit dans les travaux générativistes sur les indices référentiels en grammaire transformationnelle et a pour ambition de contribuer à dégager les contraintes formelles de coréférence au-delà de la phrase. Plus précisément, Karttunen met au jour les conditions

dans lesquelles un SN indéfini est capable d'établir un référent de discours qui puisse faire l'objet de références ultérieures dans le texte. La notion de *discourse referent* permet à Karttunen d'éviter les considérations ontologiques des philosophes à propos des référents extra-linguistiques et de demeurer dans un cadre strictement linguistique, déterminant les potentialités de coréférence dans la suite du texte. Sans réel écho en grammaire transformationnelle, cette notion, qui n'émane toutefois pas d'une réflexion sur la nature même des référents, sera exploitée dans des théories d'horizons divers, par exemple en sémantique formelle ou en linguistique computationnelle. Dans le cadre de la DRT, Kamp (1981) tente de modéliser la manière dont les interlocuteurs mettent à jour de manière dynamique leur interprétation du discours. Heim (1983) utilise quant à elle la métaphore des *file cards* (qu'on pourrait traduire par « fiches signalétiques ») pour représenter les référents de discours, sur lesquelles toutes les prédications opérées sont inscrites au fil des énoncés. Dans ces cadres d'analyse, on peut constater que la progression du discours demeure au niveau sémantique : l'évolution dynamique des référents se fait au gré des prédications et se conçoit comme une sorte de compilation de signifiés. Il en va de même dans la visée computationnelle de Sidner (1983), qui tire parti de la notion de *discourse entities* afin d'élaborer des algorithmes de résolution automatique qui soient informatiquement viables. En termes informatiques, une entité discursive représente une variable qui permet d'indexer l'information extraite des énoncés à la représentation mentale correspondante (Wolters 2001 : 11). Dans ces approches formelles, la notion de référent discursif se voit exploitée pour cataloguer les diverses informations accumulées dans le flux verbal, à des fins de représentation logique ou de traitement automatique des énoncés.

## 6.2 Une représentation du discours

Une réflexion inédite se développe parallèlement autour de la nature même des processus mentaux à l'œuvre dans les opérations de raisonnement naturel. Dans cette perspective, Johnson-Laird élabore la notion de *modèle mental* (1983, 1996, 2010). Les modèles mentaux sont la représentation cognitive iconique – ce qui veut dire que la structure de la représentation mentale est du même type que celle de la situation qu'on se représente – de situations, réelles ou hypothétiques, que construisent les êtres humains en vue de raisonner. Cette représentation est élaborée sur la base de leur perception, de leur imagination, de leurs connaissances du monde ou encore du discours. La nature des représentations ainsi produites est donc bien distincte d'une représentation « propositionnelle » logique des modèles formels. L'input discursif, notamment, donne lieu à un *modèle du discours* :

A discourse model is a mental object that constitutes an individual's knowledge of a discourse. It is constructed on the basis of what has occurred in the discourse supplemented by general and specific knowledge. Questions may be answered by recourse to the information that it contains; assertions may be evaluated by reference to it ; utterances and actions may be based upon it. We do not suppose that mental models are constructed only from discourse: they are also created from memory, perception, imagination, and the operation of other mental processes. (Johnson-Laird & Garnham 1980 : 371)

L'originalité de cette approche est d'intégrer des inputs de diverses natures à la construction des représentations. La prise en compte de l'incidence mentale du discours s'observe également dans toute une série de travaux sur la manière dont les éléments du texte servent d'instructions pour la construction du modèle :

(...) let us say that a TEXT is a set of instructions from a speaker to a hearer on how to construct a particular DISCOURSE-MODEL. The model will contain DISCOURSE ENTITIES, ATTRIBUTES, and LINKS between entities. A discourse entity is a discourse-model object [...]. (Prince 1981: 235)

Il en ressort notamment que les référents du discours y figurent avec des statuts cognitifs distincts, comme *donné, nouveau* (Chafe 1976, Clark & Haviland 1977), ou encore *inférable* Prince (*ibid.*). C'est sous cette impulsion que voient le jour les échelles cognitives présentées *supra* (Ariel 1988, 1990, 2001, Gundel *et al.*1993), à propos desquelles nous avons déjà soulevé l'inconvénient d'assortir les différentes expressions référentielles d'un statut déterminé. L'un des aspects lié à la réalité psychologique que recouvre la notion de modèle discursif est que chaque interlocuteur construit son propre modèle sur la base des indices à disposition. Nous reviendrons *infra* (§6.3.4) sur la question du caractère individuel ou commun des représentations des interlocuteurs.

Dans une optique distincte de l'approche « psychologique », Grize (1974, 1982, 1993 1996) élabore à la même période la notion de *schématisation*, qui vise pour sa part la modélisation d'opérations *logico-discursives* :

Si dans une situation donnée, un interlocuteur A adresse un discours à un interlocuteur virtuel B (dans une langue naturelle), je dirai que A propose une schématisation à B, qu'il construit un micro-univers devant B, univers qui se veut vraisemblable pour B (1982 : 172)

La notion de schématisation recouvre à la fois l'action de schématiser et son résultat (Grize 1996 : 69) : un locuteur construit une représentation du discours pour son allocutaire et celui-ci s'attache à la reconstruire. Grize (1993) se démarque par là explicitement de la logique formelle, vouée à représenter la pensée mathématique, en proposant des instruments pour analyser la « pensée commune » : un discours en langue naturelle est le fait de sujets interagissant dans un cadre spatio-temporel, guidés par une visée argumentative s'inscrivant dans un contexte social. En intégrant ces aspects énonciatifs, le contenu d'une schématisation se révèle ainsi complètement réfractaire à une formalisation en termes de valeurs de vérité.

Le modèle auquel nous adhérons, présenté ci-dessous, s'inspire de cette conception de la schématisation, en conservant sa dimension créative et représentationnelle, ainsi que la complexité de la structuration discursive, irréductible à un fonctionnement vérifonctionnel (*supra* §6.3.4). Berrendonner (1997 : 219-220) formule les avantages suivants du modèle « inédit » de Grize :

Derrière le terme schématisation, il y a l'idée que le discours n'a pas pour fonction de restituer un tableau vérifonctionnel de quelque réalité préexistante, absolue et indépendante de lui, mais plutôt d'imposer ses propres objets, en

construisant une fiction conceptuelle originale, provisoire et évolutive [...] La notion porte donc en elle une théorie de la référence et du contexte qui tranche de façon radicale avec un certain chosisme ambiant. Elle suppose en effet qu'au lieu d'assimiler les référents du discours aux *realia* en se recommandant du bon sens ou de Frege, on leur reconnaisse le statut de représentations cognitives, de schémas mentaux, dotés d'une structure formelle dont la description est affaire de logique et de sémiologie. Quant au contexte, elle conduit à y voir non pas un cadre informationnel ou situationnel fixé à titre de préalable, mais le produit dynamique de l'activité de communication : un capital évolutif de connaissances, d'hypothèses et d'assumptions partagées, assimilable à une sorte de mémoire collective des interlocuteurs.

### 6.3 Le modèle du discours fribourgeois

Nous prenons pour cadre d'analyse le modèle du discours fribourgeois (Groupe de Fribourg : 2012) dont certains aspects ont fait l'objet d'études particulières (Berrendonner 1983, 1990a, 1990b, 1994, 1995, 1997, 2002b, 2005, 2014, à par. ; [Reichler-]Béguelin 1988, 1989, 1993a, 1995, 1997a, 1997b ; Apothéloz 1995a et b ; Apothéloz & Béguelin 1995, 1999 ; Corminboeuf 2011), notamment en ce qui concerne la description des procédés référentiels.

L'objectif des chercheurs fribourgeois est de proposer un modèle général du discours pour l'étude des faits linguistiques. Ils proposent d'une part, face à l'inconsistance des notions usuelles et graphocentriques (*mot*, *phrase*, etc.) pour la segmentation des unités linguistiques, une refonte complète de la théorie des articulations du discours et de ses constituants<sup>97</sup> (*cf. infra* §6.3.3) ; d'autre part, une description des opérations qui sont exécutées dans le modèle. Nous nous intéresserons avant tout dans ce travail aux opérations de pointage au moyen de désignateurs (§6.3.5) et à leurs motivations (§6.3.6). Avant de présenter les composantes spécifiques du modèle, il est nécessaire de comprendre quelques hypothèses fondamentales du programme de modélisation fribourgeoise (§6.3.1).

#### 6.3.1 Principes méthodologiques

##### 6.3.1.1 Le rapport oral-écrit

Pour prétendre à la généralité scientifique, l'un des choix est tout d'abord de renoncer à dissocier « langue orale » et « langue écrite » ou à une approche en termes de « diglossie »<sup>98</sup>. Les spécificités observées au niveau des structures produites via l'un ou l'autre médium tiennent avant tout aux conditions de production respectives, en particulier liées à l'opposition entre « produit fugitif » vs « produit permanent » (Béguelin 2012 : 47) : dans le premier se manifestent les traces d'élaboration et de planification en cours, qui se traduisent notamment par la présence de bribes, hésitations, abandons ou autres phénomènes de « disfluence ». Ces caractéristiques du discours improvisé sont régulièrement perçues comme des « accidents de performance », alimentant le préjugé selon lequel la langue orale serait par nature fautive (*cf.*

---

<sup>97</sup> L'une des conséquences en est l'abandon de la notion de *phrase* (Groupe de Fribourg 2012 : ch. 1) au profit de la notion de *clause*, unité micro-syntaxiquement autonome. *Cf.* aussi Blanche-Benveniste *et al.* (1984 : 24-25) et leur notion d'*unité de construction*.

<sup>98</sup> *Cf.* aussi Blanche-Benveniste (1997 : 65).

Blanche-Benveniste 1997 : 11) ; le second type de production se voit quant à lui généralement débarrassé de ces « scories » (possibilité de préparer, retoucher, effacer, etc.), guidé par des normes rédactionnelles et stylistiques sous-jacentes à des genres centrés sur les intérêts de l'interprétation (Béguelin 2012). Les différences relevées sont ainsi davantage à mettre au compte des conditions propres à la variété des genres discursifs, où le degré d'improvisation, de préparation ou de compétence, le « niveau de langue », etc. sont loin d'être dédiés à l'un des deux média<sup>99</sup> (Koch & Oesterreicher 1985). Sans pour autant négliger l'existence et la prise en compte, dans un second temps, de ces paramètres contextuels, le modèle part ainsi de l'hypothèse fondamentale d'une grammaire unifiée<sup>100</sup>, en accordant une égalité de traitement aux données issues de l'oral, longtemps laissées pour compte en linguistique (à l'exception de l'attention particulière qu'y a porté le GARS<sup>101</sup>).

### 6.3.1.2 *Le statut des données*

Une autre conséquence de la vocation généralisante du modèle réside dans le choix de considérer des données effectivement produites par les usagers : il s'agit « d'en prendre acte, et de rechercher une explication au fait qu'on les utilise pour communiquer » (Berrendonner 1990a : 151). De la sorte, les chercheurs reconnaissent l'égalité de « l'existence linguistique » ([Reichler-Béguelin 1993a : 328) de toutes les occurrences langagières, y compris celles qui peuvent être considérées comme atypiques, voire déviantes par rapport à une norme ambiante. Plutôt que de les mettre à l'écart, ils cherchent à mettre au jour leur raison d'être, susceptible de révéler des régularités sur les comportements des usagers, voire aussi l'explication des réactions de censure dans des genres normés, comme à l'égard de l'exemple ci-dessous produit dans un cadre scolaire :

(103) Il est vrai que lorsque nous lisons, nous ne pensons pas que **cette histoire** est en train de vivre, de prendre forme grâce à nous. (copie de bac < ([Reichler-]Béguelin 1988 : 27)

Dans cette copie d'élève, on peut visiblement imputer l'appréciation de l'évaluateur au coût de traitement requis au décodage, étant donné qu'en début de texte, l'histoire en question n'a pas été introduite de manière explicite. Il n'en reste pas moins que le choix d'une expression démonstrative crée ici des effets discursifs d'intersubjectivité qu'il est pertinent de mettre en regard, par exemple, avec d'autres extraits d'incipit où le démonstratif ne prête sans doute pas à discussion, notamment du fait de la reconnaissance littéraire dont jouit l'auteur :

---

<sup>99</sup> On peut penser à la parole publique, au dialogue théâtral, aux interventions sur les forums ou blogs, au *chat*, aux SMS, etc.

<sup>100</sup> Les spécificités de chaque médium sont toutefois bien connues à d'autres niveaux : au niveau des démarcations respectives, ponctuation et prosodie ne se recouvrent pas ; s'agissant de la morphologie, les marques de personne (flexion verbale) ou de genre et de nombre (*e.g.*, flexion adjectivale) pour ne prendre que ces exemples, se manifestent différemment à l'oral qu'à l'écrit. Voir à cet égard Béguelin (2012 : 46) qui synthétise les positions de Claire Blanche-Benveniste sur le rapport oral-écrit.

<sup>101</sup> Groupe aixois de Recherche en Syntaxe, dont faisait partie Claire Blanche-Benveniste, qui a publié la revue *Recherches sur le français parlé* de 1974 à 2004.

(104) **Cette histoire** ne m'appartient pas, elle raconte la vie d'un autre. (Amin Maalouf, *Les échelles du Levant*)

Dès lors, les partisans du modèle fribourgeois sont extrêmement réservés à l'égard des méthodes « de fauteuil » et de leur recours aux données décontextualisées et fabriquées :

une donnée linguistique *in vitro* est une donnée incomplète, sur laquelle il est extrêmement délicat de porter un jugement d'acceptabilité ; [...] *la seule voie scientifiquement valide et épistémologiquement acceptable est l'observation des comportements linguistiques effectifs* (Apothéloz & Béguelin : 234)

Le critère du jugement d'acceptabilité, en particulier, est mis en cause :

[il] autorise à trier entre données « pertinentes » et données « non pertinentes » sur une base qui, bien souvent, reste purement intuitive. Le problème, insidieux, naît chaque fois qu'est déclarée mal formée, et donc indigne d'être prise en compte dans la description, telle ou telle structure pourtant utilisée, parfois couramment, par les sujets parlants. Le procédé a ceci de dangereux qu'il aboutit à confondre sous le même astérisque disqualificateur certains artefacts proprement irréalistes, n'ayant aucune chance d'être un jour performés, avec des énoncés qui sont au contraire accessibles à l'observation empirique, même si leur forme ne correspond pas en tout point à la norme dominante. [Reichler-]Béguelin (1993b)

Les chercheurs fribourgeois critiquent non seulement le privilège octroyé aux données forgées, mais aussi l'intérêt porté presque exclusivement à des genres de discours consensuels (en particulier l'écrit standardisé), peu représentatifs de la réalité des discours. Cette « représentation déformée » des faits de langue comporte ainsi un risque de « circularité du rapport entre les modèles et les données, les premiers légitimant les secondes et celles-ci n'étant reconnues comme valides et donc comme "acceptables" que dans la mesure où elles sont conformes à ce que prévoient les premiers ! » (Apothéloz & Pekarek Doehler 2003 : 112)

Dans le domaine de la sémantique référentielle, une pratique commune est la délimitation du champ d'observation à deux phrases, où « l'introducteur » et « le pointeur » interviennent dans des positions typiques :

À force de neutraliser les paramètres énonciatifs et contextuels, l'approche classique accrédite aussi une conception étroitement textualiste et segmentaliste des phénomènes anaphoriques (*ibid.*)

Il est important de préciser que l'approche fribourgeoise, bien qu'empirique car se fondant sur l'observation de faits attestés, ne s'inscrit pas dans la tradition de la linguistique de corpus : l'ensemble des données examinées ne répondent pas, pour l'instant du moins, à la définition stricte de *corpus* telle que l'entend par exemple Sinclair (1996) : « [a] corpus is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language ». D'abord, les données retenues, comme on l'a vu, ne sont pas systématiquement homogènes (en termes de genre discursif, de type de locuteurs, etc.) et ne prétendent pas à la représentativité d'un type de population. La méthode de récolte des exemples repose chez les Fribourgeois sur des critères qualitatifs de pertinence, à savoir leur capacité à révéler des aspects sur le fonctionnement général du discours. Bien

que l'extraction de données soit parfois utilisée pour l'étude de telle ou telle structure ciblée, elle n'en représente pas pour autant une exigence fondamentale. En effet, le risque d'une recherche exclusivement automatisée est de négliger tous les phénomènes qui ne sont pas détectables par une machine et qui, pourtant, révèlent des régularités sur le fonctionnement de la langue :

Si la recherche en linguistique ne porte que sur des faits langagiers extractibles de manière automatique (en particulier les « marqueurs de discours »), on se prive du fonctionnement d'une bonne partie du système. (Corminboeuf 2014 : 2378).

En somme, la valeur heuristique accordée aux données attestées, sélectionnées sur une base qualitative, distingue l'approche fribourgeoise de nombreux modèles. En particulier, celle-ci se démarque nettement de démarches consistant à mettre au jour les contraintes qui pèsent sur l'emploi de marqueurs, constructions, enchaînements, etc., via des manipulations d'énoncés dans le but de réduire les « possibilités » théoriques. Les Fribourgeois renversent la méthode en se fondant sans a priori sur les observables à disposition pour y déceler les régularités de fonctionnement. En d'autres termes, ils partent du principe que les comportements locutoires, rares ou fréquents, sont susceptibles de révéler les mécanismes sous-jacents de l'activité discursive, et notamment, sur la façon dont on catégorise les objets dont on parle.

### *6.3.1.3 Une vision non déterministe du rapport monde-langue*

Dans le domaine des procédés référentiels, un exemple de la démarche esquissée ci-dessus est la manière dont sont traités les faits d'« indiscrétion » (*cf. supra* §3.3). La conséquence qui en est tirée est une conception non logiciste de l'identité et une certaine malléabilité dans la façon dont les usagers représentent les référents du discours, dont certains formats présentent une disposition à la confusion référentielle. Cette approche se démarque de nombreux travaux traitant de la problématique des référents dits *évolutifs*.

Pour résumer, la question des référents évolutifs concerne des processus de transformation encourus par des référents par voie de prédication au fil d'un discours et la manière dont se comportent en particulier les marqueurs anaphoriques à cet égard. L'enjeu est « de mesurer jusqu'à quel point certaines formes de reprise sont sensibles au changement » (Charolles 1997 : 71). Si dans le contexte des « métamorphoses fictionnelles (du genre Dr Jekyll/Mr Hyde) », l'emploi des expressions référentielles a été, à juste titre, expliqué en partie par des effets de point de vue ou de changement d'univers de croyance (Schneidecker & Charolles 1993, Achard-Bayle 2001), le cas des modifications affectant une entité dans les textes de type injonctif (recettes de cuisine, notice de fabrication, etc.) est généralement abordé comme impliquant des transformations de type « ontologique » (Charolles 1997).

L'un des points de départ de la réflexion est la critique de Brown & Yule (1983) à l'encontre du traitement du pronom par Halliday & Hasan (1976) dans l'énoncé :

(105) Lavez et évidez six pommes à cuire. Mettez-*les* dans un plat passant au four. (trad. par Charolles & Schnedecker 1993 : 108)

Brown & Yule (1983) reprochent à la conception « segmentale » d'Halliday & Hasan le fait de ne pas tenir compte de l'évolution des référents entre les deux étapes décrites. Il en va de même dans le fameux exemple du poulet :

(106) Tuez un poulet actif et bien gras. Préparez-*le* pour le four. Coupez-*le* en quatre morceaux et faites-*le* rôtir avec du thym pendant une heure. (*ibid.*)

Brown & Yule plaident pour une approche interprétative « cumulative », c'est-à-dire qui tient compte des changements apportés au référent dans la progression du texte. Charolles & Schnedecker (1993) nuancent cependant ce point en avançant que, contrairement à un SN défini à même de rendre compte du résultat des transformations en vertu de son contenu lexical, le pronom se distingue précisément parce qu'il permet de « faire abstraction » de ces changements et de maintenir la coréférence malgré le contexte évolutif. Pour résoudre le paradoxe posé par le cas des pronoms, les auteurs formulent une contrainte de maintien de l'identité sortale du référent. On doit donc conclure de l'emploi des pronoms ci-dessus « qu'il existe au moins un trait sortal sous lequel ils résistent aux mauvais traitements qu'on leur inflige » (p. 123), en l'occurrence, la contrepartie massive saillante qui en subsiste. Ceci expliquerait pourquoi l'exemple ci-dessous est jugé impossible :

(107) Prenez quatre morceaux de sucre. Faites-les fondre dans de l'eau et portez-*les* à ébullition. (Charolles & Schnedecker 1993 : 123)

Les auteurs invoquent dans ce cas la différence suivante : « la contrepartie massive des morceaux de sucre dissous (à savoir « du sucre ») n'est pas saillante, alors que la « chair » des pommes l'est encore dans la compote ». Charolles (1997) assouplit néanmoins la contrainte « ontologique » du trait sortal en accordant un rôle plus prépondérant à l'emploi même du pronom capable de « forcer » la coréférence au-delà des transformations. Mais l'enjeu demeure, dans cette optique, de déterminer « jusqu'à quel point et à l'aide de quels moyens linguistiques on peut encore prétendre parler de la même chose alors que l'on expose les changements dont elle fait l'objet » (*ibid.* : 91).

Les Fribourgeois prennent clairement leurs distances à l'égard de cette approche des « transformations ontologiques ». C'est non seulement le critère subjectif d'acceptabilité qui est mis en cause, mais aussi et surtout la recherche de contraintes ontologiques – au détriment de facteurs linguistiques – qui consiste à

chercher *dans le réel extralingagier* les principes censés régler les usages linguistiques (ceux-là mêmes dont on cherche à mettre au jour le fonctionnement). Or, sans vouloir refaire la querelle des universaux, il nous paraît que la démarche qui vient d'être décrite, si elle n'est pas scrupuleusement pondérée, aboutit à faire du langage un décalque de la réalité. Et, conséquence plus grave encore, elle conduit presque inévitablement à

substituer à l'investigation linguistique une analyse ontologique ou une physique naïves. (Apothéloz & Béguelin 1995 : 235)

Sans nier l'influence de facteurs « mondains » – autrement dit nos connaissances, notre expérience et notre représentation du monde et ses objets – dans la manière dont les locuteurs catégorisent les référents en discours, les auteurs regrettent toutefois la tendance à simplifier la relation mot-chose et partant, à négliger la prise en compte d'autres facteurs dans le choix des expressions référentielles. A l'appui d'exemples diversifiés, ils montrent notamment que les procédés anaphoriques, lexicaux comme pronominaux, sont susceptibles de tenir compte ou non des modifications subies antérieurement par voie de prédication, selon les objectifs du locuteur<sup>102</sup>. Parmi ces stratégies largement dépendantes des conditions de production, les auteurs signalent des enjeux argumentatifs, sociaux (respect des faces), des effets connotatifs, polyphoniques, la gestion d'ambiguïtés potentielles, etc. sur lesquels nous reviendrons *infra* (§6.3.6).

La prise en compte des changements de point de vue, déjà invoquée pour l'analyse d'autres types de métamorphoses discursives (Schnedecker & Charolles 1993, Achard-Bayle 2001), mériterait par ailleurs d'être exploitée plus généralement, car tout procédé de désignation implique par définition un point de vue particulier. Béguelin (1997b) se penche à cet égard sur l'emploi de marqueurs référentiels dans des contextes d'hétérogénéité énonciative et montre que des phénomènes de non-congruence entre deux catégorisations d'un même référent peuvent parfois être mis au compte d'un surmarquage de la polyphonie :

(108) Mira, elle, ne croit plus en personne ni à rien. L'opposition ? « **Ils** ne sont pas assez forts ! » (Nouveau Quotidien, 09.11.94 < *ibid.*)

Outre la recatégorisation implicite opérée (individu collectif – classe), le pointeur pronominal accentue le contraste entre les deux univers de croyance en jeu (il va de soi que deux locuteurs ne recourent pas forcément aux mêmes dénominations).

Bref, la question des référents évolutifs a tout l'air, dans cette optique, d'un faux problème. Du moins, elle n'est pas un cas particulier mais concerne tous les référents, qui, par définition, sont amenés à évoluer, par les connaissances qu'on en a, au fil de la construction du discours (*cf. infra* §6.3.4). Pour les chercheurs fribourgeois, le problème à creuser n'est ainsi pas celui des transformations subies par les objets réels, mais « celles qui affectent le bagage de connaissances dont disposent, à chaque moment du discours, les interlocuteurs à propos d'un référent donné » (Apothéloz & [Reichler-]Béguelin 1995 : 239-240). Il convient donc d'examiner les moyens mis en œuvre par les usagers pour faire évoluer un objet dans le discours. Dans cette perspective, l'acte référentiel proprement dit s'accompagne de visées stratégiques qu'il convient de mettre au jour (*cf. infra* §6.3.6).

---

<sup>102</sup> Ils évoquent à cet égard les stratégies opposées de désignation des personnages chez Zola vs chez Flaubert, relevées par Corblin (1983) : Zola exploite fréquemment des désignateurs contingents (*le peintre, le commis*), évoquant parfois l'acte éphémère d'une seule prédication (*e.g. le dormeur*), tandis que Flaubert ne tient généralement pas compte des propriétés attribuées aux personnages dans sa façon de les désigner.

### 6.3.2 Discours et texte

Le modèle fribourgeois se voulant une théorie de la combinatoire discursive, il s'agit de comprendre ce que recouvre la notion fondamentale de *discours*. Dans ce cadre, le discours est conçu comme « l'ensemble des matériaux sémiotiques mis en œuvre par les partenaires d'une [...] interaction » (Groupe de Fribourg 2012 : 21) :

Un discours est donc un complexe pluri-codique, qui comprend non seulement des énonciations en langue naturelle, mais aussi des gestes, des actions, des images, des perceptions communes, des savoirs partagés tacites mutuellement manifestes, etc., combinés selon des modes de planification spécifiques.

Quant au texte, il constitue la trace des données graphiques et acoustiques d'une séquence verbale, dont doit bien souvent se contenter le linguiste pour tenter d'étudier « *le discours appréhendé à travers les textes* » (p. 22). Le texte ne constitue ainsi pas une notion fonctionnelle dans le modèle fribourgeois, mais seulement l'instrument fragmentaire à partir duquel le chercheur tâche de faire des hypothèses sur le fonctionnement du discours, qui « supporte[nt] autant que possible sans catastrophe la prise en compte ultérieure ou épisodique des composantes non verbales et environnementales ».

Cette conception se distingue de celles qui assimilent texte et discours<sup>103</sup> ou qui assimilent ce dernier à la « chaîne parlée ». De plus, en considérant le texte comme un simple témoignage de l'événement discursif, elle se distingue aussi d'approches qui le considèrent comme une unité sémantique (e.g. Halliday & Hasan 1976) ou comme un élément constitutif de l'activité discursive<sup>104</sup>.

### 6.3.3 Les articulations du discours

Le modèle fribourgeois s'inspire de la notion d'articulation du langage de Martinet (1967)<sup>105</sup> pour élaborer sa théorie des unités constitutives du discours. Il distingue ainsi différents ordres de combinatoire, se caractérisant par des fonctions, des unités et des règles d'assemblage propres. La première articulation (F1)<sup>106</sup> correspond aux unités dont la fonction est distinctive, à savoir les phonèmes et les syllabes. La seconde articulation (F2) est celle qui concerne les unités à fonction significative, c'est-à-dire morphèmes, syntagmes et *clauses*<sup>107</sup> (« "îlots" de dépendances morpho-syntaxiques », Groupe de Fribourg 2012 : 47).

---

<sup>103</sup> Voir aussi Cornish (2009) sur les risques de cette confusion.

<sup>104</sup> C'est par exemple le cas de Cornish (2009), qui appelle *texte* l'ensemble des traces sémiotiques (verbales et non verbales) qui, mises en œuvre dans un certain contexte, contribuent à créer le discours. La conception fribourgeoise ne s'éloigne pas sur le fond de cette approche de l'activité discursive. Mais la notion de *texte* n'y recouvre pas la même réalité.

<sup>105</sup> Qui ne va toutefois pas au-delà de la « proposition ».

<sup>106</sup> Chez Martinet, la première articulation correspond aux unités significatives et la seconde aux unités distinctives. Le Groupe de Fribourg inverse cet ordre pour ajouter des articulations de rang supérieur dont les unités « incorporent » celles des unités inférieures (p. 27 *sqq.*).

<sup>107</sup> L'introduction de ce terme s'explique par la volonté de se distinguer de la notion scientifiquement problématique de *phrase*.

L'une des originalités du modèle est de postuler une troisième articulation (F3), dont les unités constitutives sont des actions à fonction communicative exécutées sur le mode « ostensif-inférentiel » (Sperber & Wilson 1986). On quitte à partir de ce niveau les aspects traditionnellement dévolus à la morpho-syntaxe. Tout comme les unités significatives « incorporent » des unités distinctives, les unités communicatives peuvent « incorporer » des unités significatives : l'actualisation d'une *clause* mène à son *énonciation*, action qui comprend d'autres composantes, comme des caractéristiques prosodiques et de potentiels gestes ou signes concomitants relevant d'un autre code (« images, iconogrammes, vêtements, odeurs... », *ibid.* : 29). Les énonciations sont ainsi des manifestations mimo-gestuelles, sémiotiquement hétérogènes, exécutées selon les intentions communicatives des participants de l'interaction. Elles se regroupent en programmes discursifs selon une combinatoire propre – la pragma-syntaxe – pour composer des *périodes*, qui se terminent par un mouvement mélodique conclusif. Ce type de contour indique ainsi la fin du programme et partant, une place transitionnelle pour l'interlocuteur.

Il reste à savoir si la période constitue l'unité démarcative d'une quatrième articulation (F4) à fonction interactive (en tant que composante d'un tour de parole, Groupe de Fribourg 2012), ou si elle permet de composer des programmes praxéologiques du niveau de la pragma-syntaxe (F3) mais de plus grande ampleur, c'est-à-dire des « macro-structures narratives, argumentatives, cognitives-pratiques, etc. » (Berrendonner 2016). La question n'est pas tranchée. Quoi qu'il en soit, le fait qu'on puisse considérer soit la composition interne des périodes, soit leur participation à une intervention au niveau interactionnel, ou alors à une macro-structure d'un autre type, permet déjà d'éclairer les différences d'enjeux, visées et objet des domaines de « la pragma-syntaxe » vs de « l'analyse conversationnelle » ou de « l'analyse du discours ».

L'organisation complexe en niveaux d'analyse distincts<sup>108</sup> explique pourquoi le discours ne peut se concevoir, dans cette approche, comme la simple concaténation de segments verbaux ou son résultat compositionnel. Chaque niveau d'analyse présente une logique de structuration propre. En passant de l'un à l'autre,

on change en fait d'articulation, c'est-à-dire d'univers combinatoire. On rencontre d'autres types d'objets sémiotiques, investis d'autres fonctions, répondant à d'autres modes d'assemblage, et dont on a toutes les chances de

---

<sup>108</sup> Cette hypothèse de stratification nette distingue le modèle fribourgeois d'autres modèles théoriques de type « modulaire » comme l'approche du GARS (Blanche-Benveniste *et al.* 1984), où les articulations micro- et macro-syntaxiques se superposent et s'associent dans la construction des énoncés, ou encore comme celui de Roulet (1999 *sqq.*), où les différents modules (lexical, syntaxique, hiérarchique, référentiel et interactionnel) se combinent de manière complémentaire pour constituer le discours. Par ailleurs, l'approche aixoise fournit des outils pour l'étude de la syntaxe uniquement, et non pour l'analyse des constituants du discours. Quant au modèle de Roulet, il porte essentiellement sur l'interaction (très peu sur la syntaxe ou les procédés de référence), dont les unités peuvent être aussi bien monologiques que dialogales, alors que le modèle fribourgeois distingue radicalement les deux ordres. Bien que reprenant la notion de *mémoire discursive*, Roulet ne précise pas son rôle dans les relations interactionnelles, tandis que l'approche fribourgeoise l'exploite à des fins de modélisation des opérations d'inférence.

méconnaître les spécificités, si l'on se contente de les envisager trivialement comme des « segments » mis bout à bout. (Groupe de Fribourg 2012 : 37)

Quand on évoque la question du *contexte* pour un item étudié, autrement dit l'environnement dans lequel celui-ci apparaît, il s'agit de tenir compte du niveau auquel on le considère pour déterminer la nature des facteurs environnants. Si l'on considère uniquement un élément dans sa dimension phono-syntaxique ou morpho-syntaxique, le contexte considéré sera uniquement le matériau linguistique environnant de la combinatoire respective. Par contre, lorsque l'on s'intéresse à la production d'une occurrence, qu'il s'agisse d'une unité distinctive, significative, communicative ou interactionnelle, des facteurs contextuels hétérogènes influent sur sa réalisation. Du reste, ces paramètres contextuels ne sont pas forcément à la disposition du linguiste, qui doit s'en tenir à des hypothèses sur la base des indices à sa disposition. Outre l'environnement linguistique (souvent appelé co-texte) et le discours produit en amont, le contexte lié à la réalisation d'un élément discursif comprend ainsi les dimensions suivantes (Auer 2009, Cornish 2009) : la situation physique à la portée des interlocuteurs, la dimension sociale de l'interaction (rôles sociaux et interactionnels des participants, type d'activité engagée, etc.), le savoir partagé ou encore le genre auquel appartient le discours, etc.

Les chercheurs fribourgeois se penchent principalement sur l'analyse des faits situés à la frontière entre la morpho-syntaxe (ou micro-syntaxe) (F2) et la pragma-syntaxe (ou macro-syntaxe) (F3). À ce dernier niveau, ils s'intéressent en particulier à la description des routines praxéologiques (*i.e.* des schémas d'action) qui composent les périodes, mais laissent (pour le moment, *cf.* Berrendonner 2016) à l'analyse conversationnelle le soin d'étudier la gestion des tours de parole au sein d'un échange et à l'analyse du discours la façon dont se composent les macro-structures de périodes.

#### **6.3.4 Mémoire discursive et nature des objets-de-discours**

Partant du point de vue de Grize (1974, 1982, 1993) sur la visée représentationnelle de l'activité discursive, Berrendonner (1983 : 230-231) nomme *mémoire discursive* (désormais M) l'ensemble des représentations partagées par les participants, que ceux-ci se donnent pour but de faire évoluer dans un échange. Chaque action communicative, en particulier chaque énonciation, permet ainsi de passer d'un état à un autre. Ces transformations ne sont pas exécutées par l'incrémentation de « contenus propositionnels » mais par des opérations d'inférences jugées pertinentes à partir de chaque acte.

La mémoire discursive s'oppose en cela à une conception « sémantique » d'un univers d'interprétation. Elle se situe en effet au niveau du « cognitif-extralinguistique » (Groupe de Fribourg 2012 : 24) sans pour autant s'apparenter à une propriété psychologique de chacun des interlocuteurs, se distinguant par là du concept de *modèle discursif* (*cf. supra* §6.2). M représente « la modélisation d'une réalité interlocutivement neutre : l'ensemble de référents construits dans et par le discours » (Berrendonner 2016). Le fait que les connaissances en M soient communes ne signifie pas qu'elles soient partagées à l'identique chez les interlocuteurs.

M constitue en effet un ensemble « flou » au sens mathématique, c'est-à-dire que la validité des objets qui y figurent se montre graduable. Autrement dit, la validité des objets en M peut se montrer fiable à divers degrés, selon le type d'inférences dont ceux-ci sont le résultat (*cf. infra* Ch.2 §4.1.2 les raisonnements abductifs). Il existe donc une zone de flou permanente en M, dont s'accommodent les interlocuteurs. Mais cela explique également que des divergences puissent surgir sur la teneur précise de M, comme on le verra à maintes reprises.

Les éléments de M s'appellent les *objets-de-discours* et ils ont donc la propriété d'être manifestes et publiquement partagés par les interlocuteurs. En outre, leur nature est insensible au code qui a conduit à leur introduction (code linguistique, iconographique, gestuel, inférentiel, etc.) (*ibid.* : 23) : ils relèvent tous « d'un seul et même type logique naturel » (p. 123), bien que pouvant se présenter sous des formats et modalités variés<sup>109</sup>. La fonction communicative de chaque énonciation consiste à effectuer des opérations sur les *objets-de-discours*, notamment en attribuant à ceux-ci des propriétés par le biais de prédictions ou d'inférences. Ces attributs constituent l'*intension* d'un objet à un état donné de M.

Il convient cependant de ne pas négliger la dimension socio-culturelle qu'un objet-de-discours acquiert au cours de son évolution en M (*cf. Grize 1993*) :

[L'identité des objets] devient le produit d'une interaction entre le sujet humain et son environnement [...], dans la mesure où ces objets ont acquis le statut de construits culturels, et où par conséquent leur « essence » comporte forcément un paramètre anthropologique. (Apothélos & [Reichler-]Béguelin 1995 : 239)

C'est donc l'ensemble de ces aspects qui contribue à faire l'identité d'un objet-de-discours donné en M. On est ainsi loin de la conception « mondaine » traditionnelle des référents. Par ailleurs, la mémoire discursive ne renferme pas seulement la représentation des objets « ordinaires » (qu'ils soient discrets ou non, fictifs ou non, etc.), elle enregistre également toutes les données liées au déroulement de l'échange en cours, autrement dit, des informations de nature méta-communicative sur l'énonciation même (Groupe de Fribourg 2012 : 131 *sqq.*). Par exemple, chaque acte d'énonciation opéré en discours a pour corollaire son inscription automatique en mémoire discursive. En témoigne le fait qu'on puisse y référer de la même façon qu'aux objets du monde :

(109) Oser tuer la formule magique ? **La question** n'a plus rien de rhétorique sous la coupole fédérale en ébullition (presse).

De fait, on peut subdiviser la mémoire discursive en modèle du monde (MM) vs modèle des actions communicatives (MAC), la première se distinguant par la nature extra-linguistique des objets, l'autre par la nature méta-énoncative des informations livrées par le discours même. Cela dit, l'exemple montre que l'opération fondamentale de référence peut se décrire de la

---

<sup>109</sup> Tel que le montrent notamment l'usage de déterminants (*e.g. un vs du*), de pronoms (*il vs ça*), de prédicats (*être nombreux, s'additionner*, etc.). Par exemple *individu vs classe vs continuum, individu spécifique vs type, classe vs collection*, etc. (Berrendonner 2002, 2005).

même manière dans les deux sous-ensembles. Mais un examen attentif des spécificités linguistiques respectives mériterait d'être mené à bien.

Une autre caractéristique de la mémoire discursive est que l'attention accordée à certains objets, autrement dit leur importance cognitive, ou saillance, varie au fil du discours. On peut par exemple imaginer que pour deux interlocuteurs qui sortent d'une séance de cinéma, le film visionné (de même que les divers personnages et étapes clés de l'intrigue) sera particulièrement saillant à ce moment-là dans M. En outre, la multiplication de relations établies en discours entre un référent (via les images, le scénario, les pointages référentiels, les prédications, les inférences qu'on peut en tirer, etc.) (par exemple le personnage principal du film) et d'autres référents, dans un discours, contribuera à en faire un objet central dans l'état courant de M et à ce titre, le promouvra *centre organisateur* d'un *sous-graphe* ou réseau référentiel complexe. De ce point de vue, « les représentations que nous nous formons des objets s'organisent selon un principe plutôt gravitationnel que taxinomique » (Berrendonner 1990a : 155, Apothéloz 1995a : 169-170).

### 6.3.5 Opérations de pointage

Dans l'approche fribourgeoise, il est ainsi évident qu'une expression référentielle ne renvoie pas à un objet de la réalité tangible, mais bien à un objet de la mémoire discursive. Parmi les opérations référentielles, il est utile de distinguer entre une opération d'introduction et une opération de pointage. La première sert à installer un objet-de-discours inédit dans la mémoire discursive via une expression référentielle (typiquement un SN indéfini, cf. *supra* §4.5), tandis que la seconde consiste à désigner un référent supposé y être mutuellement manifeste pour les interlocuteurs. Lorsqu'on parle communément d'« anaphore » ou de « deixis », c'est donc au second type d'opération qu'on fait référence.

Plus précisément, un pointeur référentiel comporte, outre son aspect descriptif, « l'instruction d'unifier sa variable anonyme X avec un objet  $O_i$  valide<sup>110</sup> présent dans M » (Berrendonner 1995 : 242), en s'appuyant sur des indices de diverses natures (lexicale, morphologique, contextuelle, inférentielle, etc.) :

(110) Et là que vois-je ? Un très très petit chaton tout mignon. **Il** est noir et blanc et un petit peu touffu (roman-feuilleton en ligne, <https://www.wattpad.com/91089475-wonderful-holidays-chap-8-rom%C3%A9o-et-juliette>)

(111) Bonjour, je voudrais savoir où je pourrais trouver un vétérinaire pour **cette petite boule de poils** 'demandais-je en désignant Apoplexie du doigt' (*ibid.*, quelques paragraphes plus loin)

Dans l'exemple (110), le pronom conjoint *il* accompagné de son prédicat présuppose l'existence d'un objet correspondant en M auquel le rattacher. Il se trouve qu'un objet

---

<sup>110</sup> La notion de validité ne reposant pas sur un principe de vérité, mais sur la reconnaissance mutuelle de l'existence d'un objet en M.

particulièrement congruent vient d'être introduit en M à travers les propos qui précèdent (*un très très petit chaton tout mignon*). Quant au SN démonstratif de l'exemple (111), il implique, via l'acte de monstration qu'il opère, l'appartenance à M de l'objet ainsi désigné. Pour rappel (*cf. supra*), la manière dont est alimentée M ne détermine pas la nature de ses objets, ceux-ci relevant tous du même type : M a la propriété de « neutraliser les différences de nature entre ses multiples sources » (Berrendonner 1983 : 231). En d'autres termes, les deux références exécutées ci-dessus procèdent fondamentalement de la même façon : il s'agit d'instancier la variable introduite par l'expression en l'unifiant avec un objet présenté comme valide dans M.

Certes, chaque expression implique une procédure interprétative spécifique, dans le premier cas, via le présupposé d'existence d'un objet porteur d'un nom masculin ; dans le second cas, via la nature token-réflexive du SN démonstratif, sans négliger la prise en compte de sa signification lexicale, de son interprétation métonymique, etc. Néanmoins, le fait que tous les objets-de-discours se situent au même titre dans M, *i.e.* l'ensemble évolutif des connaissances partagées, rend caducs les critères traditionnels reposant sur la localisation du référent (*cf. supra* §5.2). Aux yeux des chercheurs fribourgeois, un seul mécanisme permet ainsi de recouvrir le fonctionnement général des procédés référentiels habituellement distingués avec zèle les uns des autres.

Il est fondamental de reconnaître la nature inférentielle des procédés de pointage, à l'instar d'ailleurs de tout phénomène interprétatif (Sperber & Wilson 1986). Dans l'opération en question, cela concerne en particulier le processus de repérage de l'objet pour l'instanciation de la variable introduite par le pointeur, et cela, à partir des divers indices hétérogènes en présence.

Il n'est à cet égard pas rare que l'objet supposé manifeste soit en fait absent de M au moment du pointage et que le recours aux calculs inférentiels soit particulièrement sollicité. C'est par exemple le cas lorsqu'un pointeur renvoie à un référent implicite, dont l'existence même est à inférer à partir d'indices concomitants (*cf.* le cas de l'anaphore associative, présenté *infra* Ch.II §4.1) :

(112) BL: Ah non, et plutôt que, la, à la Maison des saveurs, comme tu dis, j'aime mieux d'aller, au Basilic ici à Chièvremont

E: Ah oui.

BL: Moi.

E: Nous on est allés mais c'était une communion et c'était pas terrible, mais, (tu sais), c'est pas la même chose quand c'est un, un repas pour, beaucoup et <BL: Oui, et maintenant.>

BL: c'est, de nouveau, plus la même chose non plus, parce que **la femme, sa femme** est partie **elle l'a quitté**. (X) c'était un, il paraît que c'était un coureur de jupon, je n'en sais rien c'est leur problème. Mais, elle manque, au bazar, elle n'est plus là. (oral < pfc)

Il faut d'abord remarquer que cet extrait provient d'une conversation spontanée entre proches et comme c'est souvent le cas, nous n'avons accès qu'à une quantité minimale de leur savoir partagé (leurs expériences et connaissances communes), de la situation d'énonciation courante et son environnement physique, des relations que les participants entretiennent, etc.

On peut toutefois remarquer que référence est ici faite à des objets qui, vraisemblablement, ne figurent pas au préalable dans l'état courant de la mémoire discursive. Un lieu, en l'occurrence un restaurant, est récupéré apparemment de la mémoire à long terme par le locuteur BL au moyen d'un nom propre accompagné de coordonnées spatiales précises (*le Basilic ici à Chièvremont*). BL évoque plus loin le départ d'un individu désigné via le SN *la femme*, reformulé *sa femme*, le déterminant possessif *sa* impliquant l'existence d'un tiers à interpréter en relation avec le premier et surtout, à partir des *scénarios* évoqués (Sanford & Garrod 1981) : par exemple le fait que les restaurants se gèrent par des tenanciers, le fait que les couples peuvent passer par des crises, etc. Diverses opérations inférentielles à partir de ces éléments devraient permettre à l'interlocuteur de conclure que les individus désignés sont quelque chose comme un ex-couple de patrons du restaurant. On voit à travers cet exemple que les opérations de pointage induisent des tâches de reconstitution, ou *catalyse*<sup>111</sup>, pour les destinataires et qu'elles font intervenir des processus inférentiels dont la fiabilité, par ailleurs, n'est pas toujours garantie<sup>112</sup>.

### 6.3.6 Stratégies référentielles

On a déjà vu *supra* (§4.6.3) que les opérations référentielles étaient en fait des *poly-opérations*, en ce sens que la procédure de repérage d'un objet s'accompagne d'un acte de dénomination qui est fonction de stratégies référentielles particulières. Le choix pour une expression référentielle repose ainsi sur la conjonction opportune de divers objectifs communicationnels. Le modèle se distingue en cela d'approches qui composent avec des critères d'adéquation au monde, de dépendance au texte, ou encore d'ordre essentiellement cognitif. Les chercheurs fribourgeois se sont penchés sur le fonctionnement de ces stratégies référentielles à l'œuvre au long d'un discours. Sans toutefois en exposer un inventaire exhaustif, ils proposent ainsi certains principes généraux sur les comportements des locuteurs en matière de référence.

En s'inspirant de la notion de *pertinence* de Sperber & Wilson (1986), Berrendonner (1990a) dissocie deux principes régissant les activités locutoires. Le premier correspond à une norme sociale de *coopération* particulièrement visible dans les métadiscours aussi bien institutionnels que particuliers : autrement dit, il incombe au locuteur de prévenir toute difficulté susceptible d'engendrer des coûts interprétatifs importants pour le destinataire. Ce

---

<sup>111</sup> [Reichler-]Béguelin (1995a) et Berrendonner & [Reichler-]Béguelin (1996) empruntent, en l'adaptant, la notion de *catalyse* à Hjelmslev (1968). Elle leur sert à décrire les tâches interprétatives de reconstitution d'éléments manquants dans M, qui ont été « indûment » présupposés par le locuteur via des *coups de force présuppositionnels* (Ducrot 1972 : 51).

<sup>112</sup> Cf. *infra* (Ch.II §4.1) pour les parcours inférentiels possibles.

principe de coopération (appelons-le O1, conformément aux abréviations de Béguelin 1997a *infra*), privilégiant les intérêts de l'interprète, est généralement invoqué comme celui qui régit les comportements des interlocuteurs en conversation (Grice 1975), si bien qu'on a vite fait, si on l'assimile aux règles mêmes du langage, de considérer comme déviant tout comportement qui s'en écarte (Berrendonner 1990a : 151). Néanmoins, la communication ne vise pas la quantité ou l'exactitude référentielle, mais bien, comme le suggèrent Sperber & Wilson (1986), la pertinence informationnelle. Il s'agit dès lors d'adapter la transmission d'information aux objectifs communicationnels à atteindre. Il arrive constamment que les locuteurs produisent des énoncés approximatifs (*looseness*). Sperber & Wilson (1990) donnent à cet égard l'exemple d'une rencontre entre deux individus, nommés Peter et Marie, à San Francisco. L'homme demande à son interlocutrice où elle habite, à quoi celle-ci répond :

(113) I live in Paris. (*ibid.*)

Or, il se trouve que la locutrice habite à Issy-les-Moulineaux, juste en dehors des limites de Paris. Son but n'est manifestement pas de tromper son interlocuteur, mais plutôt de lui fournir une réponse pertinente pour les inférences qu'il pourra en tirer. Sa réponse ne vise pas ainsi un objectif de littéralité (*literalness*) ou de vérité, mais de présomption de pertinence. L'approximation peut ainsi refléter un ajustement à ce principe, comme dans l'exemple ci-dessous :

(114) c'est vrai que + à Paris c'est cher quand même tous les cafés *tout ça* ça revient cher on a pas trop les moyens (oral, cfpp)

L'emploi du syntagme *tout ça*, spécialisé dans la condensation et l'indistinction référentielles (*cf. infra* ch. IV), est ainsi jugé suffisamment pertinent par le locuteur pour atteindre son but. On doit probablement considérer également la mise en œuvre d'une loi du « moindre effort » incitant le locuteur à dépenser le moins d'énergie possible en vue de la réalisation de ses objectifs. On peut donc dégager un second principe contrebalançant le premier, consistant à ménager ses efforts en réduisant ses coûts de production (O2), que Berrendonner (1990a : 150) appelle malicieusement « principe de nonchalance ». L'activité interactionnelle se voit donc guidée par un calcul constant de la pertinence informationnelle résultant d'un compromis effectué entre deux « principes » antagonistes, à savoir d'une part l'optimisation de l'interprétation, d'autre part, l'économie des moyens mis en œuvre<sup>113</sup>.

Béguelin (1997a) se propose de compléter les principes dégagés par Berrendonner (1990a) en examinant le choix de certaines expressions référentielles au détriment d'autres marqueurs possibles. Elle relève d'abord une stratégie consistant à proscrire les répétitions lexicales (O3) pour éviter certains effets de monotonie ou de saturation perceptive. Une autre conduite

---

<sup>113</sup> Le second principe est d'ailleurs souvent négligé par les chercheurs qui s'efforcent de déterminer exclusivement les conditions d'*interprétation* des expressions référentielles, dans la seule perspective de l'allocutaire. Nous nous efforçons d'accorder une place tout aussi importante, dans ce travail, aux circonstances de *production* des pointeurs.

locutoire est celle qui vise à optimiser la mise à jour de M en surdéterminant une expression référentielle par des informations encore inédites (O4). Ainsi, l'exemple ci-dessous manifeste ces deux objectifs, au détriment du principe de coopération (O2) qui réclamerait l'usage de pointeurs morphologiquement congruents :

(115) On y trouve [*i.e.* dans mon sac] aussi bien des oursons broche ou barrette, un requin – mon mari qui est pourtant l'homme le plus doux de la terre en raffole et *les* collectionne –, que des yeux de lion ou de chat en verre que j'achète chez Deyrolle, le célèbre empaillleur de la rue du Bac (presse < Béguelin 1997a : 107)

Dans cet exemple l'emploi pronominal « discordant » au pluriel *les* permet à la fois d'éviter une répétition lexicale (*requin*) susceptible d'être stylistiquement sanctionnée et d'implémenter implicitement dans M la classe, en tant qu'objet inédit, à laquelle l'individu récemment activé ('un requin') appartient.

Enfin, Béguelin repère la manifestation d'un comportement locutoire (O5) tendant au « respect des règles de bienséance et de sauvegarde de la face » (p. 104), qu'elle illustre par l'exemple suivant :

(116) En réponse, « lui » et surtout « elle » (comme on dit en Roumanie pour éviter de prononcer les noms du dictateur et de sa femme)... (presse, < *ibid.*108)

L'usage des pronoms minimalement informatifs (donc *contra* O2) est explicitement justifié par la parenthèse comme stratégie de cryptage et d'évitement par peur, honte, souvenir désagréable ou encore pudeur associés à l'évocation des individus dont il convient, dès lors, d'inférer l'identité.

L'observation de données authentiques met ainsi en lumière une gamme d'objectifs communicationnels avec lesquels jonglent les locuteurs en fonction de leurs besoins dans une interaction donnée. Les emplois répondent donc à des stratégies plus diversifiées que l'optimisation du décodage ou le maintien de la vérité. Le résultat est clairement l'œuvre d'un compromis de la part du locuteur, certains objectifs prenant momentanément le dessus sur d'autres. Bien des cas réputés non standard (car enfreignant O1) s'expliquent par cette gestion des stratégies mises en place par les locuteurs.

## 7. Conclusion

L'objectif de ce chapitre était d'exposer la manière dont sont traitées quelques questions centrales en matière de référence en linguistique, afin d'en dresser un bilan critique. Nous avons ainsi mis en évidence l'héritage logico-philosophique de la sémantique telle qu'elle est aujourd'hui largement pratiquée, mais aussi la prégnance d'autres modèles dans le vaste domaine d'étude des expressions référentielles. En les mettant à l'épreuve des faits linguistiques que l'on rencontre au quotidien, nous avons dès lors pu constater leurs limites à plusieurs égards.

A l'heure où l'accès aux bases de données attestées orales ou écrites est considérablement facilité, ces modèles pourraient tirer profit de la multitude d'observables désormais à disposition. Ils pourraient bénéficier d'ajustements fondamentaux, un peu à la manière dont la syntaxe du français a été renouvelée à l'appui des ressources disponibles – notamment les données orales – depuis plusieurs décennies (*cf.* les modèles du GARS et du Groupe de Fribourg).

Partageant les hypothèses fondamentales du modèle du discours fribourgeois, nous proposons ainsi d'examiner ce que les faits, sans préjuger de leur conformité à une norme préétablie, révèlent sur le fonctionnement général de la référence. Nous nous démarquons ainsi d'approches qui tendent à promouvoir *ce qui devrait être*, en favorisant *ce qui est effectivement attesté*. Par là, nous entendons les faits dans leur diversité, qu'ils soient fréquents ou plus « insolites ».

En inscrivant notre étude dans le cadre fribourgeois, nous pouvons en outre avantageusement rendre compte du fait que les objets dont parlent les locuteurs ne sont pas des *realia* mais des constructions discursives dotées d'une certaine malléabilité. A travers les notions d'*objets-de-discours* et de *mémoire discursive*, le modèle met en avant la nature fondamentalement cognitive, évolutive et construite des représentations communes des interlocuteurs. Dans ce cadre, les opérations référentielles que ces derniers exécutent permettent de tenir compte de facteurs praxéologiques aussi bien que de paramètres liés aux aléas de l'interaction.



## **Chapitre II**

### **Fonctionnement de l'anaphore pronominale**



## 1. Introduction

Dans le premier chapitre, nous avons dressé une synthèse critique des notions en vigueur dans les théories de la référence et des expressions référentielles et en avons soulevé quelques inconvénients. Ensuite, nous avons formulé les avantages d'une approche constructiviste du discours et avons présenté dans le détail le modèle qui sert de cadre à la présente étude. Nous passons à présent à un point considéré comme central dans l'étude des procédés de référence, à savoir le fonctionnement du pronom de 3<sup>e</sup> personne, largement présenté comme un prototype en matière de maintien de la référence. Cela nous permettra d'examiner de plus près la notion d'anaphore déjà esquissée *supra* (§5.1) et de poser les bases de la réflexion sur le pronom de 3<sup>e</sup> personne en vue de l'étude empirique menée au chapitre V sur l'emploi de la forme pluriel *ils* à valeur référentielle sous-déterminée.

En effet, le pronom personnel *clitique* (ou *conjoint*) de 3<sup>e</sup> personne *il* est considéré partout comme le modèle de l'anaphore dont la résolution référentielle opère de manière quasi-automatique. C'est « l'expression anaphorique par excellence » (Kleiber 1990a : 26) : dépourvu de tête lexicale, il indiquerait la continuité par défaut. Avant d'évaluer ce point de vue, nous proposons de commencer par un rappel des propriétés morpho-syntaxiques du pronom clitique de 3<sup>e</sup> personne (§2). Seront ensuite présentées les principales conceptions de l'anaphore pronominale de ces dernières décennies, d'abord les approches « antécédentistes » (§3.1, §3.2), puis l'approche cognitive (§3.3), à propos desquelles nous soulèverons les problèmes en suspens (§3.4). Enfin, nous aborderons le cas de l'anaphore pronominale indirecte (ou associative) (§0), qui se montre réfractaire aux modèles dominants et qui ouvre des pistes intéressantes pour l'analyse des expressions référentielles en général.



## 2. Propriétés morphologiques et sémantiques du pronom clitique de 3<sup>e</sup> personne

Le pronom *il* et ses variantes allomorphiques appartiennent à la catégorie des *pronoms conjoints* (ou *clitiques*) du français, ainsi appelés parce qu'ils figurent toujours dans l'environnement immédiat d'un verbe (seule la négation *ne* ou un autre pronom clitique peuvent les en séparer) et forment avec lui une même unité prosodique. Ils diffèrent en cela des *pronoms disjoints* (cf. *supra* § 4.4) qui, par leur caractère phonologique accentué (on les appelle aussi *toniques*), peuvent constituer une énonciation à eux seuls, par exemple en réponse à une question (117), figurer en tant que sujet d'une clivée ou encore être disloqués (118) :

(117) La grandre prêtresse : Qui va m'escalader cette colline ?

Les Amazones : **Moi ! moi !** (Gracq, J., *Penthésilée*, 1954 < Frantext)

(118) C'est vraiment **lui** qui m'a initié à la musique j'ai envie de dire parce que c'est c'est **lui** qui a commencé euh | \_ | à faire du tambour déjà **lui** à la il avait il avait dix ans je crois | quand il a commencé le tambour la percussion (ofrom)

Si les deux types de pronoms représentent des arguments du verbe, seuls les pronoms disjoints sont admis dans des positions typiques de SN. Comme ils partagent la distribution des SN et qu'ils commutent avec ceux-ci, les pronoms personnels disjoints de la série {*Moi, Toi, Lui, Elle, Nous, Vous, Elles, Eux*<sup>114</sup>}, on peut les considérer comme des *Pro-SN*<sup>115</sup> (plutôt que comme des *pronoms*), au côté, notamment, des pronoms possessifs *le mien, le tien*, etc. ou des démonstratifs *celui-ci/-là, cela*. De leur côté, les clitiques personnels {*je, tu, il, elle, on, nous, vous, ils, elles*}, en dépit de l'étymologie du terme *pronom*, ne constituent pas des substituts syntaxiques (ni d'ailleurs sémantiques, voir *infra* §3) de N ou de SN, malgré l'étiquette qu'on leur attribue régulièrement dans les grammaires scolaires ou dans certains courants générativistes (entre autres De Cat 2007, Laenzlinger 2003, Rizzi 1986a)<sup>116</sup>.

Du point de vue morphologique, certaines formes se déclinent selon leur fonction syntaxique : les clitiques *il/elle/ils/elles* en position de sujet deviennent *le/la/les* en position d'objet direct, *lui/lui/leur* en position d'objet indirect et *en* ou *y* aux régimes obliques. Outre leur capacité à marquer la personne verbale (1<sup>e</sup>, 2<sup>e</sup>, 3<sup>e</sup>, 4<sup>e</sup>...), le genre et le nombre, les clitiques manifestent ainsi une déclinaison casuelle, ainsi que le montre le tableau ci-dessous :

---

<sup>114</sup> La majuscule indique ici le caractère accentué des Pro-SN, permettant, le cas échéant, de les différencier des clitiques.

<sup>115</sup> La notion de *pro-forme* sous-jacente est ici empruntée à l'*Approche pronominale* (Blanche-Benveniste *et al.* 1984, 1990) où elle recouvre le concept de proportionnalité syntaxique faisant ressortir un paradigme de formes et partant, un ensemble de contraintes de réaction. Le terme recouvre donc exclusivement une classe distributionnelle et non un procédé de dépendance à une forme dans le contexte linguistique. Pour une critique de la notion de *pro-forme* comme substitut d'un segment mentionné, voir Cornish & Salazar Orvig (2016). Voir *infra* (§3.1) pour une mise au point à ce sujet.

<sup>116</sup> Pour un bilan de la controverse sur le statut des pronoms sujets, voir Zumwald Küster (2014).

**Tableau 1** : Flexion des pronoms personnels conjoints en français

Rang	Nombre	Sujet	Régime accusatif	Régime datif	Régime locatif	Régime ablatif
1	Singulier	je	me			
2		tu	te			
3		il / elle / ce, ça	le / la	lui		
4	Non marqué	on	nous			
5						
6	Pluriel	ils / elles	les	leur		

On voit que la plupart des personnes ont des déclinaisons défectives ou neutralisent tantôt le paramètre du genre (*e.g. lui, leur*), tantôt celui du nombre (*on, nous, vous*). Il est à noter que *ça* clitique est à distinguer d'un *Ça* tonique (Creissels 1995 : 30). Le premier ne fonctionne qu'en position sujet et commute avec *ce* devant le verbe *être*, tandis que le second commute avec *cela* notamment en position de complément (*manger ça, avec ça, pour ça, etc.*) ou en position détachée (*Ça c'est pas juste*). Les clitiques *ce* et *ça* partagent avec *on* une déclinaison largement défective.

Le comportement morphologique et syntaxique des clitiques a, depuis plusieurs décennies, conduit certains chercheurs à les traiter comme des affixes verbaux (*e.g. Strozer 1976, Rivas 1977, Jaeggli 1982, Miller 1992, Auger 1995, Roberge 1995, Miller & Monachesi 2003, Culbertson, 2010, Berrendonner 1993, 2008, Zumwald Küster 2014*) servant à marquer les différents arguments syntaxiques du verbe. En outre, parmi les clitiques, le fonctionnement des « sujets » se distingue de celui des régimes : comme les suffixes verbaux des langues à « sujet nul », les clitiques sujets du français ont pour caractéristique de marquer l'accord du verbe avec le sujet (Darmesteter 1877). Dès lors, de nombreux linguistes ont suggéré qu'ils étaient en fait (ou en passe de devenir) des préfixes verbaux marqueurs d'accord (Auger 1995, Berrendonner 2008). Par ailleurs, certains étendent ce traitement aux clitiques objets (*e.g. Berrendonner 2008*), tandis que d'autres mettent en évidence les différences de statuts entre clitiques sujets et clitiques objets (*e.g. Auger 1995*). En tout cas, l'hypothèse du statut affixal des clitiques « sujets » représente un parti-pris audacieux au vu des théories dominantes, car elle remet en cause la pertinence du célèbre *pro-drop parameter*<sup>117</sup>, obligeant à « recatégoriser le français comme "langue à sujet nul" » (Berrendonner 2000 : 30).

Quoi qu'il en soit, nous laissons aux syntacticiens la tâche de tirer les conséquences théoriques du comportement et de la distribution des clitiques du français, et nous nous concentrons à présent sur chacun des traits morphologiques véhiculés par les clitiques de 3<sup>e</sup> personne.

<sup>117</sup> C'est-à-dire le critère qui caractérise les langues pouvant se passer d'un pronom sujet réalisé (Chomsky 1981).

## 2.1 La personne

Les pronoms se déclinent en personne afin de marquer les rapports entre les actants impliqués dans le procès exprimé par le verbe. Dans un état de langue ancien, les personnes étaient marquées par les désinences verbales et le verbe représentait un « prédicat-sujet » (Moignet 1981 : 91). En français contemporain au contraire, la personne s'exprime en grande partie « en dehors » du verbe au moyen du pronom sujet conjoint : sauf pour quelques verbes fréquents où la morphologie est particulièrement discriminante quant à la personne (*avoir, être, aller*, etc.), la désinence verbale se montre de nos jours peu discriminante, en particulier à l'oral pour les trois premières personnes<sup>118</sup>. C'est à ce titre que les pronoms conjoints peuvent être perçus comme une forme d'accord avec le verbe, marquant avec la désinence verbale une concordance quant à la personne (*cf.* ci-dessus).

D'après une conception héritée de la grammaire grecque, on reconnaît généralement trois personnes, une série au singulier et une au pluriel. Parmi celles-ci, Benveniste (1966 : 228-230) traite *je* (celui qui parle) et *tu* (celui à qui *je* s'adresse) ensemble, fonctionnant de manière réflexive par rapport à l'énonciation, *i.e.* renvoyant toujours aux mêmes instances énonciatives (c'est leur « unicité spécifique », p. 230), respectivement au locuteur et à l'interlocuteur<sup>119</sup> (*cf. supra* Ch.I §4.4). Il oppose à ces deux premières la « 3<sup>e</sup> personne » *il*, totalement extérieure à la relation instaurée entre *je* et *tu*. Moignet (1981 : 92) parle à cet égard du « délocuté », personne « absente du système de l'interlocution ». Sa référence se montre pour sa part variable et dépourvue d'un trait d'animation. C'est pour cette raison que Benveniste remet en question son statut de *personne* : « la "3<sup>e</sup> personne" n'est pas une "personne" ; c'est même la forme verbale qui a pour fonction d'exprimer la *non-personne* » (1966 : 228). C'est d'ailleurs, rappelle-t-il, à celle-ci que l'on recourt dans les tournures dites impersonnelles.

On déjà a vu *supra* aussi (Ch.I §4.4) que Benveniste ne considère pas la série des personnes dites « plurielles » comme une simple pluralisation des personnes du singulier. En effet, *nous* ne consiste pas en une multiplication de *je*, mais en « une *jonction* entre "je" et le "non-je", quel que soit le contenu de ce "non-je" » (1966 : 233). La personne *nous* représente en quelque sorte un « "je" dilaté au-delà de la personne stricte, à la fois accru et de contours vagues » (p. 235). Cela explique les emplois de *nous* de majesté où *je* est amplifié, rendu plus solennel, ou les emplois de modestie où *je* s'estompe dans une expression plus « diffuse ». De même, *vous* exprime un *tu* amplifié par un ensemble indistinct de personnes, ou par métaphore, utilisé comme une forme de politesse. En ce qui concerne le pluriel de la non-personne *ils*, Benveniste lui reconnaît deux interprétations : un pluriel « régulier », donc multiplicateur ; mais également un fonctionnement diffus où il « exprime la généralité

---

<sup>118</sup> Par exemple, les verbes du 1<sup>er</sup> groupe (infinitif *-er* sauf *aller*) présentent la même forme acoustique pour 4 personnes sur 6 au présent de l'indicatif.

<sup>119</sup> Benveniste (1966) nuance toutefois cette conception, qui correspond à l'emploi ordinaire. La 2<sup>e</sup> personne *tu* (notamment dans ses emplois génériques) se définit plutôt comme « la personne non-je » ou « la personne non-subjective » (p. 232).

indécise du *on* (type *dicunt, they say*) » : « C'est la non-personne même qui, étendue et illimitée par son expression, exprime l'ensemble indéfini des êtres non-personnels » (*ibid* : 235). Pour toutes ces raisons, on se gardera de considérer *vous, nous* et *ils* comme de simples correspondants pluriels des trois premières personnes ; il paraît dès lors plus sage de leur réserver une position ordinale particulière sous le nom de 4<sup>e</sup>, 5<sup>e</sup> et 6<sup>e</sup> personnes.

## 2.2 Le cas

Le marquage casuel des pronoms conjoints reflète la structure argumentale du verbe : chaque verbe impose à son environnement des configurations syntaxiques particulières. Tesnière (1959) parle de *valence* d'un verbe pour désigner les différents arguments qu'il admet. On postule généralement une corrélation entre fonctions syntaxiques et rôles sémantiques (Tesnière 1959, Fillmore 1968, Jackendoff 1990). Ainsi, à chaque argument syntaxique correspondraient certains rôles sémantiques, l'ensemble de la structure argumentale reflétant un schéma actanciel, ou « petit drame » (Tesnière 1959) :

(119) *Il la lui a prêtée.* (Charolles 2002 : 190)

Les différents arguments en (119), à savoir les clitiques sujet, objet direct et objet indirect expriment les rôles respectifs d'*agent* (celui qui fait l'action, à savoir le *prime actant* dans la terminologie de Tesnière), de *patient* ou *objet* (actant qui supporte le procès, ou *second actant*) et de *bénéficiaire* (celui au bénéfice/détriment duquel se fait l'action, ou *tiers actant*).

La liste des rôles sémantiques varie d'une langue à l'autre, mais également au sein d'une même langue et en fonction des cadres d'analyse. On peut par exemple établir un inventaire très détaillé des rôles actantiels, en affinant le plus possible les catégories, ou à l'inverse, prétendre à un grand degré de généralité, en ne notant que des oppositions minimales. Du fait qu'elles reposent sur des critères intuitifs, de telles typologies ne peuvent donc pas échapper à une part d'arbitraire, comme le relèvent Riegel *et al.* (2009 : 237).

A un même argument syntaxique peuvent correspondre plusieurs rôles sémantiques, ce qui se voit en français notamment à travers le cas des diathèses : dans les constructions passives, par définition, l'argument sujet se voit doté d'un rôle de patient, par opposition à sa version active où le sujet représente un agent (*e.g. le gâteau<sub>suj./patient</sub> a été mangé par les enfants<sub>compl./agent</sub>*). Quant aux constructions avec *il* dit *impersonnel*, elles expriment des verbes sans agent (*il pleut*). Sans même envisager les diathèses, la structure actantielle est liée aux constructions potentielles d'un lexème verbal donné :

(120) *le jardinier a reçu un coup dans la figure* (Lazard 1995 : 153)

Le SN *le jardinier* en position sujet dénote ici, conformément aux traits de sélection manifestés par le verbe *recevoir*, un bénéficiaire. L'interprétation de la structure actantielle d'un énoncé est donc d'abord étroitement liée au sens lexical d'un verbe donné. Cette conception est sous-jacente à la notion de *coercition* (*cf. supra* Ch.I §3.2) : pour rappel, une

situation d'*accommodation* ou de *forçage* de nature sémantique sur un élément dont les propriétés sont en contradiction avec les traits sélectionnels d'un autre terme (voir par ex. Lauwers & Willems 2011). Soit l'exemple suivant (ou sa traduction en français) :

(121) I began a book (*ibid.* : 1219) = J'ai commencé un livre

Le verbe *commencer* sélectionne pour son régime un objet de type 'procès', alors que le SN lexical *un livre* dénote en principe un simple ouvrage écrit. La contradiction sémantique se résout ainsi par la réinterprétation du SN *un livre* comme un procès, en l'occurrence l'activité de lecture ou de rédaction d'un livre (*ibid.*). Ce sont les traits sélectionnels du verbe *commencer* qui imposent une telle interprétation sur l'argument objet. On voit donc qu'un schéma actantiel dépend clairement d'un lexème verbal donné.

Dès lors, il vaut mieux éviter de généraliser une forme de correspondance entre arguments syntaxiques et actants sémantiques indépendamment de la prise en compte d'un lexème verbal particulier. On ne dira donc pas que la déclinaison casuelle des pronoms code des rôles sémantiques, mais plutôt qu'elle reflète la structure argumentale propre à un verbe donné – ou mieux, à l'un de ses usages conventionnels. La structure actantielle, pour sa part, reste liée au sens lexical du verbe en question (notamment à ses traits sélectionnels).

### 2.3 *Le nombre*

La plupart des clitiques du français sont en outre variables en genre et en nombre. La variation en nombre se manifeste en français à travers l'opposition morphologique singulier/pluriel. Selon la conception courante, cette variabilité reflète un trait de comptabilité (Lauwers 2014 : 117) : on associe à la marque du 'singulier' un trait d'unicité, tandis qu'on dote le 'pluriel' d'un trait de multiplicité (Corbett 2000 : 4). Dans le cas des noms, Guillet (1978 : 2) rapporte ainsi que « tout substantif N est donc a priori susceptible d'être employé au singulier et au pluriel avec les interprétations respectives un seul N et plusieurs N ». Le nombre serait sélectionné selon que le locuteur projette de référer à un ou une somme d'individus dans son discours. C'est ce qui fait dire à Cornish (1999 : 131) qu'il est à interpréter comme un indice extensionnel, par opposition au genre (voir *infra* §2.4).

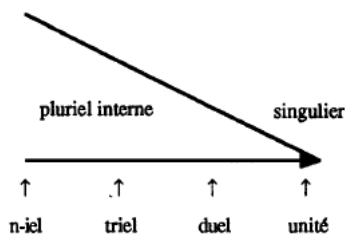
Or, la question du nombre et de ses implications sémantico-référentielles s'avère bien plus complexe. On a vu ci-dessus (§2.1) avec Benveniste (1966) que pour les pronoms de 1<sup>ère</sup> et de 2<sup>e</sup> personnes, les formes dites du « pluriel » s'expliquent mal en termes de multiplicité, au vu des emplois de modestie, de majesté ou de politesse. Le pluriel est à envisager dans ce cas comme « une personne amplifiée et diffuse » et comme un « facteur d'illimitation, non de multiplication » (*ibid.* : 235). Benveniste reconnaît également la possibilité d'une lecture « amplifiée » du pluriel de la « non-personne », à savoir *ils*, dont il rapproche l'interprétation de celui du *on*. Celui-ci se révèle d'ailleurs particulièrement polysémique, apte à représenter aussi bien un individu indéterminé qu'un ensemble de personnes incluant ou non le locuteur (*cf. infra* Ch.V §3.4.2).

D'autre part, on peut relever que certaines langues ne se limitent pas à un système du nombre binaire, mais intègrent davantage de membres de l'opposition à travers les nombres duel, triel, quatriel (voire encore paucal pour certaines langues exotiques, cf. Corbett 2000). Ces catégories manifesteraient selon Guillaume (1964, 1985) une forme de *pluralité interne* :

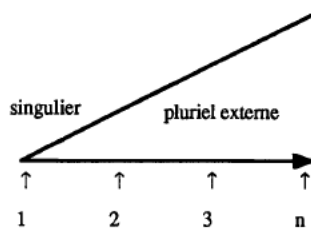
La pluralité interne est celle qui, sous une unité enveloppante, saisit un pluriel contenu, plus ou moins apparent. Le duel des langues anciennes est un vestige de la pluralité interne. Il consiste à saisir « deux » sous un seul regard unique. De là sa convenance avec les choses formant naturellement paire. (Guillaume 1985 : 103)

Le pluriel interne s'oppose ainsi au *pluriel externe*, le seul dont on rend généralement compte, qui pluralise par addition ou multiplication d'unités. Guillaume schématise les concepts de *pluralité interne* vs *pluralité externe* de la manière suivante :

**Figure 6** : Pluralité interne d'après Curat (1988)



**Figure 7** : Pluralité externe d'après Curat (1988)



La figure 7 illustre donc le mouvement de la « pluralité arithmétique », tandis que la figure 6 montre un mouvement de regroupement solidaire « qui aboutit à l'unité » (*ibid.*).

On peut également voir la trace d'une pluralité interne dans le cas de formations de collectifs au neutre pluriel du grec ancien ou du latin, par exemple à travers la fameuse « règle » Τὰ ζῷα τρέχει<sup>120</sup> selon laquelle, de surcroît, le verbe s'accorde au singulier (Colombat 1993 : 30). Ce type d'emploi ne serait pas voué à livrer « l'image d'une série additive, mais plutôt celle d'une collection, d'où le verbe au singulier lorsque le sujet de celui-ci apparaît s'identifier à

<sup>120</sup> Τὰ ζῷα τρέκει = les animaux courent. Le verbe τρέκει est conjugué à la 3<sup>e</sup> personne du singulier, et le sujet Τὰ ζῷα est au neutre pluriel.

une telle collection » (De Carvalho 1993 : 99). On associe ainsi généralement la pluralité interne à une interprétation dite collective et la pluralité externe à une interprétation appelée distributive. Une interprétation se révèle collective si un prédicat s'applique à un ensemble plutôt qu'à ses membres pris séparément. Elle apparaît au contraire distributive si une propriété est attribuée à chacun des individus (Mari 2006) :

(122) Les Suisses qui ont marqué l'année 2014 sont nombreux (titre d'un article de presse, <http://www.illustr.ch/people/National/fil-de-linfo>, 24.12.2014)

(123) Beaucoup de volontaires sont jeunes, autour des 18 ans. (<http://www.lematin.ch/>, 19.09.2014)

Dans (122), le prédicat 'être nombreux' ne peut s'appliquer qu'à l'ensemble, et non à chacun des individus, tandis qu'en (123), le prédicat 'jeune', par ailleurs spécifié après coup ('avoir environ 18 ans'), s'applique à chaque individu, et non au groupe lui-même. Néanmoins, les contextes sont loin d'être tous désambiguïsants :

(124) Suisse : les politiciens chantent pour Noël (<http://www.rts.ch/play/tv/19h30/video/>, 25.12.2014)

Dans ce cas-là, rien dans le titre ne permet de trancher entre une interprétation collective (le chœur ainsi constitué) ou une interprétation distributive (chacun des politiciens a chanté l'un après l'autre), même si, vérification audio-visuelle faite, c'est la seconde lecture que le journaliste avait manifestement à l'esprit.

Le trait collectif ou de pluralité interne peut aussi se manifester lexicalement. En français comme dans d'autres langues, certains noms ont la particularité de s'employer presque exclusivement au pluriel<sup>121</sup>. Pour cette raison, on les appelle *pluralia tantum* ou *pluriels lexicaux* (Lauwers 2014, Lammert 2015, Wierzbicka 1988, Wisniewski 2009) Selon certains auteurs (Jespersen 1924, Furukawa 1977, cités par Lauwers 2014), le pluriel représente dans ces cas-là paradoxalement la marque du non-nombre, ce qui se traduit par la non-dénombrabilité des noms en question.

En fait, Lauwers (2014) remarque que les pluriels lexicaux ne manifestent pas tous le même degré de non-dénombrabilité. Il propose ainsi une typologie de ces N<sup>122</sup> sur la base de leur apparition respective dans un corpus d'écrits du web avec des déterminants de quantification (les numéraux cardinaux, *plusieurs*, *divers*, *différents*) et de jugements d'acceptabilité avec des expressions de réciprocité (*l'un après l'autre*, *les uns après les autres*) et des adjectifs « délimitatifs » ou « stubbornly distributive » (Schwarzschild 2011, cité par Lauwers 2014) comme *grand*, *petit*, *long*. Lauwers relève ainsi, du côté du minimalement dénombrable, les N

---

<sup>121</sup> Ou alors la variation en nombre ne correspond pas à l'opposition entre unité et somme d'unités, cf. : *lunette vs lunettes*, *ciseau vs ciseaux*, etc.

<sup>122</sup> La sélection des N est basée sur les N régulièrement évoqués dans les études antérieures et sur les indices de fréquence opposant la forme au sing. et au pl. de la base *Lexique* de B. New (<http://www.lexique.org/docLexique.php>) (*ibid.* : 120).

lexicaux appelés « compacts » (p. 121), qui se montrent les plus réfractaires à la dénombrabilité. Il s'agit de noms tels que *oreillons*, *alentours* qui apparaissent systématiquement avec des déterminants définis et dont le contenu sémantique se présente comme opaque ou compact. Ensuite, l'auteur distingue les pluriels « denses » (p. 122) comme *arrhes*, *épinards*, *loisirs*, *mœurs*, etc. qui admettent la « quantification imprécise » avec *des* ou *quelques* mais pas la quantification discrète avec *plusieurs*, *différents* ou des numéraux. Ces N représentent des entités sémantiquement homogènes. Ils se démarquent en cela de la catégorie des pluriels « agrégatifs » (p. 123) tels que *vivres*, *viscères*, *frais*, *données excréments*, *déchets*, etc. qui apparaissent plus régulièrement avec des déterminants individuants (*plusieurs*, *divers*, *différents* et les numéraux) et qui manifestent un plus grand degré d'hétérogénéité sémantique entre les individus<sup>123</sup>, bien que conservant un trait [+collectif]. Lauwers mentionne pour finir une catégorie à part, celle des « pluriels internes transposés » (p. 127), qui se laissent volontiers dénombrer par les marqueurs ci-dessus. C'est le cas de *toilettes*, *archives*, *pourparlers*, *représailles*, etc. Cependant, le dénombrement opère au niveau externe, il intervient donc exclusivement au niveau du tout et non de ses membres (« deux toilettes », « trois archives », « plusieurs pourparlers », p. 127-128). Les éléments qui composent le tout demeurent ainsi inaccessibles à toute saisie quantifiée.

On peut mettre les *pluralia tantum* en perspective avec les noms dits collectifs<sup>124</sup>, qui tout en étant au singulier, désignent un ensemble d'individus dont les propriétés communes peuvent être plus ou moins lâches (Lecolle 2013). Ainsi en va-t-il des noms comme *foule*, *constellation*, *public*, *bouquet*, *peuple*, etc. La « pluralité interne » de ces N est particulièrement visible à travers l'apparition d'accords dits « ad sensum », étudiés par Berrendonner & [Reichler-]Béguelin (1995) :

(125) **Le jeune couple très étonné** remercia Mathias et lui **dirent** au revoir. (copie d'élève, < *ibid.* : 37)

Selon Lammert (2010 : 91), ce genre d'accord entre un sujet collectif singulier et un verbe au pluriel est fréquent en ancien français mais disparaît peu à peu par la suite. Elle rapporte un exemple de Pascal en français classique :

(126) Et ainsi **ce peuple**, déçu par l'avènement ignominieux et pauvre du Messie, **ont été** ses plus cruels ennemis. (Pascal, *Les Pensées* < Lammert 2010 : 91)

Néanmoins, Lammert juge cet accord impossible en français moderne. L'attestation d'exemples comme (125) de Berrendonner & Béguelin (1995) met donc à mal cette prédiction. D'un point de vue normatif, d'aucuns soutiendront que ces données proviennent de scripteurs malhabiles. Cependant, des données d'autres genres attestent la persistance de ce type d'accord au pluriel en français contemporain :

<sup>123</sup> Lauwers (2014) distingue en fait trois sous-classes manifestant des degrés d'hétérogénéité référentielle variables.

<sup>124</sup> Voir Lammert (2015) pour un examen de leurs différences sémantiques.

(127) bon après y a des y a des affinités qui se font machin | \_ | mais après c'est c'est c'est un groupe **qui sont** tout le temps tout le temps tout le temps tout le temps ensemble y a pas des d'échappatoire (ofrom)

(128) Jean Dujardin et Alexandra Lamy forment **le couple qui font** le plus rêver les Français (web, légende d'une photo du couple, <http://www.ohmymag.com>)

Lorsqu'un tel phénomène apparaît chez un écrivain, il reçoit le statut de figure de style, en l'occurrence de syllepse du nombre. Ainsi en va-t-il de l'exemple suivant d'Aragon cité dans le *Bon usage* :

(129) Ce couple tenait peu de place dans **leur** coin (Aragon, cité par Grevisse & Goosse < Berrendonner & Béguelin 1995 : 24)

A noter qu'avec le déterminant possessif, l'accord au pluriel dans des écrits « surveillés » semble mieux toléré par la norme :

(130) Le groupe Cartier espère donc que **leurs** collègues et concurrents reviendront sur leur décision l'an prochain. (presse, [Reichler-]Béguelin 1993a)

(131) Il va sans dire que la controverse a fait rage au sein d'une bonne partie de la population norvégienne lorsque le couple princier annonça en décembre dernier **leur** intention de se marier. (presse, Belga, 23.08.2001)

Berrendonner & Béguelin décrivent le procédé à l'œuvre comme l'assimilation d'un tout à la classe de ses membres (*ibid.* : 39), qui représente une *dualité* (*cf. supra* Ch.I §3.3) : pour rappel, un type d'objet à géométrie variable appréhendable de deux points de vue différents, ici au format d'individu collectif et de classe.

Quant à l'anaphore collectif-classe via l'indice de 6<sup>e</sup> personne, elle est en général répertoriée par les linguistes, pour certains selon des conditions strictes (*cf. infra* §4.1.1) :

(132) y avait une soir/ un soir euh | \_ | jeu de nuit donc c'est un groupe qui a dû organiser des un jeu de nuit dans la forêt | \_ | qui faisait assez peur d'ailleurs mais conduit | \_ | hein | \_ | ben conduit organisé alors **ils** ont dû | \_ | organiser donc des jeux de nuit (ofrom)

En somme, au vu des différentes situations présentées dans cette section (indices personnels, pluriels lexicaux, N collectifs), la question du nombre en français ne peut se résumer à une simple opposition entre unité et somme d'unités. La définition traditionnelle de l'opposition en nombre nécessite une réelle remise en question. Parmi les pistes intéressantes qu'on peut évoquer, Lauwers (2014) attribue au pluriel, de par l'analogie entre les pluriels lexicaux et les N massifs, un « effet massifiant ». D'autre part, les notions d'« amplification » ou d'« illimitation » de Benveniste (1966) concernant les indices personnels au pluriel mériteraient davantage d'attention pour d'éventuelles généralisations sur l'interprétation du pluriel des N. Enfin, il n'est peut-être pas inutile de faire un petit détour par les fondements mathématiques de la notion de pluriel, inspirée de la notion d'ensemble. Un ensemble au sens

mathématique représente une collection d'éléments, qu'on peut par exemple représenter par les diagrammes de Venn. Les éléments sont liés à l'ensemble par un rapport d'appartenance. On considère généralement que le pluriel recouvre l'ensemble ou les éléments qui le composent. Comme le cas le plus typique consiste en un ensemble contenant plusieurs éléments, autrement dit dont le cardinal est supérieur à 1, c'est la somme des éléments qu'on retient généralement pour l'interprétation du pluriel. On laisse de côté le fait qu'en mathématique, un ensemble peut contenir un seul élément (un singleton), voire aucun élément (ensemble vide). On pourrait se demander si la langue permet de rendre compte d'un ensemble à un ou zéro élément en langue, par exemple au vu de faits comme celui-ci :

(133) - Les planètes carrées, ça n'existe pas. (Brisson, D., *Gros sur la tomate*, p. 12)

Cet énoncé ne reflète-t-il pas la référence à un ensemble ('les planètes carrées') à propos duquel on nie l'existence des membres qui le composent ?

Un autre phénomène intéressant à envisager de ce point de vue est la notion de personne amplifiée (Benveniste 1966), qui pourrait être regardée comme un ensemble dont le cardinal est égal ou supérieur à 1 (*cf. supra* §2.1, *nous* et *vous* de modestie, de majesté, etc.). Dans tous les cas, une étude empirique sur les marques du nombre en français mériterait de voir le jour pour remettre en cause l'idée répandue que le pluriel correspond à une simple addition d'éléments.

Nous terminons par une brève remarque sur le pluriel à l'oral. Le trait du pluriel peut demeurer non spécifié<sup>125</sup>, si les déterminants, les suffixes, les prédicats ou encore le contexte ne sont pas discriminants. En témoigne cet exemple où Charles de Gaulle, lors d'un discours public, opère une reformulation métalinguistique pour désambiguïser son propos :

(134) « Je m'adresse au(x) peuple(s)... aux peuples au pluriel. » (de Gaulle < Blanche-Benveniste (2000 : 15)

Aucun indice ne permet en effet dans ce cas d'opter pour une interprétation au pluriel. Blanche-Benveniste (*ibid.*) note toutefois que si le discours n'avait pas été préalablement rédigé, les propos auraient probablement été formulés autrement pour garantir une interprétation au pluriel. Quant à la marque du pluriel de l'indice de 6<sup>e</sup> personne, elle n'est audible que par le biais de la liaison stricte [iz] devant un élément vocalique (*e.g.* [izɔ̃]).<sup>126</sup> Sans cela, d'autres éléments morphologiques (désinence verbale, suffixes adjectivaux, etc.),

---

<sup>125</sup> Une bonne partie des substantifs et adjectifs sont homophones au singulier et au pluriel : *maison(s)*, *cheveu(x)*, *bateau(x)*, *grand(s)*, *beau(x)*, *bleu(s)*, etc.

<sup>126</sup> A noter que cette liaison n'est pas stricte en français québécois, où la 6<sup>e</sup> personne peut être prononcée sans liaison devant voyelle, *e.g.* [jɔ̃].

pour autant qu'ils soient distinctifs<sup>127</sup>, lexicaux (prédicats collectifs, réciproques, etc.) ou contextuels peuvent concourir à l'interprétation du nombre.

## 2.4 Le genre

Contrairement au trait du nombre, celui du genre grammatical est en principe inhérent à un nom commun. Le genre des noms est arbitraire, souvent hérité de l'étymon de ceux-ci (Riegel *et al.* 2009 : 329). Ainsi, rien ne motive l'opposition de genre entre *vélo/bicyclette*, *fleuve/rivière*, etc. Néanmoins, pour ce qui concerne les noms d'animés, en particulier d'humains, la question est plus délicate, comme en témoigne le débat toujours actuel autour de la féminisation du lexique qui a émergé dans les années 1970 (Elmiger 2008, 29 *sqq.*). En schématisant un peu la problématique, nous constatons que le genre grammatical ne reflète pas univoquement un trait biologique, au vu des exceptions bien connues comme les noms féminins *recrue*, *estafette* ou *sentinelle*, qui représentent des rôles généralement dévolus au sexe masculin. De même, les noms masculins *mannequin*, *tendron*, *laideron*<sup>128</sup> s'appliquent typiquement à des femmes. A cet égard, il faut distinguer entre le sexe biologique et le genre social ou socioculturel (*gender* en anglais). Ce dernier tire ses fondements des rôles, activités, comportements, expériences, bref des traits typiquement associés à un sexe donné (Elmiger 2008 : 47). On peut encore relever les noms féminins du type *personne*, *vedette*, *connaissance*, etc., qui ne se restreignent pas à la désignation de l'un des deux sexes ou genres socioculturels.

D'un autre côté, sexe naturel ou *gender* ne sont pas totalement sans rapport, aussi indirect soit-il, avec les phénomènes d'accords grammaticaux dans le cas des noms épïcènes (*i.e.* des noms qui n'ont pas de genre spécifié au niveau lexical mais qui, en discours, s'adaptent à des accords aussi bien masculins que féminins) tels que *pianiste*, *élève*, *concierge*. En outre, certains suffixes sont spécifiquement dédiés à l'opération d'une telle distinction (*vendeuse*, *inspectrice*, etc.). Nous verrons plus bas comment résoudre cet apparent paradoxe.

Il faut encore évoquer le phénomène du masculin non marqué, qui concerne aussi bien la référence aux animés qu'aux inanimés : la coordination de deux SN dont les noms sont de genres différents entraîne en principe un accord au masculin pluriel (*le garçon et la fille sont grands / le verre et la tasse sont grands*). Michard (1996) conclut des travaux de Jakobson (1971) sur les *corrélations* entre genres marqué et non marqué<sup>129</sup> que « la signification générale du masculin est donc posée comme première : c'est celle qui s'oppose au féminin

---

<sup>127</sup> Au présent et à l'imparfait, la morphologie verbale des 3<sup>e</sup> et 6<sup>e</sup> personnes sont fréquemment homophones pour les verbes en *-er* (*il mange vs ils mangent, il mangeait vs ils mangeaient*), par opposition au passé composé ou au futur (*il a mangé vs ils ont mangé, il mangera vs ils mangeront*). Quant au pluriel des adjectifs attributs, il est souvent inaudible (*ils semblent contents*).

<sup>128</sup> Mais *laideronne* est également attesté !

<sup>129</sup> « La corrélation est l'opposition d'une catégorie marquée, caractérisée par la présence de A et d'une catégorie non-marquée, caractérisée par le manque de signalisation de A. » (Jakobson 1971, cité par Michard 1996 : 30)

comme genre non-marqué par rapport au genre marqué, et c'est de cette signification générale que dérivent les significations spécifiques » (p. 31)<sup>130</sup>.

A propos des marques du genre dans le cas des pronoms conjoints, on peut admettre que les 1<sup>e</sup> et 2<sup>e</sup> personnes fonctionnent à la manière des noms épicènes animés, dans le sens où ils sont compatibles avec des accords au masculin et au féminin, visibles ailleurs dans le contexte linguistique :

(135) Comme tu es grande, ma petite Louise ! (titre d'un livre pour enfant)

En (135), on observe une redondance des marques féminines au sein de l'énoncé via les différents suffixes et la forme du possessif : elles s'interprètent ici manifestement comme le reflet d'un genre biologique.

Ci-dessous, le pronom clitique pluriel *nous* apparaît de même non spécifié en genre. Seul le suffixe de l'adjectif attribut comporte une marque, en l'occurrence formellement le masculin, dans un usage qu'on peut inférer non-marqué.

(136) Nous sommes loyaux envers nos clients, dont les intérêts sont notre priorité (slogan d'une entreprise)

En effet, il désigne ici un référent indifféremment homme ou femme. Il en va d'ailleurs de même pour le SN *nos clients*.

Le cas des indices de 3<sup>e</sup> (*il, elle*) et 6<sup>e</sup> personne (*ils, elles*) est différent, car ceux-ci sont intrinsèquement marqués en genre. En outre, en tant que « non-personnes », ils s'attachent à désigner aussi bien des animés que des inanimés. De nombreux auteurs se sont efforcés d'identifier le critère qui détermine la marque de genre du pronom de 3<sup>e</sup> personne. Nous expliquerons ci-après en quoi nous considérons en réalité cette question comme un faux problème. Mais il est nécessaire pour cela de comprendre le fondement des hypothèses avancées afin de pouvoir ensuite nous en distancier.

---

<sup>130</sup> Dans le cas de la référence aux animés, ce « masculin à valeur générique » représente l'une des cibles de la critique féministe, du fait que son interprétation nécessite des efforts cognitifs de désambiguïsation pour savoir s'il inclut ou non la référence à des femmes, en particulier dans l'emploi des N lexicaux animés. S'ajoute également au débat une dimension symbolique, selon laquelle la langue est susceptible de refléter les inégalités sociales entre hommes et femmes (Elmiger 2008 : 111-112). Diverses alternatives sont proposées pour éviter l'ambiguïté du masculin générique, parmi lesquelles on peut mentionner les définitions légales dans les textes à caractère juridique du type « le masculin générique est utilisé pour désigner les deux sexes », Chancellerie fédérale 2000 (Elmiger 2008 : 123), l'emploi des doublets intégraux (« les étudiants et étudiantes », *ibid* : 127) ou abrégés (« les traducteurs/-trices », *ibid.* : 121) (jugés fastidieux et peu lisibles par leurs détracteurs) ou les innovations graphiques (« les étudiantEs » ; « illes », *ibid.* : 121). Néanmoins, ces aspects sortent du domaine linguistique et concernent essentiellement les représentations sociales. On peut à cet égard relever la polysémie de la notion de « genre » (grammatical, biologique, sociologique, sans oublier son acception en analyse du discours) et le nombre de disciplines ainsi touchées (linguistique, sociologie, anatomie, psychologie, sciences de la communication, pédagogie, sciences politiques, etc.), conduisant inévitablement à des traitements très divers de la question pouvant donner l'impression d'une confusion notionnelle.

Certains invoquent la présence d'un contrôleur linguistique, autrement dit d'un antécédent responsable de l'accord du clitique. Lorsque le contrôleur explicite fait défaut, on invoque le contrôle par un N *absentee*, à récupérer de manière implicite (Tasmowski-De Ryck & Verluypen 1982, 1985) :

(137) (John essaie de mettre une table dans le coffre de sa voiture ; Mary dit :) Tu n'arriveras jamais à *la* /\**le* faire entrer dans la voiture (Kleiber 1990a : 36)

(138) (Même situation, mais avec un bureau :) Tu n'arriveras jamais à *le* /\**la* faire entrer dans la voiture (*ibid.*)

Selon les partisans du contrôle linguistique, ces exemples montrent que l'adéquation du genre dépend d'un nom, même si celui-ci n'est pas introduit verbalement. Cette thèse s'avère cependant moins convaincante dans le cas où le clitique renvoie à un être humain :

(139) Mon docteur vient d'agrandir son cabinet de réception. *Elle* avait trop de clients (< *ibid.* : 37)

(140) (un automobiliste fonce sur vous :) Mais *il* est fou ! (*ibid.* : 29)

L'hypothèse d'une récupération nominale (*femme* ? pour (139) ; *homme* ? ou *automobiliste* ?) pour (140) montre ici ses limites. Dès lors, d'autres chercheurs proposent d'y voir, à l'inverse, un contrôle *pragmatique*, se passant d'un intermédiaire linguistique (Lasnik 1976, Wiese 1983, Bosch 1986) : le pronom, en l'absence d'un antécédent explicite, tirerait ses marques directement des propriétés référentielles, en vertu d'un accord dit *conceptuel*, *ad sensum* (ou encore par *syllapse*) respectivement au sexe du référent (*cf.* aussi *supra* Ch.I §4.4 et ci-dessus §2.3). Mentionnons d'emblée que cette notion sémantique de l'accord nous paraît tout à fait inappropriée, le phénomène de l'accord étant confiné au domaine des contraintes purement formelles de cooccurrences.

Mais plus généralement, le débat autour de la question du contrôle, que ce dernier soit linguistique ou pragmatique, est la conséquence d'une vision où la langue est subordonnée à des influences extérieures, en l'occurrence, le réel : dans cette perspective, la langue est vouée à reproduire la réalité, elle est déterminée par celle-ci. Ainsi, les partisans de l'un ou l'autre camp cherchent à prouver ce qui, du sexe ou du N du référent, régit l'accord grammatical.

Comme nous l'avons déjà exposé (*supra* Ch.I §6.3.1), notre vision des choses prend le contrepied de cette approche : à nos yeux, il s'agit de partir des usages de la langue et de les décrire, pour observer la manière dont les locuteurs configurent leur représentation du monde, qui n'en est pas une simple reproduction fidèle. Nous ne nions pas toute influence du monde réel sur la manière dont les locuteurs se le représentent; il va de soi que notre perception du monde réel a un rôle prépondérant sur nos représentations en général. Mais il faut prendre en considération le fait que les usagers disposent d'une liberté de configuration des référents discursifs, affranchie des seules contraintes de la réalité. Dans le domaine en question, celui

des marques de genre, nous partons donc du principe que c'est le locuteur, via les marques morphologiques *ad hoc*, qui attribue des traits en fonction de l'interprétation qu'il souhaite véhiculer. Ainsi, ce n'est pas le référent du monde qui détermine univoquement l'emploi du genre, mais c'est l'usager qui indique, à travers la convocation d'un genre, l'appartenance de l'objet à une catégorisation nominale masculine ou féminine.

En outre, une partie des problèmes posés par ces théories pourrait être résolue par une conception alternative du phénomène de l'accord (*cf.* la notion *supra* de *contrôle*), généralement abordé comme une relation orientée, où un terme (source) impose ses marques à l'autre (receveur). Ainsi, dans une approche générativiste-transformationnelle, l'accord observé en surface résulte de l'application en structure profonde d'une règle de copie des traits du premier sur le second (Creissels 2006 : 24 *sqq.*). C'est ainsi que la théorie du contrôle linguistique postule que le pronom tire sa marque de genre d'un nom donné. Or, dans une approche strictement descriptive, on pourrait se contenter de constater que certaines informations à propos d'un référent se voient répétées ou réparties sur plusieurs formes, sans préjuger d'une quelconque dépendance (*ibid.*), *a fortiori* dans le cas des pronoms non régis comme (137), (138), (139) et (140), nullement soumis à des contraintes formelles. Blinkenberg (1950) soulignait déjà l'incongruité de la notion d'accord pour l'usage des pronoms conjoints :

Un pronom personnel peut toujours se rapporter directement au sens qu'il représente, sans l'intervention préalable dans la chaîne parlée d'un substantif. Il ne s'agit plus d'un groupement syntactique constitué par des termes réciproquement dépendants ; au contraire, le pronom personnel est un terme syntactiquement indépendant. (p. 17)

A défaut d'un phénomène d'accord, on peut observer dans l'usage des pronoms celui de redondance des marques morphologiques. Ainsi, dans les cas de référence pronominale, on observe souvent une congruence des marques de genre à travers le ou les pronom(s) (successifs), sans que la présence d'un SN lexical soit obligatoire, comme le montrent (137), (138), (140). De la sorte, le locuteur réitère le genre de l'*attribut de dénomination* (Berrendonner & [Reichler-]Béguelin 1995) du référent sur ces divers éléments linguistiques (les pronoms et, le cas échéant, la réalisation effective de la dénomination). La congruence en genre au fil d'un discours entre les pronoms successifs et, éventuellement, un SN coréférentiel<sup>131</sup>, s'explique ainsi non pas en termes d'accord, mais par un principe pragmatique d'*isonymie* : les sujets parlants « conservent par défaut la même dénomination courante, et donc la même catégorisation, aussi longtemps du moins que ces attributs demeurent distinctifs et qu'il n'y a pas un intérêt stratégique particulier à procéder autrement » (Berrendonner & [Reichler-]Béguelin 1995 : 34). On peut illustrer ce principe par cet exemple :

---

<sup>131</sup> Congruence traitée ailleurs en termes de « chaîne anaphorique » (Chastain 1975, Corblin 1995, *cf. supra* §5.3). Nous renonçons à l'emploi de cette notion du fait de la vision de dépendance qu'elle suppose.

(141) sur les rochers en fait y a des des vers | \_ | qui se s'appellent des vers sapins de Noël | \_ | *ils ils* sont vraiment en forme de sapin de Noël enfin c'est drôle | \_ | un peu en spirale comme ça | \_ | et euh *ils* vivent *ils* vivent sur les cailloux | \_ | et quand tu passes à côté d'*eux* ou quand tu claques des doigts à côté d'*eux* | \_ | euh *ils* rentrent hyper vite dans leur trou | \_ | c'est trop drôle quoi euh tu claques des mains pis bloup | *ils* sont *ils* sont à l'intérieur tu *les* vois plus *ils* disparaissent | \_ | pis de pis si t'attends un moment | \_ | *ils* remontent tout doucement (ofrom)

La « suite » pronominale coréférentielle au masculin pluriel n'est pas grammaticalement régie par l'antécédent en amont (*des vers qui s'appellent des vers sapins de Noël*), mais répète simplement le genre de l'étiquette lexicale (*ver*) associée au référent évoqué. Ce principe pragmatique d'isonymie, reflétant une stratégie d'économie mémorielle, permet d'expliquer avantagusement pourquoi les emplois coréférentiels des pronoms sont prototypiques.

La situation complexe de la référence aux animés mérite qu'on s'y attarde. On reconnaît en général que la catégorisation biologique ou socio-culturelle, communément partagée, est particulièrement prégnante : « c'est un fait bien connu, appuyé par de nombreuses données linguistiques, qu'une des catégorisations premières opérées par l'homme est celle qui discrimine parmi toutes les entités les êtres humains avec une subdivision perceptive directe en hommes et femmes » (Kleiber 1990a : 39). Sans entrer dans des considérations ontologiques, [Reichler-]Béguelin (1993a : 345) remarque une forme de lexicalisation des pronoms (conjoint ou disjoint) de 3<sup>e</sup> personne masculin et féminin, susceptibles de véhiculer « le contenu descriptif de "être humain du sexe masculin" (*il*) ou "être humain du sexe féminin" (*elle*) », à l'appui d'exemples de ce genre :

(142) Si vous optez pour le rasoir électrique, il vaut mieux en acheter un qui sera réservé à votre usage personnel : « *ils* » ont horreur qu'on leur emprunte leur rasoir (*Bien-Etre-Santé*, 07.08.90 < [Reichler-]Béguelin 1993a : 345)

Ici, la marque de genre contribue à construire un référent porteur d'une dénomination masculine, dont on peut inférer, au vu des prédications 'avoir horreur' et 'avoir un rasoir', une identité biologique<sup>132</sup>. Néanmoins, il arrive qu'une situation de conflit survienne entre les différentes interprétations possibles du genre (*cf.* aussi *supra* (139)). On peut invoquer la volonté du locuteur de privilégier tantôt l'inférence d'un genre biologique ou social, tantôt le marquage d'une dénomination spécifique. Ci-dessous, les deux exemples montrent des stratégies inverses : isonymie (143) *vs* changement d'appellation sous-jacente favorisant une interprétation biologique (144) :

(143) Vendredi de la semaine passée, une recrue de l'ER inf 2-12 a soudain été victime d'un arrêt cardiaque à Bière. Malgré les mesures de réanimation qui avaient été prises sur-le-champ, la recrue est décédée aujourd'hui au CHUV à Lausanne où *elle* avait été

<sup>132</sup> Mais rien n'empêche qu'on en infère, selon le contexte, une classe de 'barbiers', 'marchands de rasoirs', etc.

hospitalisée. ([www.news.admin.ch](http://www.news.admin.ch), 13.07.12 ; d'autres médias précisent qu'il s'agit, dans les faits, d'un jeune homme)

- (144) Le mannequin [Naomi Campbell] est célèbre pour son comportement colérique et parfois violent. En 2007, *elle* avait même dû effectuer des travaux d'intérêt général après avoir agressé une femme de chambre. (<http://www.closevmag.fr/>, 04.02.14)

L'attribut de dénomination dont le pronom reflète le genre présente *de facto* les référents correspondants comme des référents *classifiés, nommés* ou encore *catégorisés* (Kleiber 1990a : 39). Lorsque la catégorisation demeure implicite, on suppose généralement qu'elle correspond à un nom de niveau de base de la sémantique du prototype (Rosch *et al.* 1976), qui constitue le niveau utilisé par défaut et présentant une saillance référentielle importante du point de vue perceptif et fonctionnel, comme le montrent de nombreux travaux sur les tâches de dénomination d'objet (Cornish 1999 : 132). Au niveau sémantique, il représente un compromis entre la sous-spécification et la surcharge informationnelle (*cf. infra* 4.2). L'exemple de Cornish (*ibid.*, adapté en français) illustre cette question : en voyant un chien pénétrer dans une maison par la porte d'entrée laissée ouverte, un locuteur sera susceptible de s'écrier *Qu'est-ce que ce chien fiche là ?*, plutôt que *Qu'est-ce que cet animal fiche là ?* (niveau superordonné) ou *Qu'est-ce que ce golden retriever fiche là ?* (niveau subordonné) (*ibid.*). Néanmoins, avant de juger un emploi comme « marqué » sur la base de l'échelle de Rosch *et al.* (*ibid.*), il convient de toujours tenir compte de chaque contexte d'énonciation : c'est avant toutes choses l'état de M à un moment donné du discours qui rend pertinent tel ou tel emploi lexical plus ou moins spécifié, en fonction des coûts cognitifs et des effets recherchés par le locuteur.

Nous terminons cette section sur le genre en évoquant le cas des pronoms conjoints de 3<sup>e</sup> personne *ce, ça*, (parfois aussi *il* « impersonnel »), et les pronoms compléments *le* (référant à des procès), *en* et *y*, souvent qualifiés de *neutres*, dans la mesure où ils neutralisent l'opposition de genre (entre autres Brunot 1922 : 91, Grevisse 2011 : §461, §240, Riegel *et al.* 2009 : 377, Cornish 1991, Carlier 1996, Bartning 2006). A notre avis, cette caractérisation prête à confusion, dans le sens où le terme « neutre » fait croire à un fonctionnement d'un système à trois genres : masculin, féminin, neutre. Or, le français ne connaît pas un tel système<sup>133</sup> : les pronoms en question s'accordent, le cas échéant, exclusivement, au masculin<sup>134</sup> :

- (145) La mer, **c'est beau** (Titre de l'article de Willems 1998)

- (146) J'ai toujours pensé que cela finirait ainsi. Et je **l'ai espéré**. (Dumas, Cl., *L'Herbe chaude*)

---

<sup>133</sup> Par ailleurs, le seul cas où il serait légitime d'invoquer une forme de « neutralisation » du genre serait celui des noms épiciques, où le trait demeure indéterminé au niveau lexical.

<sup>134</sup> Sur l'existence d'un genre neutre en ancien français, voir Marchello-Nizia (1989).

(147) **Il** est **certain** que Constantin aimait le faste. Il développa encore les comptes de la Cour et il entendait que les hauts fonctionnaires l'imitassent. (Lot, F., *La fin du monde antique et le début du Moyen Âge*)

Quoique dans un emploi sémantiquement non marqué, on a donc bien affaire à un accord au masculin, comme en témoignent les éléments avec lesquels les pronoms sont en relation d'implication d'occurrence. Ce qui est plus intéressant, c'est le fait que contrairement aux pronoms *il/elle*, ceux-là sont non-marqués quant au trait d'individuation : ils n'indiquent pas, comme *il* et *elle*, l'appartenance du référent à une catégorie nominale conventionnelle individualisable. De ce point de vue-là, ils demeurent donc non spécifiés, ce qui les rend productifs dans la référence à toutes sortes d'objets aux contours flous et sans dénomination propre (cf. *infra* Ch.III §5.2.2).

Signalons pour finir que certaines variétés connaissent des emplois de la 6<sup>e</sup> personne non-marqués en genre, comme en français québécois où la prononciation en usage ne distingue pas entre *ils* et *elles* (les deux pouvant être prononcés [i] (ou [j] devant voyelle, comme dans *ils/elles ont*, prononcés sans liaison [jɔ̃]). De son côté, Bauche (1920 : 111) mentionne les prononciations dites « populaires » *ma femme il est venu* ou *les vieilles femmes ils sont toujours à causer*. La question d'une neutralisation peut également se poser pour l'exemple SMS suivant<sup>135</sup> :

(148) Non ici **ils** mettent tous des jupes et les mecs ils mettent des pantalons de survêtement (88milSMS<sup>136</sup>).

On peut enfin remarquer que de manière générale à l'oral, la distinction entre les clitiques [i] et [ɛ] (régulièrement prononcés sans [l] final), sans autres indices d'accord, n'est pas toujours évidente à opérer.

## 2.5 *Bilan*

A travers cet exposé relatif aux marques morphologiques des pronoms conjoints, notre objectif était de mettre en évidence la complexité et la subtilité de leur interprétation, en dépit de l'apparente simplicité avec laquelle ils sont généralement traités. En ce qui concerne le pronom référentiel<sup>137</sup> *il* (et, dans une certaine mesure, de sa « variante » au pluriel, la 6<sup>e</sup> personne) on peut retenir les points suivants :

---

<sup>135</sup> En l'absence de contexte supplémentaire (e.g. le message auquel celui-ci réagit), on ne peut toutefois prouver que *ils* renvoient ici à des individus de genre biologique/social féminin.

<sup>136</sup> Panckhurst R., Détrie C., Lopez C., Moïse C., Roche M., Verine B. (2014) : <http://88milsms.humanum.fr/index.html>.

<sup>137</sup> Nous excluons par là évidemment le *il* « impersonnel », qui ne commute avec rien. Dans les constructions dites « impersonnelles », « la position d'argument sujet n'existe pas. Elle n'est pas simplement vacante, mais impossible à instancier » (Berrendonner 2000 : 50), d'où l'explication de la vacuité référentielle du *il*. Ce dernier a tout l'air d'y fonctionner comme un marqueur de diathèse *asubjectale*. On ne peut dès lors l'intégrer à la classe des pronoms conjoints.

- i) Le clitique *il* signale une non-personne (au sens de Benveniste), un « délocuté » : il représente un actant individué (par opposition à d'autres indices de 3<sup>e</sup> personne) distinct des personnes du discours (*je, tu, nous, vous*).
- ii) Son cas (ou régime) reflète la structure argumentale d'un verbe donné, à mettre prudemment en rapport avec la structure actantielle du lexème verbal au niveau sémantique.
- iii) La variation en nombre ne signale pas tant une opposition purement quantitative qu'une différence entre unité et ensemble, l'ensemble pouvant être appréhendé de différentes manières.
- iv) La marque de genre reflète le genre de l'attribut nominal du référent. Le masculin sert de genre non-marqué.

Une caractéristique évidente et définitoire des indices personnels est leur nature non lexicale. A ce titre, outre les caractéristiques mentionnées, les indices personnels sont pauvres en indications « descriptives ». C'est la raison pour laquelle on les traite généralement comme des formes sous-spécifiées par rapport aux expressions lexicales. Nous proposons à présent d'observer le fonctionnement du pronom conjoint de 3<sup>e</sup> personne au-delà des marques morphologiques, autrement dit dans sa dimension anaphorique, unanimement mise en avant dans la littérature sur la question.

### 3. Approches de l'anaphore pronominale

Étant donné la prégnance de l'approche « antécédentiste » de l'anaphore (*cf. supra* §5.3), le pronom de 3<sup>e</sup> personne n'échappe pas à une telle conception, appréhendée dans sa version forte en termes de *substitution*, comme en témoigne l'extrait du *Bon usage* ci-dessous :

Les pronoms sont des représentants (ou des *substituts*) quand ils reprennent un terme se trouvant dans le contexte, ordinairement avant, parfois après. Ce terme est appelé **antécédent**. [...] *Vous demandiez les journaux d'aujourd'hui ; je vous les apporte.* (Grevisse & Goosse 2011 : §650)

La conception substitutive reflète fidèlement l'étymologie du mot (*pronomem* en latin, traduction du grec *antonumia*) et apparaît dans les traités de grammaire les plus anciens (Denys le Thrace, II<sup>e</sup> s. avant J.-C., Apollonius Dyscole, II<sup>e</sup> s. après J.-C.), en passant par la *Logique* de Port-Royal : « L'usage des pronoms est de tenir la place des noms, et de donner moyen d'en éviter la répétition, qui est ennuyeuse » (Arnauld & Nicole 1662=1874 : 109). Nous allons voir à présent quelles sont les théories linguistiques sous-jacentes à la notion de substitution (§3.1). Cette conception sera toutefois abandonnée en linguistique (alors qu'elle perdure en grammaire scolaire), au profit d'une approche plus fonctionnelle, d'abord cantonnée à une perspective textuelle (§3.2) puis réaménagée, à des fins de généralité, dans une optique cognitive (§3.3). Enfin, nous tirons un bilan de la conception en vigueur du fonctionnement du pronom.

#### 3.1 Conception substitutive

C'est l'approche en vigueur dans les grammaires classiques ou contemporaines, où la vocation substitutive est également appelée *suppléance*, *remplacement* ou encore *représentation* (*e.g.* Arrivé *et al* 1986 : 63, Gardes-Tamine 1998 : 147, Grevisse 2011 : §220 & §650). Si l'on peut admettre qu'en ancien français, les pronoms personnels sujets fonctionnaient à la manière de SN, on ne peut pas en dire autant aujourd'hui au vu de leur distribution distincte, les apparentant comme on l'a vu *supra* (§1) à des affixes de la flexion verbale : là où une seule série de pronoms était en usage autrefois, on observe de nos jours deux classes distinctes de pronoms (Zumwald Küster 2014), la classe atone des pronoms clitiques (ou conjoints)<sup>138</sup> et la classe tonique des pronoms disjoints, cette dernière seulement occupant les positions de SN.

La conception supplétive des pronoms est étayée par les travaux des plus grands théoriciens de la linguistique, qui font de la substitution un concept clé de leur modèle, quoique

---

<sup>138</sup> Pour certains auteurs, cette évolution est le résultat d'un processus de grammaticalisation (notamment Givón 1976, Lambrecht 1981, Miller et Monachesi 2003, Culberston 2010). La notion de grammaticalisation est toutefois à manipuler avec précaution, comme le souligne Zumwald Küster (2014) du fait qu'elle met en jeu, dans certaines conceptions, des scénarios diachroniques rigides impliquant un certain nombre d'étapes et une évolution à sens unique. Pour une remise en question de la notion, voir Béguelin, Corminboeuf & Johnsen (2014).

diversement récupéré par les grammaires<sup>139</sup>. Ainsi, Bloomfield (1933) part de la définition suivante :

A substitute is a linguistic form or grammatical feature which, under certain conventional circumstances, replaces any one of a class of linguistic forms. (Bloomfield 1933 : 247)

Parmi les différents substituts<sup>140</sup>, Bloomfield appelle *anaphoriques* ceux qui remplacent une expression déjà mentionnée (l'antécédent), comme ci-dessous :

Thus, when we say *Ask that policeman, and he will tell you*, the substitute *he* means, among other things, that the singular male substantive expression which is replaced by *he*, has been recently uttered. (*ibid.* : 249)

L'auteur prend soin de préciser qu'un substitut ne désigne pas directement un référent du monde, tel que les expressions régulières, mais qu'il évoque une classe de formes linguistiques :

[...] substitutes are one step farther removed than ordinary forms from practical reality, since they designate not real objects, but grammatical form classes ; substitutes are, so to speak, linguistic forms of the second degree. (1933 : 250)

Lyons (1977 : 659) et Bosch (1983 : 15-16) voient dans cette approche l'origine du traitement initial des pronoms en grammaire transformationnelle : l'existence implicitement admise par Bloomfield de deux degrés de la langue se voit alors récupérée en termes de structures *superficielles vs profondes* dans le modèle transformationnel de Chomsky (1965). A partir de ce modèle, Langacker (1969) élabore la règle de *pronominalisation* qui permet de remplacer un SN en structure profonde par un pronom en surface si elle maintient une identité référentielle (coréférence) ou lexicale entre les deux (notée à travers des indices référentiels identiques). Ainsi, la phrase *Pierre pense qu'il est immortel* (Zribi-Hertz 1996: 37), où le sujet de la subordonnée représenterait Pierre, est dite dérivée d'une structure profonde de ce type :

(149) Pierre<sub>i</sub> pense que Pierre<sub>i</sub> est immortel. (*ibid.*)

Néanmoins, des difficultés de marquage de la coréférence des termes en structure profonde conduisent rapidement les générativistes à abandonner cette première analyse et à considérer que les pronoms sont générés de manière directe (Lyons 1977 : 663). L'intérêt se porte alors sur l'identification des contraintes syntaxiques et sémantiques de coréférence entre antécédent et pronom, à travers l'élaboration des concepts de *c-commande*, de *liage* et *gouvernement*

---

<sup>139</sup> Cf. aussi Bühler (1934=2009 : 563) à propos de l'anaphore : « Si elles pouvaient parler, toutes les flèches anaphoriques s'exprimeraient à peu près ainsi : regarde devant ou derrière toi le long de la chaîne du discours actuel ! Il y a là *quelque chose qui a en fait sa place par ici*, à l'endroit où je me trouve, de façon à pouvoir être relié avec ce qui suit » (nous soulignons).

<sup>140</sup> A titre d'exemples, Bloomfield intègre dans cette classe les pronoms personnels anglais, les adverbes spatio-temporels (*here, there, now, etc.*), des quantifieurs tels que *all, some*, les pronoms démonstratifs *this, that*, les nombres cardinaux *one, two, three*, des pro-verbes (*do, be, have, will, must, ...*), les anaphoriques zéros, etc.

(Reinhart 1976, Chomsky 1981). Nous n'entrerons pas dans les détails de ces notions<sup>141</sup>, étant donné que la grammaire générative, dont le champ d'observation est la phrase isolée, ne capte qu'une infime partie des emplois des pronoms : ceux, syntaxiquement contraints, qui possèdent un antécédent dans le même énoncé. En effet, seuls les cas d'anaphore dite liée, c'est-à-dire soumis à la contrainte syntaxique de *c-commande*<sup>142</sup>, concernent le domaine en question :

En distinguant l'anaphore libre de l'anaphore liée, Reinhart fait de la *c-commande* la ligne de démarcation entre l'anaphore discursive et l'anaphore syntaxique, qui seule concerne la théorie grammaticale. Parce qu'elle ignore la *c-commande*, l'anaphore libre peut unir un anaphorique à un antécédent situé dans une autre phrase. Limitée par la *c-commande*, l'anaphore liée n'est par contre *a fortiori* possible que dans les limites d'une phrase (Zribi-Hertz 1996 : 92).

En bref, les limites d'une conception substitutive du fonctionnement pronominal se situent à notre avis à deux niveaux : i) d'une part, dans l'hypothèse, critiquée *supra* (Ch.I §5.4), selon laquelle le pronom remplace une expression préalablement mentionnée, la relation entre un « antécédent » (s'il en y a un) et un anaphorique n'étant pas, comme déjà vu, une mécanique réglée ; ii) d'autre part – et ce problème concerne spécifiquement le français – dans l'idée erronée que les pronoms conjoints partagent la distribution des SN (*cf. supra* §1). Pour désigner un ensemble de formes grammaticales (*e.g.* les pronoms disjoints) syntaxiquement « proportionnel » à un autre, c'est-à-dire qui partage la même distribution (*e.g.* les SN), nous éviterons ainsi de parler de « substitution » pour la première raison invoquée, mais nous emprunterons la notion de *pro-forme* à l'approche pronominale du GARS (Blanche-Benveniste *et al.* 1984, 1990), qui permet de révéler, en micro-syntaxe, l'existence de paradigmes d'éléments équivalents et de relations de rection.

Au demeurant, si les pronoms conjoints ont le statut d'affixes, on peut se demander s'ils sont capables d'assumer un procédé référentiel, qu'on associe typiquement aux expressions nominales. Autrement dit, peut-on admettre qu'un simple affixe supporte un procédé anaphorique ? En fait, selon nombre de linguistes (entre autres Barlow 1988, Corbett 2006, Croft 2013, Siewierska 2004), il n'y pas lieu de faire une distinction entre marques de personne et expressions anaphoriques :

Most scholars working on agreement acknowledge that there is no good basis for differentiating between person agreement markers and anaphoric pronouns. (Siewierska 2004 : 121)

Siewierska montre que dans certaines langues, le même affixe (ici le préfixe *-i*) est utilisé tantôt comme marque d'accord du verbe, tantôt comme seul vecteur de valeur référentielle :

---

<sup>141</sup> Par la suite, certains représentants de la grammaire générative ont avancé que les pronoms s'assimilient non pas à des SN, mais à des déterminants (pour un aperçu de la question, voir Corblin 1995 : 27).

<sup>142</sup> Définition de base selon Reinhart (1976) : « Un nœud A est dit *c-commander* un nœud B si (1) A ne domine pas B, et inversement ; (2) le premier nœud à ramifications qui domine A domine également B. (Zribi-Hertz 1996 : 56)

- (150) Gumawana<sup>143</sup>  
 Kalitoni i-paisewa  
 Kalitoni 3SG-work  
 ‘Kalitoni worked’  
  
 I-situ vada sinae-na  
 3 SG -enter house inside-3 SG (INAL)  
 ‘He entered the inside of the house.’ (*ibid.*: 122)

En fait, on pourrait faire la même analyse en français :

- (151) Kalitoni *il* a travaillé  
*Il* est entré à l’intérieur de la maison.

Cet exemple montre que le français connaît le même phénomène, l’indice personnel pouvant tantôt être redondant sur le sujet et surmarquer la 3<sup>e</sup> personne, tantôt endosser seul le rôle référentiel.

Cette mise au point nous permet, au final, de rejeter la conception substitutive de l’indice de 3<sup>e</sup> personne en français – à considérer syntaxiquement comme un affixe verbal – sans que cela remette en cause sa capacité à assumer un fonctionnement référentiel.

### 3.2 *Conception textualiste*

On a vu *supra* (Ch.I §5.3) le rôle qu’a joué l’ouvrage d’Halliday & Hasan (1976) dans la conception textualiste de l’anaphore. Les auteurs distinguent explicitement le processus de substitution de la référence :

Substitution is a relation between linguistic items, such as words or phrases ; whereas reference is a relation between meanings. In terms of the linguistic system, reference is a relation on the semantic level, whereas substitution is a relation on the lexicogrammatical level, the level of grammar and vocabulary, or linguistic ‘form’<sup>144</sup>. (p. 89)

Situant le pronom de 3<sup>e</sup> personne parmi les procédés de référence, ils considèrent son fonctionnement comme anaphorique (au sens textuel) par défaut et soulignent sa fonction particulièrement cohésive dans les textes<sup>145</sup>, où les occurrences pronominales sont susceptibles de s’accumuler pour renvoyer à un même référent :

<sup>143</sup> Langue austronésienne parlée en Papouasie-Nouvelle-Guinée.

<sup>144</sup> Même si certaines occurrences peuvent s’interpréter des deux manières, relevant d’un phénomène à la fois grammatical et sémantique (*ibid.* : 88), les auteurs situent par exemple l’emploi de *one*, *do* et *so* du côté de la substitution (respectivement nominale, verbale et clausale) et le cas des pronoms personnels, démonstratifs et formes de comparaison du côté de la référence. Pour ce qu’ils appellent substitution, on peut penser en français à certains emplois de *en*, *celui/celle* « mentionnels » ou *faire* parfois traités en termes d’anaphore « lexicale », c’est-à-dire qu’ils n’interviendraient qu’au niveau de la dénomination (*e.g.* Riegel *et al.* 2009 : 1031) (et non de la référence). Ce genre de phénomène nécessiterait toutefois un examen plus approfondi.

<sup>145</sup> Pour le français, Corblin (1995 : 194) note également la fonction « maximale cohésive » du pronom de 3<sup>e</sup> personne et sa capacité à « monopoliser les chaînes de référence ».

One occurrence of *John* at the beginning of a text may be followed by an indefinitely large number of occurrences of *he*, *him* or *his* all to be interpreted by reference to the original *John*. This phenomenon contributes very markedly to the internal cohesion of a text, since it creates a kind of network of lines of reference, each occurrence being linked to all its predecessors up to and including the initial reference. (p. 52)

Or, c'est précisément cette vision de connexité segmentale opérant à reculons qui a été critiquée *supra* (Ch.I §5.4). Les objections, on s'en souvient, sont l'in vraisemblance d'un processus aussi réglé et analytique dans des genres de parole moins planifiés, la difficulté, dans bien des cas, de cerner un segment délimitable, ou encore le traitement à part du fonctionnement « exophorique » des formes concernées. Malgré les limites et le défaut de généralité de la conception textualiste, celle-ci reste sous-jacente aux recherches appliquées, par exemple en TAL ou en psycholinguistique, focalisées sur le processus de résolution référentielle reposant sur l'identification d'un antécédent, susceptible de causer des « dégradations de performance du système » ou des « coûts de traitement ».

### 3.3 Conception cognitive

Dans la foulée des travaux cognitivistes sur l'anaphore et les expressions référentielles (*supra* §4.6, §3.3), des hypothèses censées parer aux inconvénients soulignés ont été formulées à propos du pronom de 3<sup>e</sup> personne.

Ainsi, pour éviter un « traitement éclaté » liée à l'opposition traditionnelle anaphore-deixis (*cf. supra* §5.2), Kleiber (1994b) propose de voir dans l'occurrence pronominale ci-dessous la manifestation d'une forme de continuité thématique :

(152) (en voyant une connaissance passer :) Je ne *l'*ai pas vu depuis des mois (p. 103, traduit de Bosch 1983)

A ses yeux, il s'agit d'un prolongement non seulement de la référence, mais aussi de la situation « manifeste ou saillante » (*ibid.* : 82) où le référent doit jouer « le rôle d'un argument » (*ibid.*). Kleiber propose la glose suivante de la situation qui fait l'objet du prolongement (en italiques) :

(153) *C'est Paul qui passe / Voici Paul.* Je ne *l'*ai pas vu depuis des mois. (*ibid.* : 84)

Il ajoute que ce n'est pas seulement la perception de l'individu en question qui rend celui-ci ou la situation manifestes, mais surtout le fait que les interlocuteurs connaissent préalablement l'individu, et il résume son explication de la sorte :

« [L]e pronom *il* dans cette nouvelle optique reste, en somme, un marqueur de continuité thématique, mais non pas d'une continuité thématique simplement nominale : il indique que le référent est saisi en continuité avec la situation manifeste » (p. 123).

Nous avouons, comme Demol (2010 : 51-52), une certaine gêne à l'égard des notions employées, en particulier à l'égard des notions intuitives de *situation manifeste ou saillante*,

de *continuité thématique* ou de *de rôle d'argument*. La première est expliquée comme renvoyant à une situation « disponible ou présente dans le focus d'attention de l'interlocuteur ». Cependant, la notion reste vague et fait de surcroît appel à une autre notion problématique, celle de *focus*. Il en va de même pour l'expression *continuité thématique*, renvoyant à la notion de *thème*, habituellement opposée à celle de *focus* (cf. discussion *infra* §3.4). Quant à la notion d'*argument*, issue de la tradition logique, elle n'est pas non plus explicitée. On ne sait pas si elle est ici à fondement syntaxique (cf. la valence verbale), si elle repose seulement sur une sémantique intuitive ou si elle mélange les deux niveaux, avec tous les problèmes que cela suppose. D'ailleurs, Kleiber (1994b) admet le caractère « provisoire » de sa description et les questions ouvertes qu'elle laisse : « la question : quand est-ce qu'un locuteur peut présumer qu'un référent est manifeste pour son interlocuteur ? reste ouverte » (p. 130).

L'auteur souligne également le rôle joué par la phrase-hôte dans la détermination du « bon » référent :

(154) Fred enleva son manteau. *Il* était fatigué. (p. 87)

(155) Fred enleva son manteau. *Il* était sale. (*ibid.*)

La prédication opérée par la phrase-hôte, ou *segment indexical* (Cornish 1999 : 70) se révèle donc cruciale pour l'interprétation du pronom (Gundel *et al.* 2000, Yule 1982), *a fortiori* en cas d'absence d'antécédent :

(156) Attention ! Tu vas *le* casser (Cornish 1999 : 77)

Le prédicat '*casser*' ne sélectionne pas n'importe quel type d'objet, mais un actant qui a précisément la propriété de se casser et doté d'une classification usuelle au masculin.

Cornish propose une vision similaire du fonctionnement du pronom de 3<sup>e</sup> personne, se traduisant par le maintien de l'attention portée sur le référent en question :

As far as ordinary third-person pronouns are concerned, their chief discourse function is to signal referential and attentional continuity, thereby marking the stability of a given referent's existence within a given discourse model. This is the case where they are unaccented in English, and clitic in French » (*ibid.* : 63)

On se souvient également (cf. *supra* Ch.I §4.6.1) que pour Ariel (1988, 1990), les pronoms marquent une haute accessibilité, autrement dit, ils sont considérés comme spécialisés dans la référence à des entités très accessibles en mémoire : cette accessibilité ne reposant pas seulement sur des critères textuels, mais aussi sur le nombre de candidats potentiels et sur la « topicalité » (ou « saillance ») du référent.

Quant au modèle de Gundel *et al.* (1993 : 279) (cf. *supra* Ch.I §4.6.2), il requiert un statut *en focus* pour l'emploi des pronoms non accentués, c'est-à-dire que le référent visé doit être situé

dans le centre d'attention des interlocuteurs (issu de la mémoire à court terme). Les objets en « focus » sont à leurs yeux les plus susceptibles de constituer le « topique » des énoncés subséquents<sup>146</sup>.

On pourrait multiplier les témoignages en ce sens à l'égard du fonctionnement du pronom de 3<sup>e</sup> personne<sup>147</sup>, reconnu de la sorte comme le marqueur typique de la continuité, qualifiée tantôt de thématique, tantôt de topicale, ou encore d'attentionnelle, etc. Avant de mettre en évidence des faits qui s'y montrent réfractaires (§4.2), nous présentons ci-dessous quelques réserves à l'égard des notions invoquées pour justifier celles que nous retenons pour la suite de ce travail.

### 3.4 *Bilan*

Il nous paraît à ce stade utile de formuler quelques remarques et mises au point sur des problèmes d'ordre notionnel. En effet, la plupart des théories invoquent des termes dont l'extension demeure passablement floue. C'est notamment le cas des notions de *focus/topique* (et leurs dérivés).

On peut d'abord constater que la notion de *focus* est exploitée dans plusieurs domaines : dans un sens psychologique, le *focus* constitue le centre ou foyer d'attention des interlocuteurs (Sanford & Garrod 1981, Gundel *et al.* 1993, Cornish 1999). Il représente le point de repère en fonction duquel les participants gèrent le flux d'informations échangées au cours du discours, afin de s'assurer d'être « sur la même longueur d'ondes ». Dans une perspective cognitive, l'anaphore et la deixis constituent précisément des moyens de gérer le *focus*, en y plaçant ou en y maintenant des référents. Selon Cornish (1999 : 25), le *focus* d'attention est réactualisé à chaque nouvelle prédication par les interlocuteurs, ceux-ci étant amenés à déterminer ce qui fait l'objet de chacune d'entre elles – et partant ce qui y figure. Sidner (1981: 220) adopte une vision similaire à l'égard de la notion de *focus* : « [t]he items in focus are those that are talked about for a part of the discourse ». On ne peut s'empêcher de remarquer ici des chevauchements avec la notion de *topique* ou de *thème*, envisagée par Lambrecht (1994) en termes d'*aboutness* ou d'*à-propos*, concept inspiré par Strawson (1964 : 97) : « we intend in general to give or add information *about what is a matter of standing current interest or concern* » (nous soulignons). Le *topique* est en ce sens « the matter of current interest which a statement is about » (Lambrecht 1994 : 119). Les notions de *focus* et de *topique* se recouvrent donc chez certains auteurs : « topics are referentially given in the

---

<sup>146</sup> La notion de *focus* est visiblement utilisée dans un sens psychologique et non en termes de structure informationnelle, où *topique* et *focus* en principe s'opposent. En fait, les deux notions de *focus* ne seraient apparemment pas sans rapport : « elements tend to be linguistically focussed because the speaker wants to bring them into the focus of attention » (Gundel *et al.* 1993 : 279). Le terme *topique* semble quant à lui envisagé ici en terme d'*aboutness* au rang de la phrase : « what the speaker intends a sentence to be primarily about » (*ibid.*).

<sup>147</sup> Cf. Lambrecht (1994 : 132) : « in coherent discourse the overwhelming majority of subjects are unaccented pronouns, *i.e.* expressions which indicate topic continuity across sentences ». Voir aussi l'état de la question de Demol (2010 : 48-60) sur le marquage de la continuité propre au pronom *il*. D'autres travaux vont dans le même sens : entre autres Apothéloz (1995a : 280), Charolles (1997), Givón (1983), Schnedecker (1997).

sense of being in the current focus of attention » (Gundel & Fretheim 2004 : 180). Or, dans le domaine de la structure informationnelle, les notions de *topique/focus* sont supposées s'opposer, du moins se montrer distinctes : « Topic is what the sentence is about ; focus is what is predicated about the topic. » (Gundel & Fretheim 2004 : 176). Du point de vue psychologique, la notion de focus recouvre celle de topique, mais dans son acception informationnelle, elle s'y oppose !

La notion de *topique*<sup>148</sup> à son tour ne va pas sans poser problème : elle peut être appréhendée, au niveau de l'énoncé, selon les modalités de livraison de l'information : « topic is *given* in relation to focus and focus represents the *new information* predicated about the topic » (Gundel & Fretheim 2004 : 4, nous soulignons). En cela, l'opposition se superpose à celle de *thème vs rhème*, liées à la structure informationnelle. Le topique peut en outre être envisagé d'un point de vue positionnel comme le « point de départ » de l'énoncé (Demol 2010 : 172). Au niveau syntaxique, certains assimilent le topique au sujet de la phrase (Givón 1984). D'autre part, le topique peut être interprété au niveau discursif, en tant que participant le plus central du discours (Givón 1983, 1992). Kleiber (1994b) parle également de *topique discursif*, en ne recourant paradoxalement qu'à des exemples de deux phrases, comme le relève Demol (2010 : 173). En définitive, on constate que les définitions se montrent souvent intuitives et à l'origine de problèmes d'identification (Grobet 2002 : 27-29). Par ailleurs, les niveaux d'analyse du topique – syntaxique, cognitif, sémantique – sont fréquemment confondus (Demol 2010 : 173). Cette imprécision se reporte sans surprise sur le dérivé *topicalité*. Demol (2010 : 141) calcule le *degré de topicalité* par le nombre de renvois opérés sur un même référent. D'ailleurs, elle remplace indifféremment l'expression par celle de *degré de saillance*, à l'image d'Ariel (1988, 1990), qui recourt à tantôt à la notion de topicalité, tantôt à celle de saillance comme fondement de l'un des critères d'accessibilité.

Tout compte fait, les conceptions concurrentes de focus et topique et leurs délimitations tantôt superposées, tantôt croisées, tantôt distinctes, contribuent à entretenir le flou scientifique. Pour cette raison, nous préférons nous passer de ces notions dans nos analyses. Lorsque nous évoquerons des questions strictement liées à la structure informationnelle, nous conserverons l'opposition *thème-rhème*, où le *thème* est envisagé exclusivement en termes d'*aboutness* (ce dont on parle) et le *rhème* désigne l'information nouvelle pertinente, sur laquelle portent notamment les modalités d'énoncé (négation, interrogation, etc.) ou dont la réalisation est susceptible d'entrer dans un dispositif de clivage. Pour ce qui concerne l'importance cognitive accordée à un objet-de-discours, nous parlerons de son degré de *saillance* ou d'*activation* pour un état donné de M (Groupe de Fribourg 2012 : 126-127). La notion de saillance a le mérite de jouer sur un seul tableau, le niveau cognitif<sup>149</sup>, là où celles de *topique* et de *focus* se rencontrent sur de nombreux terrains, de manière parfois indifférenciée et prêtant de la sorte à confusion.

---

<sup>148</sup> Voir Demol (2010) pour un exposé récent des diverses définitions proposées.

<sup>149</sup> Selon Landragin (2004), la notion émane d'une analogie avec le domaine de la perception visuelle.

Landragin (2004) définit la saillance comme le caractère prépondérant, prégnant de certains éléments : un élément saillant *ressort en premier* dans les représentations cognitives du sujet parlant. Il est vrai que la notion met en jeu une multiplicité de facteurs de niveaux d'analyse différents (Kibrik 2011 : 390 *sqq*<sup>150</sup>), dont certains se montrent difficilement mesurables et impliquent forcément une part d'intuition : facteurs sémantiques, prosodiques, syntaxiques, inférentiels, perceptifs, récence d'introduction d'un référent, intentions et attention des interactants, schémas culturels, etc. (Landragin 2004.). Au vu de notre conception du discours (*cf. supra* Ch.I §6.3.2), on doit supposer que toutes sortes de paramètres contextuels, même si le linguiste n'y a pas accès, influent sur l'attention portée à un référent en M. Et parmi les objets à la portée des interlocuteurs dans un état donné de M, il va de soi que certains apparaissent relativement saillants alors que d'autres demeurent plus en retrait.

On peut tenter de clarifier cette façon intuitive de voir les choses en distinguant, avec Apothéloz (1995a : 169-170), deux sortes de saillance : une *saillance locale*, qui dépend de facteurs « accidentels » de la situation immédiate, tels que la récence d'activation (par des moyens verbaux ou non) ou les propriétés perceptives particulières de l'objet, etc. La *saillance cognitive*, quant à elle, recouvre la centralité d'un objet sur le plan de l'organisation des représentations cognitives en mémoire discursive et a trait à la notion de pertinence :

l'objet le plus saillant cognitivement est alors celui qui occupe la position la plus centrale dans l'univers d'objets considéré : c'est aussi, vraisemblablement, celui dont l'effet organisateur est le plus fort dans cet univers. (*ibid.* : 170)

La saillance locale peut être représentée comme une pile d'objets au sommet de laquelle se tient momentanément le plus saillant d'entre eux, tandis que la saillance cognitive s'apparente davantage à une structure en graphe ou treillis au centre de laquelle se trouve l'objet le plus saillant au niveau de l'organisation de M, qui constitue le nœud vers lequel converge le plus grand nombre de liens (Apothéloz 1995a : *ibid.*) : il devient ainsi le *centre organisateur* (*cf. supra* Ch.I §6.3.4). La manière dont ces deux types de saillance interagissent reste néanmoins assez floue et théorique, mais leur distinction peut déjà permettre d'affiner et de nuancer certaines situations.

Nous proposons un exemple pour illustrer la manière dont peut se manifester la saillance d'un objet en M. La devinette ci-dessous présente la particularité de jouer sur le caractère implicite d'un référent pourtant bien activé en M :

(157) Qu'est-ce que c'est :

- Lorsque je marche, je mets parfois ma main dans la poche de mon pantalon pour le tripoter. - Il n'est pas rare qu'en m'endormant, je le touche
- Il m'arrive aussi de le soupeser involontairement quand je lis mon journal, comme

---

<sup>150</sup> Kibrik (2011) adopte une approche pluri-factorielle du degré d'activation référentielle. Il dégage des facteurs liés à la présence d'un antécédent (en termes de distance, de syntaxe et de sémantique) et d'autres liés aux propriétés mêmes du référent (animation, protagoniste, etc.). Nous nous distinguons toutefois de cette approche qui se fonde essentiellement sur des paramètres propres à l'écrit et qui considère uniquement le degré d'activation pour le choix de l'expression référentielle.

pour m'assurer qu'il est toujours à sa place

Ben... mon téléphone portable, quoi ! A quoi ton esprit tordu t'a mené ?

(<https://fr.answers.yahoo.com/question/index?qid=20111115232751AAKrGwU>)

Dans cet exemple, la question initiale introduit un objet momentanément indéterminé, le plaçant à ce moment-là non seulement au « sommet de la pile », mais également, au vu du genre de discours à l'œuvre (une devinette) et des attentes à cet égard, directement au centre d'un réseau référentiel à élaborer. Malgré la justification finale du locuteur se déresponsabilisant de l'interprétation « commise » par le destinataire, de nombreux paramètres contribuent à maintenir au plus haut l'activation de l'objet incriminé tout au long du discours : au genre de la devinette (dont l'enjeu est la découverte d'un référent), on peut ajouter des facteurs culturels (le caractère tabou du référent insinué, dont l'inférence est appuyée par les verbes *tripoter x*, *toucher x*, *soupeser x*, etc. induisant des schémas d'actions typiques) et surtout, le nombre de renvois opérés sur lui (*ce*, *le*, *sa*, *il*). Ces éléments concourent à rendre le référent implicite plus important que d'autres objets introduits dans cet extrait, comme la 'main', le 'pantalon' ou le 'journal' qui, tout en modifiant tour à tour le sommet de la pile au niveau local, se greffent et gravitent ainsi autour de lui au niveau cognitif, faisant de lui le *centre organisateur* du réseau référentiel construit. Cet extrait illustre en outre une situation de désaccord entre les interlocuteurs sur le contenu exact de M (*cf. supra* Ch. §6.3.4) : l'auteur de la devinette pousse délibérément son destinataire à inférer un type d'objet en fonction des indices à disposition pour finalement mettre en cause l'interprétation produite et fournir une solution alternative et insoupçonnée.

\*

Dans la section suivante, nous allons voir que l'hypothèse « cognitive », à savoir le maintien de la référence à un objet déjà saillant, au-delà des problèmes terminologiques pointés, rencontre aussi des obstacles au niveau théorique. En effet, certains faits s'y montrent réfractaires, en particulier les emplois « indirects » du pronom de 3<sup>e</sup> personne que nous allons aborder ci-après, ainsi que l'emploi « sous-déterminé » de la 6<sup>e</sup> personne (*ils*) étudié au chapitre V. Certes, il ne s'agit pas de remettre en cause l'aptitude du pronom à fonctionner dans des circonstances de continuité, appuyée par nombre de données et d'études<sup>151</sup>. Mais le maintien de la saillance est une motivation – fût-elle la plus fréquente – parmi d'autres possibles, au recours à la forme en question. Notre ambition, dans la section suivante, est ainsi de mettre en lumière d'autres facteurs potentiels favorisant le choix du pronom.

---

<sup>151</sup> Voir Demol (2010) pour un état de la question et une étude empirique de *il/celui-ci* sur un corpus journalistique.

#### 4. L'anaphore pronominale indirecte ou associative

Dans le cas où aucun objet en mémoire ne correspond au pointeur linguistique, l'interprétation peut passer par l'inférence d'un nouvel objet cible à partir d'un objet-support déjà valide en M. Pour décrire cette procédure, on peut faire appel à la notion d'*anaphore associative*, parfois aussi appelée *indirecte*. Cornish (2001) en fournit la définition suivante :

Ce terme [*anaphore indirecte*] dénote des configurations où un anaphorique ne reprend pas le référent évoqué par son antécédent textuel, mais pointe vers un autre qui pourra lui être associé d'une façon ou d'une autre (relation « partie-tout », de métonymie, de « classe-ensemble », ou plus généralement « associative »). (p. 1)

L'origine du concept d'*association* utilisé pour caractériser la relation référentielle remonte à Guillaume (1919), qui qualifie d'*associatives* les expressions en italiques ci-dessous (Kleiber *et al.* 1994) :

- (158) Et comme le voyageur passait alors devant l'église, les saints personnages qui étaient peints sur *les vitraux* parurent avoir de l'effroi. *Le prêtre* agenouillé devant *l'autel* oublia sa prière. (p. 163)

Le caractère défini de ces SN découle ainsi de leur *association* avec un référent mentionné en amont via le SN *l'église*. Pour certains auteurs, le terme d'*anaphore associative* est réservé à l'emploi des SN définis lexicaux de ce type (*cf. infra* Corblin 1987, Kleiber 1990a, Kleiber *et al.* 1994). En ce qui concerne les pronoms, leur aptitude à fonctionner de la sorte est tantôt contestée (Erkū & Gundel 1987, Sanford & Garrod 1981, Ariel 1990), tantôt reconnue, comme on le verra, sous réserve de conditions d'emploi strictes (Kleiber 1990a, Cornish 2001, Cornish *et al.* 2005) ou comme des « ratés de l'émission » (Charolles 1990, Gundel *et al.* 2000). Néanmoins, quelques auteurs s'efforcent de dégager les régularités d'emploi des pronoms associatifs (Béguelin 1993a, Johnsen 2013, Berrendonner 2014). Avant d'aborder la question des pronoms de ce type (§4.2), il est toutefois nécessaire d'examiner de plus près la notion générale d'anaphore associative (§4.1), qui a en effet suscité une vive polémique en linguistique française à la fin du XX<sup>e</sup> siècle, qu'il s'agisse de sa délimitation ou du mécanisme inférentiel qu'elle met en jeu.

##### 4.1 L'anaphore associative

Le procédé référentiel de l'anaphore associative comprend des facteurs de différents ordres, à savoir sémantiques, formels, cognitifs et pragmatiques. Prenons des exemples canoniques tels que :

- (159) J'ai acheté un stylo hier, mais *la plume* est cassée (Azoulay 1978)

- (160) Il s'abrita sous un vieux tilleul. *Le tronc* était tout craquelé (Fradin 1984)

- (161) Nous arrivâmes dans un village. *L'église* était située sur une hauteur (Kleiber 1992c)

Par sa nature « définie », le SN anaphorique introduit un référent nouveau « sur le mode du connu ». C'est en vertu d'une relation associative existant entre le référent disponible et celui à inférer (ici, la plume du stylo, le tronc du vieux tilleul, l'église du village dont on vient de parler) que l'on peut conclure à l'existence de ce dernier. On reconnaît généralement deux éléments en jeu dans le processus référentiel : une dépendance interprétative entre l'expression référentielle et le contexte préalable, et une disjonction entre le référent implicitement introduit et celui qui a servi de point de départ à l'inférence (Charolles 1990).

Bien que mis au jour par Guillaume (1919), le phénomène de l'anaphore associative alimente surtout au début les recherches sur l'article défini et n'est pas étudié pour lui-même, considéré comme n'ayant « au fond rien de bien mystérieux » (Kleiber *et al.* 1994). Ce n'est que vers la fin du XX<sup>e</sup> siècle que l'anaphore associative fait l'objet d'une attention particulière (entre autres Hawkins 1978, Azoulay 1978, Fradin 1984, Erkü & Gundel 1987, [Reichler-]Béguelin 1989, Charolles 1990, 1994, 1999, Kleiber 1990b, 1992b, 1999, 2001a, 2001b, 2004, Berrendonner 1995, Apothéloz & Béguelin 1999, Cornish 2001), et que se manifestent des tentatives de délimitation et de restriction, en même temps que des tentatives de généralisation du phénomène, démarches aboutissant à des conceptions divergentes, à l'origine d'un large débat.

#### 4.1.1 Conception étroite

La conception qualifiée d'*étroite* ou de *standard* par Kleiber *et al.* (1994) dissocie les notions d'anaphore associative et d'anaphore indirecte. Ses tenants considèrent que l'anaphore associative représente un cas particulier de l'anaphore indirecte. Ils font intervenir dans la définition de l'anaphore associative des critères formels et sémantiques restrictifs (entre autres Corblin 1987, Kleiber *et al.* 1994). A noter qu'au sein de cette approche, les contraintes peuvent varier d'un auteur à l'autre. Un aspect important se situe dans la différence de comportement des diverses expressions référentielles, l'hypothèse étant qu'un changement de catégorie linguistique entraîne un changement de configuration discursive (*ibid.*). Aussi, sur la base de jugements d'acceptabilité, les auteurs soutiennent que toutes les formes ne sont pas pareillement aptes à supporter une anaphore associative :

(162) ? Nous arrivâmes dans un village. *Cette église* était située sur une butte (*ibid.* : 48)

(163) ? Harry roulait vers Londres. *Elle* tomba en panne à mi-chemin (*ibid.* trad. de Sanford *et al.* 1983)

Ces exemples sont supposés révéler des restrictions d'emploi définitives de l'anaphore associative, réservée à l'emploi des SN définis. De ce point de vue, il ne suffit pas que l'on comprenne de quel référent il s'agit pour que l'on puisse approuver de tels enchaînements. Au contraire, leur malformation prouve, aux yeux des auteurs, l'exclusion des SN démonstratifs et des pronoms en site associatif. A leurs opposants qui exhibent des données authentiques infirmant cette position (*e.g. Il neige et elle tient* < [Reichler-]Béguelin 1988, *cf. infra* (185),

ils mettent en cause le caractère discutable des énoncés, répliquant que « l'authenticité n'est pas un gage de bonne formation, sinon il n'y aurait plus d'erreurs ». Ils leur reprochent la surpuissance de leur modèle, qui prédit « plus d'enchaînements mal formés que d'enchaînements bien formés » (Kleiber *et al.* 1994 : 50). On peut néanmoins rappeler les risques de circularité d'un tel modèle (*cf. supra* Ch.I §6.3.1), les auteurs jugeant eux-mêmes de l'acceptabilité des données inventées par leurs soins, s'imposant de la sorte comme les garants du bon usage. L'opérationnalité du modèle en question s'explique, au fond, par des jugements de valeur et la mise à l'écart des données récalcitrantes...

Dans l'approche étroite, seules les opérations inférentielles de déduction sont tolérées, en particulier celles qui vont « du tout vers la partie », car « une entité se définit généralement par les ingrédients qui la constituent » (Kleiber *et al.* 1994 : 42). Cela explique le monopole du SN défini, impliquant selon les théories dominantes, au vu du présupposé d'existence et d'unicité qu'il véhicule, que l'on dispose des informations nécessaires à la saturation référentielle (Milner 1982, Corblin 1987, Kleiber *et al.* 1994). Outre le caractère orienté de la relation méronymique invoquée, celle-ci doit, selon Kleiber *et al.* (1994), être inscrite au niveau lexical : il peut s'agir d'un savoir conventionnel nécessaire (*e.g.* le lien entre une 'lame' et un 'couteau') ou à caractère stéréotypique (*e.g.* le lien entre une 'église' et un 'village').

Ajoutons pour terminer<sup>152</sup> un critère textuel à la définition étroite de l'anaphore associative : le caractère délimitable de l'antécédent, idéalement représenté par un SN, ou alors un « prédicat », mais avec des conditions strictes<sup>153</sup> (Kleiber *et al.* 1994, Kleiber 1997b) :

(164) Pierre se coupa du pain puis rangea *le couteau* (Corblin 1987)

Dans ce cas, la relation exprimée via l'antécédent et l'expression anaphorique est de type actantiel, où le référent du SN *le couteau* représente un actant (l'instrument) du prédicat de *couper*, cette relation devant être préconstruite au niveau lexico-stéréotypique : « il suffit de la mention du prédicat pour que ses arguments soient également disponibles » (Kleiber 1997b : 95).

En somme, toutes les données qui s'écartent des critères invoqués sont exclues du domaine de l'anaphore associative. Parmi les faits refoulés, on peut noter :

- Les SN définis sous la portée d'un restricteur de domaine :

---

<sup>152</sup> Deux contraintes sémantiques supplémentaires sont ajoutées par Kleiber *et al.* (1994) et Kleiber (1999), à savoir la *condition d'aliénation* du référent à inférer et le *principe de congruence ontologique* entre les référents mis en relation. La première contrainte indique que le référent à inférer doit être envisageable comme « un individu autonome » (Kleiber 1999 : 85). La seconde, qui explique la difficulté d'inférer des entités syncatégorématiques comme les propriétés ou les événements, exige que les deux référents impliqués dans le processus d'inférence soient du même type ontologique : selon l'auteur, il est ardu d'opérer l'aliénation, à partir de l'existence d'une voiture, de sa couleur ou de sa course, contrairement à son volant (*ibid.* : 89).

<sup>153</sup> *E.g.*, le nom de l'expression anaphorique doit être de niveau basique, comme *couteau* dans l'exemple subséquent, par opposition à *canif* (hyponyme) ou *instrument* (hyperonyme).

- (165) En France, *le président* voyage beaucoup (Kleiber *et al.* 1994)
- les emplois inférables « situationnels » :
- (166) *Ce train/il* a toujours du retard (A à B qui attend comme lui le Vintimille sur le quai de la gare, Kleiber 2004 : 290)
- les SN démonstratifs « textuellement » inférables :
- (167) J'ai planté un acacia, parce que *ces arbres* ne craignent pas le froid (*ibid.*)
- les SN démonstratifs mémoriels :
- (168) Souviens-toi de nos vacances : *cette mer, ce sable, ce ciel !* (Wilmet 1986)
- les *ils* « génériques textuels indirects » :
- (169) Paul a acheté une Toyota, parce qu'*elles* sont robustes (Kleiber 2004 : 290)
- les *ils* dits collectifs :
- (170) Paul a été à l'hôpital. *Ils* l'ont soigné énergiquement (*ibid.*)

Parmi ces procédés référentiels, certains (les quatre derniers) sont considérés comme des anaphores indirectes, à défaut d'être des anaphores associatives (l'anaphore associative se révélant comme un sous-type de l'anaphore indirecte). Sans entrer dans le détail de chacune de ces exclusions, on peut percevoir dans l'approche étroite la volonté de traiter séparément des faits en présence, sur la base de critères sémantico-formels stricts que révèlent, entre autres, les jugements d'acceptabilité.

#### 4.1.2 Conception large

Dans une conception dite *large* (Kleiber *et al.* 1994) au contraire, les chercheurs observent une variété de relations associatives<sup>154</sup> mises en œuvre au moyen de diverses expressions référentielles (entre autres [Reichler-]Béguelin 1989, 1993a, Berrendonner 1994, 1995, Berrendonner & [Reichler-]Béguelin 1995, Apothélos & Béguelin 1999, Charolles 1990, Charolles & Choi-Jonin 1995). L'un des points de départ est le fait qu'on n'observe pas de différence de nature dans l'interprétation d'une anaphore coréférentielle et celle d'une anaphore associative, si bien que si l'on en venait à considérer la fréquence des données, on pourrait être amené à « inverser l'ordre des préséances » (Charolles 1990 : 3).

C'est d'ailleurs le parti-pris risqué qu'illustre le modèle fribourgeois, en mettant l'anaphore associative au premier plan et en ramenant la variante coréférentielle à un cas particulier. Berrendonner (1995) s'inspire de la notion abstraite d'*ingrédience* de la méréologie lesniewskienne (1989), qu'il adapte pour appréhender la relation associative de manière

---

<sup>154</sup> Voir aussi Erkü & Gundel (1987), ainsi que Clark (1977) qui répertorie de nombreux types de pontage inférentiel (*inferential bridging*).

flexible et non limitée à un strict rapport partie-tout tel que le prône la conception étroite : ainsi, le foncteur que Berrendonner (1995) nomme « ingrédient » caractérise aussi bien des liens entre actants et procès<sup>155</sup> (171), des rapports de simple contiguïté (localisation, matière, etc.) (172)<sup>156</sup> et même des relations d'identité (173), comme c'est le cas pour l'anaphore coréférentielle :

(171) Une supposition que tu attrapes le diamant, qu'est-ce que tu ferais de *tout cet argent* ? (M. Aymé < Berrendonner 1995 : 251)

(172) Il y avait une valise sur le lit. *Le cuir / les vêtements* étai(en)t tout taché(s) (Charolles 1990)

(173) Marc baissa la tête, croisa les bras, brusquement calmé. *Il* jeta un rapide regard vers la cantatrice du front Ouest. (F. Vargas, *Debout les morts*, p. 45)

Cette dernière situation illustre une conception large et réflexive de la relation d'inclusion dans une théorie ensembliste en logique naturelle, d'après laquelle on peut considérer que « tout objet est ingrédient de lui-même » (Berrendonner 1995 : 253). En bref, les situations d'identité, autrement dit de coréférence, ne représentent qu'un cas particulier de la relation d'ingrédience ainsi conçue. Charolles (1990) note également de son côté la diversité des relations associatives, parmi lesquelles « toute forme de contiguïté référentielle », y compris les rapports les plus vagues, comme ci-dessous :

(174) Sophie dormait, *l'avion* survolait l'Océan indien (*ibid.* : 13)

Dans ce cas, c'est le discours qui impose la relation d'association, poussant l'interprète à inférer que « "l'avion" a à voir avec le sommeil de Sophie »<sup>157</sup>.

Cette approche généralisante de l'anaphore s'oppose à la définition « étroite » par le fait qu'elle admet de nombreux parcours inférentiels pour l'interprétation des anaphores associatives. On peut déjà remarquer que les opérations cognitives à l'œuvre, présentées exclusivement comme des déductions dans la conception étroite (*e.g.* (159), (160) et (161)), conformément à une « règle » du genre 'si on a le tout, on a la partie', sont sans doute beaucoup plus complexes que cela, impliquant plusieurs étapes de différentes natures dans le raisonnement. Or, l'interaction des divers facteurs cognitifs en jeu ne sont jamais véritablement décrits dans les travaux. Sans pouvoir combler entièrement ce manque, nous pouvons néanmoins revenir sur le cas de (161) en signalant la nécessité d'opérer avant toute déduction une induction du type 'les villages de cette région comportent une église', qui ne

---

<sup>155</sup> Le processus interprétatif se fonderait notamment sur l'existence de structures cognitives telles que les *scénarios* (Sanford & Garrod 1981), *cadres (frames)* < Minsky 1975) ou encore *scripts* (Schank & Abelson 1977).

<sup>156</sup> Cf. les pontages inférentiels variés chez Clark (1977).

<sup>157</sup> Il reste toutefois difficile, à notre avis, d'interpréter cet énoncé indépendamment de tout contexte. C'est d'ailleurs sur les questions qui touchent à la nature et au tri des données que se distinguent principalement Charolles et le Groupe de Fribourg.

représente pas forcément un stéréotype pour tout interprète si l'on tient compte de facteurs culturels. Dans le même ordre d'idées, Berrendonner (1995) fournit cet exemple authentique impliquant manifestement, entre autres, une opération d'induction, c'est-à-dire où l'interprète est amené à construire un stéréotype à partir du discours même (p. 246) :

(175) Hier soir vers 19h15, un incendie s'est déclaré dans la cheminée d'une ancienne ferme à Neyruz<sup>158</sup>. [...] Le feu a pris au sommet de *la borne*, dans la partie non tubée (presse < *ibid.* : 238)

C'est ici le discours qui permet à l'allocutaire d'inférer une relation « à parfum générique » (Charolles 1990), en l'occurrence que la borne (en tant que *type*) est une partie de la cheminée dans les fermes fribourgeoises<sup>159</sup>, qui constitue ainsi la « prémisses nomologique » de l'interprétation associative (Berrendonner 1995 : 246). Par ailleurs, on peut noter par ce raisonnement l'enrichissement potentiel des connaissances de type générique pour l'interlocuteur (*ibid.* : 247, Charolles 1990, 1999).

Berrendonner (1995) relève en outre, à la suite de Charolles (1990), des parcours inférentiels de type abductif, contestés par d'autres (*e.g.* Kleiber 1992c) :

(176) Le cheval d'Yvain déclencha un mécanisme qui fit tomber sur eux une porte de fer. Enfermé dans *cette salle*, Yvain était complètement désorienté (copie d'élève < *ibid.* : 246)

Dans un exemple de ce type, le parcours présente une orientation inverse à celle de la déduction : il obéit à un calcul probabiliste du genre « si on a la partie, on a le tout », dont la validité n'est pas garantie. En l'occurrence, on conclut à l'existence d'une salle à partir de celle d'une porte de fer, au détriment d'autres hypothèses envisageables (*e.g.* l'existence d'un jardin, d'une cave, etc.). Ces deux derniers exemples montrent également que le discours lui-même, plutôt qu'un savoir lexico-stéréotypique, peut être à l'origine de la construction de la relation associative (Erkū & Gundel 1987, Charolles 1990, 1994, 1999, Berrendonner 1995). C'est aussi le cas de l'exemple suivant proposé par Charolles (1990), tiré d'une fable de La Fontaine :

(177) Un pâtre à ses brebis trouvant quelque mécompte,  
Voulut à toute force attraper *le larron* (La Fontaine, VI, I < *ibid.*)

Selon Charolles, un 'mécompte' n'implique pas lexicalement l'existence d'un larron. C'est bien le contexte discursif, par « présomption de cohérence » qui permet de cautionner l'inférence et de promouvoir « des relations inédites, conjoncturelles, qui peuvent n'avoir qu'une validité occasionnelle » (*ibid.* : 14). Cela montre non seulement que les relations

---

<sup>158</sup> Neyruz est le nom d'un village du canton de Fribourg en Suisse.

<sup>159</sup> On pourrait toutefois objecter que pour le lecteur fribourgeois de la presse locale, le stéréotype n'a pas à être construit, car il fait déjà partie de ses connaissances culturelles. Si le lectorat est en revanche plus vaste, il y a en effet des chances qu'il ne soit pas au fait du stéréotype.

d'association dépassent largement le savoir « a priori » (lexical, stéréotypique), mais aussi que « l'inférence fonctionne à reculons, pour établir un lien (un pont) qui ne peut pas ne pas être, puisqu'il est manifeste » (*ibid.* : 4).

Comme en témoignent (171) et (176), la conception large admet une diversité de formes supportant l'anaphore associative. A ce titre, les chercheurs fribourgeois qualifient sans scrupule d'*associatifs* des indices de 3<sup>e</sup> personne ou des SN démonstratifs fonctionnant de la sorte<sup>160</sup> :

(178) Le danger d'être piqué par une abeille est le plus grand entre mai et juillet. Pendant cette période, *elles* essaient le plus fréquemment (circulaire médicale, [Reichler-]Béguelin 1993a : 337)

(179) Il est vrai que lorsque nous lisons, nous ne pensons pas que *cette histoire* est en train de vivre, de prendre forme grâce à nous. (copie de bac < [Reichler-]Béguelin 1989 : 312)

L'exemple (178) présente l'introduction d'un individu quelconque de type N (une *abeille*), suivie du rappel de la classe extensionnelle correspondante via le pronom clitique *elles*. Rappelons que Kleiber (2004 : 290) les traite comme des anaphores indirectes de type générique, mais non associatives (*cf. ex.* (169)). En (179), l'expression démonstrative renvoie à un référent implicite représentant l'objet effectué du processus de lecture qui vient d'être évoqué. Aux yeux des Fribourgeois, les deux exemples s'appliquent parfaitement à la description du mécanisme de l'anaphore associative. S'ils reconnaissent des valeurs propres aux différentes formes d'expression référentielle, ils en jugent l'emploi relativement flexible et en concluent : « it would be vain to establish a strict correlation between types of expressions and types of uses » (Apothéloz & Béguelin 1999 : 369). Le choix des formes d'expressions référentielles répond davantage à la mise en œuvre de stratégies guidées par les objectifs et besoins en cours des locuteurs et interlocuteurs (Béguelin 1997a). Ainsi, au reproche de largesse du modèle accusé de générer des énoncés déviants, les auteurs répliquent que la prise en compte des facteurs contextuels permet justement de fournir une explication aux sentiments de « déviance » des usagers. L'approche s'efforce ainsi de « favoriser l'émergence de généralisations en rapprochant, sous le terme d'« associatives », toutes les expressions référentielles dont l'interprétation met en jeu des opérations de raisonnement » (Berrendonner & [Reichler-]Béguelin 1995 : 26).

Pour finir, la contrainte d'un antécédent nominal est elle aussi battue en brèche. Hawkins (1977), Erkü & Gundel (1987), Béguelin (1997b) et Apothéloz & Béguelin (1999) remarquent que les sources sollicitées dans le processus inférentiel ne sont pas exclusivement de nature

---

<sup>160</sup> Charolles (1990), exploitant le critère d'acceptabilité, est plus nuancé sur ce point, jugeant les exemples de Béguelin « à la limite de l'acceptabilité » ou alors, concernant les pronoms, comme « des ratés de l'émission ». Cependant, il reconnaît que les pronoms « en syllepse [...] manifestent la quintessence du processus en œuvre dans l'anaphore associative, puisque la force du rapport associatif qu'exploite l'anaphore pronominale est telle que le support du nom (N2) devient implicite. » (p. 10)

linguistique : des indices de la situation environnante, des comportements mimo-gestuels ou des savoirs d'arrière-plan sont susceptibles de se combiner indistinctement, comme ci-dessous :

- (180) [L2 a le nez rouge et souffle dans un gros mouchoir]  
- L1 : vous êtes grippé ?  
- L2 : Oui, j'essaie de *la* garder pour moi (oral < Béguelin 1997b)

Cet exemple montre bien qu'outre l'occurrence de l'adjectif *grippé* (duquel on peut inférer, par principe d'*isonymie*, la dénomination implicite du référent impliqué), les percepts situationnels (les indices mimo-gestuels exhibés) et l'expérience du monde (les effets de la grippe), contribuent ensemble au processus d'interprétation, de manière non discriminée.

### 4.1.3 Bilan

A l'issue de cette présentation des conceptions concurrentes de l'anaphore associative, on peut constater une divergence manifeste des objectifs respectifs des deux approches : l'approche dite étroite ou standard vise à inventorier et isoler des sous-types et ne retient comme associatives que les cas jugés consensuels, avec souvent l'intuition du linguiste comme garant du respect de la norme. L'approche large, dans laquelle nous nous inscrivons, tend à promouvoir un modèle général sur la base des usages observés, en faisant primer un mécanisme unique qui englobe une diversité de manifestations.

Certes, nous convenons que tous les types d'expressions référentielles ne s'utilisent pas indifféremment dans les mêmes contextes et que les composantes sémantiques jouent un rôle certain dans le choix de la forme. Néanmoins, la prise en compte des paramètres discursifs et contextuels dans le processus référentiel nous paraît tout aussi décisive dans l'explication du choix des marqueurs (*cf. supra* Ch.I §6.3.6). En outre, avec le développement exponentiel des ressources pour la collecte de données, il nous paraît indispensable de dépasser l'intuition comme critère scientifique et d'aller à la rencontre des faits réels désormais disponibles, notamment oraux, considérablement marginalisés dans les travaux sur l'anaphore. Dans le domaine des expressions référentielles en linguistique française, outre les travaux précurseurs fribourgeois, peu de chercheurs s'y sont engagés<sup>161</sup> et le présent travail a pour ambition de contribuer à combler ce manque en portant une attention particulière aux emplois en contexte d'oral spontané.

Reste un point d'ordre terminologique à clarifier. L'approche (« large ») que nous adoptons ne fait pas de différence entre les notions d'*anaphore associative* et *anaphore indirecte*. Néanmoins, afin d'éviter toute sorte de confusion, notamment celle qui associerait notre position à la conception « standard », nous emploierons désormais la notion d'*anaphore indirecte*.

---

<sup>161</sup> On peut relever l'intérêt pour l'oral de quelques travaux sur la résolution automatique des anaphores en TAL, comme Antoine (2004) et Muzerelle *et al.* (2012).

## 4.2 Contextes d'apparitions des pronoms indirects

Dans leur étude pragmatique sur l'anaphore indirecte, Erkü & Gundel (1987) soutiennent que celle-ci ne peut être réalisée par des pronoms : « the one formal property which appears to distinguish all indirect anaphors from direct ones is the fact that they cannot be pronominal » (p. 541). Sanford & Garrod (1981: 154), puis Ariel (1990) partagent cette idée : « I suggest that the reason why pronouns [...] cannot retrieve implied entites is that they mark too high and automatic a retrieval » (p. 185). Cependant, lorsqu'on se penche de plus près sur les faits, cette position se voit démentie. Dans cette section, nous présentons les travaux qui remettent en question ces conclusions et qui cherchent à éclairer les contextes d'apparition des pronoms indirects. Nous avons distingué les études selon le type d'approche, à savoir celles qui manifestent une optique « restrictive » et qui s'efforcent de déterminer les contraintes sémantiques ou cognitives pesant sur leur emploi (§4.2.1), puis celles qui s'inscrivent dans une perspective « descriptive » (§4.2.2), en tentant de dégager les circonstances récurrentes favorables à leur occurrence.

### 4.2.1 Restrictions d'emploi

Dans cette partie, nous présentons les démarches et hypothèses d'auteurs qui reconnaissent l'usage des pronoms indirects, en établissant toutefois des conditions strictes d'emploi qui font intervenir les notions de *saillance*, de *focus* ou encore de *centralité*.

Dans son article de 1990a, Kleiber admet l'existence de pronoms indirects, qu'il propose d'envisager selon un critère non pas textuel (avec ou sans antécédent), mais de présence ou d'absence préalable (linguistique ou situationnelle) du référent.

a) Emplois *in praesentia*, c'est-à-dire, où le référent est « donné » par le contexte :

(181) Fred enleva son chapeau. **Il** avait trop chaud. (ici une anaphore coréférentielle typique)

(182) Mais il est fou ! (face à un automobiliste qui fonce sur vous) = (140)

b) Emplois *in absentia*, (*il/ils* dits « inférables »<sup>162</sup>) (p. 27) :

(183) A Boston, *ils* roulent comme des fous (*ibid.*)

(184) *Il* va venir tout de suite. [un censeur d'une école s'adressant à la mère d'un élève qui attend devant le bureau du proviseur] (p. 33) = (75)

« L'acceptabilité » d'exemples comme (183)-(184) s'explique selon l'auteur par le respect des conditions sémantico-cognitives propres à l'emploi de *il*, pour rappel (*cf. supra* §3.3) :

la nécessité de renvoyer à un référent conçu comme classifié et l'indication de rechercher le référent dans une *situation* qui est manifeste (ou saillante) et dans laquelle le référent occupe une place d'argument. (p. 46)

---

<sup>162</sup> Cf. Prince (1981).

Il y oppose la malformation de (185), exemple rédigé par un élève :

(185) \*Il neige et elle tient ([Reichler-]Béguelin 1988)

Selon Kleiber, il manque une étape inférentielle déterminante, à savoir, le constat de la présence de la neige, pour pouvoir appréhender un renvoi par le clitique *elle* conformément à la thèse de la continuité. Il soutient qu'en l'état, on n'a pas affaire à une situation manifeste et justifie sa démarche par la volonté de

brider les possibilités de récupération inférentielle qu'entraîne un tel assouplissement pour ne pas donner lieu à l'excès contraire, à savoir une prolifération incontrôlée d'emplois inférables de *il*. (p. 43)

Hormis la démarche que nous avons déjà critiquée sur le tri des données (*cf. supra* §6.3.1), un aspect de l'analyse mérite une mise au point, à savoir, l'importance accordée au critère de la présence vs l'absence du référent, à l'origine d'un traitement inférentiel ou non du pronom (*cf. l'appellation il* « inférable »). Dans leur théorie de la pertinence, Sperber & Wilson (1986) montrent bien que la communication humaine fonctionne sur le mode inférentiel et que tout processus interprétatif consiste pour le destinataire à inférer un sens visé par le locuteur à partir des indices à disposition dans le contexte. La pertinence de ces inputs ne varie pas en fonction de la nature (stimuli ostensifs tels qu'énoncés, gestes ou inputs ordinaires tels que percepts, pensées, buts, souvenirs, etc.), mais en fonction des coûts et effets produits par leur traitement cognitif.

On a déjà soutenu *supra* (Ch.I §6.3.5) que tout pointage référentiel, y compris coréférentiel, implique une opération d'inférence qui mène à l'unification du pointeur avec un objet en M (que le référent soit déjà valide ou non en M). Cela peut être illustré simplement en modifiant l'exemple (181) :

(186) Fred enleva son chapeau. *Il* était mouillé.

L'ambiguïté se résout forcément à travers un pari probabiliste de type inférentiel en fonction des informations contextuelles à disposition. En fait, dans tous les cas, le pointeur référentiel donne l'instruction d'unifier son contenu avec un objet manifeste. Si celui-ci est effectivement valide dans M, l'inférence consiste à sélectionner le référent visé, tandis que s'il fait défaut, l'inférence consiste à l'introduire sur le champ, par coup de force, pour parer à ce manque. On objectera ainsi à l'analyse de Kleiber (mais aussi à d'autres analyses, *e.g.* Ariel 1990 invoquant une récupération « automatique » par les pronoms en général) que les types relevés, qu'ils soient « in praesentia » ou « in absentia », impliquent de toute façon un processus inférentiel interprétatif.

Gundel *et al.* (2000) s'intéressent quant à eux au fonctionnement des pronoms indirects en anglais, relativement à la Hiérarchie du Donné (voir *supra* Ch.I §4.6.2). Pour rappel, un pronom inaccentué requiert selon leur théorie un statut *en focus* du référent, autrement dit, il signale que le référent est déjà dans le centre d'attention de l'allocutaire au moment où il

apparaît. Un tel pronom se révèle donc *a priori* être un mauvais candidat pour l'anaphore indirecte, dans les cas où l'existence même du référent est à inférer. Cependant, les auteurs constatent que leurs données, récoltées à partir d'un corpus d'anglais oral et écrit, ne satisfont pas systématiquement aux prédictions de la Hiérarchie, étant donné l'attestation de pronoms indirects qui ne réfèrent pas à des objets « en focus » :

(187) My mum bought an exercise tape and so I'll go nuts and plays it in the morn and in the afternoon and do the added things *she* says for those that want a more intense workout (*she* = la femme sur la cassette de gym) [soutien. alt. désordres d'alimentation, p. 96]

En effet, Gundel *et al.* reconnaissent que le référent visé en (187) par *she*, à savoir 'la prof de gym' n'est pas « en focus » pour l'allocutaire au moment de l'occurrence du pronom, à l'inverse de 'la mère du locuteur', à ce moment-là bien meilleure candidate à une référence subséquente. Ici, c'est aussi bien la prédication 'the added things she says for those that want a more intense workout' que le *script* (Schank & Abelson 1977) évoqué par la cassette de gymnastique, qui mènent à l'interprétation du référent adéquat.

Gundel *et al.* remarquent que dans leur corpus, la plupart des exemples concernés par l'anaphore pronominale indirecte sont des cas de référence vague, dont l'identification référentielle n'apparaît pas pertinente :

(188) The other therapy I had to say good-bye to was day treatment (where I'd been going off and on since January) - - but it wasn't my own decision. *They* kicked me out. Why? For being TOO SICK ! ([soutien alt. désordres d'alimentation], p. 95)

Les auteurs notent à l'égard de cet exemple une quasi-équivalence sémantique de l'énoncé en *they* avec une tournure au passif sans agent exprimé (*I got kicked out*)<sup>163</sup>. Afin d'expliquer ces exemples, ils ont recours à la notion d'*accommodation* (Lewis 1979), qui consiste en une réparation de ces « violations mineures » commises envers la Hiérarchie, se traduisant par un assouplissement des conditions d'emploi de la forme utilisée. La motivation de telles « infractions » résiderait soit dans l'insouciance du locuteur à l'égard des représentations effectives de l'interlocuteur (*cf.* la maxime de nonchalance *supra* §6.3.6), soit, à l'inverse, dans sa confiance envers les capacités inférentielles de celui-ci. Dans ce second cas, le plus fréquent aux dires des auteurs, le locuteur mise sur une « accommodation » de la part de l'interprète, s'appuyant sur des facteurs tels que la reconnaissance de scripts (Schank & Abelson 1977), les connaissances d'arrière-plan, leur connivence, et surtout, la prédication hôte du pronom. On peut rapprocher la notion d'*accommodation* de celle de *catalyse* que [Reichler-]Béguelin (1995a) et Berrendonner & [Reichler-]Béguelin (1996) empruntent à Hjelmslev (1968) pour décrire la tâche interprétative de reconstitution d'éléments manquants en M. Ces traitements cognitifs peuvent être envisagés comme la conséquence d'un *coup de force présuppositionnel* (Ducrot 1972 : 51) qui désigne l'opération « illégitime » de

---

<sup>163</sup> Nous aurons l'occasion de nuancer cette « équivalence » *infra* (Ch.V §3.4.13.4.1).

présupposition de la part d'un locuteur, c'est-à-dire la présupposition d'éléments préalablement inexistant en M.

Si le processus d'interprétation décrit par Gundel *et al.* nous paraît pertinent à certains égards, nous tenons à relever le caractère passablement contradictoire de l'argumentation proposée, résidant dans le fait que les emplois vagues (selon eux la majorité des exemples) sont supposés ne pas nécessiter de résolution, alors que le processus invoqué d'*accommodation* implique précisément l'inférence d'un référent adéquat.

Enfin, malgré l'intérêt des données récoltées, nous pouvons regretter que les auteurs ne profitent pas de l'attestation de ces faits pour réviser les fondements de la Hiérarchie. Cette solution, quoiqu'esquissée, est finalement abandonnée au profit du maintien de « l'intégrité » de la théorie. Dès lors, les faits relevés sont relégués au rang d'exceptions (« violations mineures ») – aussi bien au niveau quantitatif (« relativement peu fréquents ») qu'au niveau diaphasique (« registre relâché »).

Cornish (2001) revient justement sur le statut requis par les pronoms selon la Hiérarchie de Gundel *et al.* (1993) et propose d'envisager la question en termes de *centralité* des référents, à l'appui de données comme celle-ci :

(189) Femme : Why didn't you write to me ?

Homme: I did Ø... started to Ø, but I always tore 'em up ([dialogue du film *Summer Holiday*], p. 7)

L'interprétation du pronom passe par la récupération d'une information préalablement disponible, bien qu'implicite, en mémoire. En effet, le lexème verbal 'write' du premier énoncé comprend sémantiquement dans sa structure actantielle l'objet effectué du procès, à savoir, l'objet de l'acte d'écriture, qu'on peut inférer, d'après le pointage pronominal et le contexte, comme représentant une classe de 'lettres', 'messages' ou autres 'mots doux'. Cornish juge l'extension du statut *en focus* imprécise, telle qu'exigée par l'usage des pronoms inaccentués d'après la *Hiérarchie du Donné*. Il introduit à cet effet le concept de *centralité*, qui distingue entre entité implicite nucléaire et périphérique, jugeant qu'un pronom inaccentué peut renvoyer à un référent implicite *nucléaire* comme les 'lettres' du prédicat 'écrire' en (189), mais non à un objet *périphérique*, comme dans :

(190) Mary dressed the baby. # They were made of pink wool. (Sanford & Garrod 1981 : 154)

En (190), Cornish explique l'incongruité de l'enchaînement par le fait que l'objet visé ne constitue qu'un instrument du prédicat 'dress'. L'hypothèse de Cornish (2001, 2005) se voit par la suite testée expérimentalement en français et en anglais dans Cornish *et al.* (2005) au moyen de mesures de temps de lecture de dialogues de ce genre :

- (191) L1 : Cet artiste a peint toute la journée hier.  
L2 : Oui et il *les* a vendu à bon prix aussi. (*implicite, nucléaire*)
- (192) L1 : Cet artiste a peint toute la journée hier. Ses tableaux ont vivement intéressé une passante très riche.  
L2 : Oui et il *les* a vendu à bon prix aussi. (*explicite, nucléaire*)
- (193) L1 : Cet artiste a peint toute la journée hier.  
L2 : Oui et il *les* a tous utilisés, du plus fin au plus épais. (*implicite, périphérique*)
- (194) L1 : Cet artiste a peint toute la journée hier. Ses pinceaux étaient nombreux et de tailles différentes.  
L2 : Oui et il *les* a tous utilisés, du plus fin au plus épais. (*explicite, périphérique*)

Le design expérimental permet aux auteurs de croiser deux facteurs : celui de la centralité et celui du caractère explicite ou non du référent, les prédictions étant que la référence anaphorique aux objets explicites et nucléaires faciliterait la lecture des énoncés. En effet, les résultats montrent que les énoncés présentant un pronom indirect sont lus plus lentement que ceux contenant un pronom direct et que les énoncés référant à des objets dits périphériques sont lus plus lentement que ceux référant à des objets dits nucléaires. En situation implicite, la référence aux objets périphériques provoque un temps de lecture plus long que la référence aux objets nucléaires, tandis qu'il n'y a pas de différence significative à cet égard en condition explicite. Il faut ajouter que les dialogues étaient suivis d'une assertion dont les sujets devaient juger la vérité/fausseté, du genre, à la suite de (191) : *L'artiste a pu vendre ses tableaux* (vrai). Ceci pour vérifier que la cible du pronom avait été bien interprétée, en particulier dans les conditions implicites. A cet égard, les auteurs relèvent un taux de bonnes réponses élevé (93%) et surtout aucune différence significative entre les conditions implicites nucléaire vs périphérique sur ce taux. L'expérience est répétée pour l'anglais et fournit des résultats similaires. Les auteurs en concluent<sup>164</sup> que ce n'est pas tant le caractère explicite ou implicite de l'introduction du référent en mémoire qui autorise ou non la reprise par un pronom non accentué, ainsi que le suggèrent de nombreux travaux antérieurs, mais plutôt la centralité relative du référent. Ils proposent une mise en perspective du concept scalaire de *centralité* avec la *Hiérarchie du Donné* et font figurer les référents nucléaires sur un intervalle comprenant les statuts *en focus* et *activé*, tolérant à leurs yeux une récupération via un pronom non accentué. Les référents périphériques, pour leur part, se voient plutôt attribuer le statut *familier*, stade jugé plus critique quant aux possibilités de rappels pronominaux.

Les résultats esquissés montrent qu'outre l'introduction explicite d'un référent, le schéma actantiel du verbe concerné ou le savoir conventionnel influent sur les capacités de rappel pronominal. Il est indéniable, en effet, qu'un objet nécessairement impliqué par un procès se montre plus « accessible » qu'un objet facultatif, dont il faut potentiellement inférer

---

<sup>164</sup> A l'égard des travaux expérimentaux en psycholinguistique en général, il nous semble qu'une réflexion pourrait être menée sur l'interprétation qu'on tire unanimement des mesures de temps de lecture.

l'existence. Il demeure à nos yeux cependant une certaine équivoque autour de la notion de *centralité conceptuelle*, du fait que les auteurs rattachent celle-ci aussi bien à la structure lexico-sémantique du verbe qu'à une dimension cognitive (*cf.* le parallèle fait avec les statuts de Gundel *et al.* 1993). Or, il nous paraît difficile de préjuger de la centralité cognitive des référents à partir de la seule structure actantielle d'un prédicat (ou éventuellement d'un savoir stéréotypique), l'activation cognitive d'un référent étant fonction, comme nous l'avons évoqué *supra* (§3.4), d'une pluralité de paramètres.

\*

Si les études présentées ci-avant adoptent des méthodologies (intuitive, empirique, expérimentale) et des données (construites ou attestées) diverses, elles s'efforcent toutes de tracer une limite entre emplois possibles ou non des pronoms indirects en fonction de la prégnance du référent, dans la perspective du décodage essentiellement. Nous nous tournons à présent vers une manière alternative d'envisager la question, qui met en évidence les stratégies mises en œuvre par le locuteur.

#### 4.2.2 Régularités d'emploi

L'optique consiste à prendre acte des emplois réels et à déceler les régularités de fonctionnement observées en contexte. Béguelin (1993a, 1997a) inventorie pour sa part les conditions pragmatiques qui favorisent l'emploi des pronoms indirects, en relevant les routines inférentielles récurrentes. Elle précise d'abord la nature de la relation indirecte : « le pronom ne rappelle pas une information qui serait fournie explicitement dans le contexte linguistique. Son interprétation, référentielle autant que morphosyntaxique, nécessite le recours à une information non dite, à une forme lexicale inférée » (1993a : 327). L'interprétation se fait ainsi par le biais d'un calcul implicite qui permet la construction d'une inférence non stricte du genre « s'il y a x, il y a aussi y » (p. 335-336). Pour exposer les rendements de l'emploi des pronoms indirects, nous pouvons invoquer les objectifs communicationnels présentés *supra* (Ch.I §6.3.6)

Béguelin mentionne des stratégies d'optimisation de l'encodage (*cf.* O2), illustré ci-dessous à travers un principe « nonchalant » d'économie lexicale, au détriment d'O1 (pour rappel, l'optimisation du décodage) :

(195) L1 : J'ai tout préparé pour la fondue.

L2 : Tu as rempli le réchaud ?

L1 : Non, pas ça, mais par contre il y **en** a dans le tiroir (oral < *ibid.* : 371)

La reconstitution du « coup de force », par catalyse (ou accommodation), du référent de *en* passe par la prise en compte d'un scénario conventionnel (la préparation d'un repas de fondue et des ustensiles), visiblement bien connu des interlocuteurs. L'inférence de la forme lexicale sous-jacente peut néanmoins demeurer incertaine, comme dans ces exemples d'oral que nous avons relevés ci-dessous :

- (196) BM: Il y a quand même plus qu'un épicier hein, il faut pas qu'il lui arrive un coup dur et, un boulan/ il y a deux boulangers, mais autrement ben il y a plus rien hein. Avant il y avait pff, des épiciers il y en avait quatre ou cinq euh, euh, la confection il y *en* avait deux ou trois euh, maintenant il y en a plus qu'un pff un petit qui, qui vivote hein. C'est vrai que ça, ça fait mal quand même (pfc)
- (197) L1 : ouais enfin ceci dit si on a ouvert le magasin à C- un magasin de jonglage à Caen c'est pas innocent hein c'est qu'il y en a pas mal hein  
 L2 : ah + et q- s- et comment ça se fait euh comment tu peux expliquer < ça  
 L1 : alors > là {rire} < ça je ne  
 L2 : c'est intéressant > ça  
 L1 : ouais ça c'est + je pense c'est je pense que c'est parce que depuis pas mal de temps il y a un magasin qui **en** vendait +  
 L2 : d'accord  
 L1 : donc forcément les gens euh se rassemblent autour d'un endroit où ils peuvent avoir le matériel facilement quoi + puis nous on a plus ou moins pris la relève quoi ben c'est du c'est de l'endroit d'où je viens quoi (crfp)

Dans l'exemple (196), on ne sait pas si le locuteur privilégie l'indice morphologique pour la catégorisation implicite de son référent (des *confectionneurs*, lexème toutefois peu courant) ou s'il a à l'esprit une classe de dénommés *artisans, fabricants, vendeurs*, etc. D'ailleurs, l'allocutaire n'est pas obligé d'en inférer une dénomination congruente : il tâchera d'inférer un type de référent qui soit le plus compatible avec les indices à disposition. Il en va de même en (197), où la dénomination sous-jacente du référent de *en*, tel que le conçoit l'encodeur, pourrait aussi bien être *matériel, équipement, articles, instruments* (de jonglage), etc., mais où les indices contextuels ne permettent pas de trancher. A noter que dans tous ces cas où le pointeur pronominal introduit implicitement un référent nouveau, on peut invoquer l'objectif O4, pour rappel, l'apport d'informations inédites, qui se greffe sur l'acte référentiel proprement dit.

Un parcours inférentiel fréquemment mis en œuvre via les pronoms indirects est celui qui assimile un individu collectif à la classe de ses membres. Ce procédé permet lui aussi au locuteur de s'épargner, via O2, une tâche de recatégorisation lexicale du référent :

- (198) JO1: Et euh, bon ben j'avais pas de logement ni rien à Madrid. <E: Tu es parti comme ça?> Qu'est-ce que j'ai fait? En fait oui. J'ai téléphoné avant à, à l'Ambassade de France à Madrid pour savoir si jamais bon ben **ils** pouvaient trouver quelque chose, sait-on jamais. (pfc)

Dans ce cas-là, on peut invoquer la notion de dualité (*cf. supra* Ch.I §3.3), car les indices contradictoires de formatage de l'objet interviennent au sein d'une même *clause* (*cf. supra* Ch.I §6.3.3, pour rappel, un îlot de dépendances morpho-syntaxiques). Outre l'économie à l'encodage, on pourrait aussi se demander si le « désaccord » morphosyntaxique ne

correspond pas à une stratégie de désambiguïsation (O1), permettant de prévenir l'inférence d'une interprétation biologique que pourrait induire la marque de genre d'un *elle*.

D'autres cas d'« économie lexicale » s'expliquent plutôt par la volonté d'éviter des répétitions lexicales (O3) en privilégiant visiblement le maintien, par *isonymie*, de l'attribut de dénomination d'un référent déjà introduit, sans qu'il y ait toutefois coréférence :

(199) Il paraît peu probable qu'une Suisseuse s'imposera cette fois-ci. C'est regrettable, puisqu'*elles* sont majoritaires. (*La Suisse*, courrier, 29.01.89, [Reichler-]Béguelin 1993a : 337)

En effet, une redondance lexicale pourrait provoquer un effet de saturation perceptive. Là également, en induisant subrepticement l'inférence d'un nouvel objet, en l'occurrence la classe (*elles*) à laquelle appartient l'individu membre (*une Suisseuse*), le locuteur met en œuvre le principe d'apport informationnel (O4).

Un peu dans le même genre, nous proposons l'exemple suivant, qui met en jeu plusieurs occurrences de pronoms conduisant à considérer l'individu type de la classe extensionnelle préalablement introduite :

(200) elle [la confrérie] venait surtout en aide à à l'époque il y avait beaucoup de malheureux + mais elle venait surtout en aide même + aux étrangers + qui pouvaient mourir + euh dans notre village + et l'esprit philanthropique de la confrérie + c'est que il fallait *le* prendre en charge et *l'*enterrer surtout *s'il* n'avait personne (crfp)

Toujours dans une perspective de régulation des récurrences de signal (O3), on peut signaler les rappels d'un actant morpho-sémantiquement impliqué par un prédicat-source :

(201) J'ai été très touché d'apprendre que j'ai une abonnée à mon journal qui dessine, mais j'aimerais bien *en* recevoir (Journal de Johnny le Désossé, Genève ; *en* = (des dessins). [Reichler-]Béguelin 1993a : 359)

L'emploi du pronom permet ici au scripteur de se passer de l'introduction explicite d'un référent en vue d'O4, tout en évitant l'effet de saturation lexicale que cela pourrait engendrer.

Béguelin invoque également des rendements euphémiques, au profit d'O5, liés à l'usage de pronoms indirects :

(202) Ma mère me dit qu'il ne faut pas gâter les enfants, et elle me fouette tous les matins, quand elle n'a pas le temps le matin, c'est pour midi, rarement plus tard que quatre heures. Mlle Balandreau m'y met du suif. (J. Vallès, *L'enfant*, < Béguelin 1997a : 109)

Le locuteur doit ici conjecturer quelle partie du corps, manifestement, a souffert des coups de fouets. La nature implicite du procédé à l'œuvre autorise d'ailleurs le locuteur à démentir

l'interprétation à caractère tabou induite par le scénario évoqué, comme ci-dessous, par l'intermédiaire d'une parenthèse<sup>165</sup> :

(203) En gros – je vous la fais courte – l'homme marié avait un petit créneau, juste là, enfin deux heures à tirer, quoi..., et il se demandait si par tous les hasards du monde, ce trou dans son emploi du temps ne serait pas compatible avec les siens à elle. Bon, le hic ce n'était pas tant une question de morale : Garance *en* avait tellement avalé (de coulevres) avec lui qu'elle n'en était plus à une humiliation près, le hic, enfin les deux hics, c'est que là maintenant, elle avait plutôt envie de discuter de ses doutes existentiels que de baiser [...] (Gavalda, *L'échappée belle*, p. 113)

Dans l'exemple suivant, on relève également la présence d'une parenthèse après l'occurrence du pronom à fonction euphémique, cette fois pour faciliter la tâche interprétative du lecteur :

(204) « [...] Il prévoyait ce qui allait se passer. C'est une affaire qui vient de Naples. *Ils* (la mafia) l'ont descendu parce qu'il savait quelque chose [...] » (*sic*, *La Suisse*, 18.01.1990, Béguelin 1997a : 108)

L'emploi d'un pointeur pronominal révèle une stratégie de sauvegarde de la face par opposition aux risques encourus par une désignation plus explicite. A noter que la parenthèse (*la mafia*) a également pour rôle de distinguer deux situations d'énonciation (Béguelin 1997b). Les phénomènes d'hétérogénéité sont justement un lieu propice à l'occurrence de pronoms indirects :

(205) Et puis, ce soir-là, vers 21h, sortant du cinéma du Grutli, elle cheminait le long de l'avenue du Mail, songeant au voyage qu'elle projetait de s'offrir en Egypte. Elle avait mis son sac en bandoulière, bien serré sous son manteau. Elle tenait fermement son parapluie plié sous son bras.  
Soudain, elle se sentit poussée à terre.  
- Je suis tombée sur le coude. J'ai vu les étoiles. *Il* m'a pris mon sac. *Il* a dû l'arracher parce que je le tenais bien. (*Tribune de Genève*, 23.1.90, [Reichler-]Béguelin 1993a : 373)

Face à des faits de discours rapporté, l'interprète est naturellement amené à faire des conjectures sur le déroulement de l'interaction première et à distinguer de la sorte deux univers de croyance. L'emploi du pronom indirect en (205) dans le discours cité contribue ainsi à marquer la discontinuité entre les deux situations d'énonciation imbriquées, et partant, à signaler le changement de point de vue adopté par la locutrice citée. Cette stratégie s'oppose à une stratégie de type « intégrative » favorisant la continuité des énonciations en présence (*ibid.*).

---

<sup>165</sup> Noter également ici le clitique objet *la* dans « l'aphorisme lexicalisé » *la faire courte*. Voir à cet égard Ch.III §5.2.2.3.

Un effet de point de vue est également en jeu ci-dessous, auquel il faut ajouter l'intervention d'O4, surdéterminant l'opération référentielle par l'implication d'une nouvelle étiquette lexicale, traduisant ici la rectification implicite d'une appellation jugée inadéquate :

- (206) [le gardien de prison arrive avec le repas]  
Premier prisonnier :- Qu'est-ce que c'est ?  
Gardien : - Le potage du chef aux vermicelles.  
[les hommes commencent à manger]  
Second prisonnier : - *Elle* n'est pas mangeable (*elle* = la soupe) (script du film *Le trou* < Rosenberg 1970, cité par Cornish à par.)

Dans ce dialogue, le second prisonnier manifeste une rupture de point de vue sur le référent concerné : il considère en effet la dénomination *potage* trop valorisante pour désigner le référent en question. Celui-ci se voit ainsi implicitement « rétrogradé » par une étiquette de registre plus commun, voire péjorative (Cornish à par.).

Mentionnons une dernière stratégie impliquant également un changement de dénomination sous-jacente, qui consiste cette fois à prévenir des ambiguïtés, au bénéfice de O1 (pour rappel, l'optimisation du décodage) et de O4 :

- (207) J'orchestre mes défilés moi-même pour être sûr d'être bien compris. Cette relation avec les mannequins est très importante. Tout à coup, *elles* deviennent mes interprètes, *elles* deviennent Montana, elles sont extraordinaires. (presse < Béguelin 1997a : 107)

Dans (207), le scénario évoqué contient des actants humains désignés via une dénomination masculine sémantiquement non marquée. L'occurrence du pronom indirect au féminin suppose un changement d'appellation, à propos duquel on peut inférer l'allusion au sexe des modèles. Outre cet apport informationnel, le changement de genre permet ainsi de prévenir l'hypothèse que le couturier produirait également des vêtements pour hommes.

## 5. Conclusion

Contrairement aux idées reçues déjà évoquées *supra*, on peut donc bien relever l'opérationnalité des pronoms en site d'anaphore indirecte. Cet aperçu des rendements discursifs montre que le phénomène dépasse largement le pur respect des contraintes sémantiques ou cognitives supposées encodées dans le signifié des pronoms conjoints de 3<sup>e</sup> personne. Il montre aussi une part de la variété des routines inférentielles à l'œuvre, souvent restreintes, en matière d'anaphore associative, à une opération de déduction fondée sur un savoir a priori (lexical ou stéréotypique).

Parmi les opérations inférentielles mises en œuvre par les pronoms indirects, il nous paraît utile de distinguer quatre cas de figure distincts :

- Les inférences d'objet ordinaires, avec « disjonction référentielle » (Charolles 1990) proprement dite, où le pronom indirect invite à inférer l'existence d'un objet à partir d'un autre objet, comme en (195), en (204) ou en (199).
- Les cas où le pronom est unifiable avec une « variable » déjà sémantiquement impliquée (quoiqu'à ce moment-là sous-déterminée), comme en (201).
- Les cas où le pronom signale un changement d'attribut de dénomination, sans disjonction référentielle (*i.e.* il maintient la coréférence), comme en (206), (207).
- Les cas de *dualité* (*cf. supra* Ch.I §3.3), où l'on peut envisager un même objet sous deux formats différents (198).

Le cas des pronoms « indirects » produits dans des circonstances réelles nous est apparu comme un champ d'investigation pertinent, car il révèle l'existence de facteurs et stratégies récurrents motivant le choix du marqueur, qui passent souvent inaperçus dans les approches sémantiques ou cognitives qui se concentrent sur la mise au jour de contraintes d'emploi.



## **Chapitre III**

### **La sous-détermination référentielle**



## 1. Introduction

Une idée communément partagée par les sujets parlants est que la langue sert à exprimer la pensée de manière claire et univoque. Tout flou sémantique ou toute interprétation ambiguë se voient dès lors jugés inopportuns et à éviter autant que possible. D'ailleurs, cette recherche d'univocité du sens semble relever d'une sorte de réflexe, aussi bien chez les usagers que chez les linguistes :

L'univocité est trop profondément et trop implicitement postulée, vécue comme une évidence pour qu'il soit senti nécessaire de l'énoncer, pour que la foi qu'on a en elle soit pleinement consciente. A cet égard la linguistique (c'est-à-dire ici la quasi-totalité des linguistiques, toutes écoles confondues) participe de cette conviction des sujets parlants. (Le Goffic 1982 : 84)

Prémisse inconsciente chez certains, le postulat de clarté n'est toutefois pas qu'implicite : il est régulièrement revendiqué par les philosophes, à l'instar de Descartes ou de Leibniz, qui rêvent d'une langue universelle idéalement univoque (Le Goffic 1981). Il s'ensuit que l'on voit régulièrement émerger diverses tentatives de remédiation au flou ou à l'ambiguïté, telles que les propositions de *définitions de noms* (lat. *definitio nominis*) d'Arnauld & Nicole (1662=1874 : 87 *sqq.*), le système de notation symbolique (*Begriffsschrift*) instauré par Frege (1971) ou encore le développement de structures sous-jacentes en grammaire générative, où les transformations sont responsables d'« effets de dégénérescence » par disparition d'informations en surface (Harris 1976 : 57 cité par Fuchs & Le Goffic 1983).

Cette vision de la langue comme vouée à l'expression d'idées claires et univoques est pour le moins simpliste, ainsi que le regrette Lyons (1981 : 203) :

Not only is it frequently, and erroneously, associated with the view that all sentences have precise and determinate meanings ; it is based on the equally erroneous assumption that clarity and the avoidance of vagueness and equivocation are always desirable, regardless of what language game we are playing.

En effet, l'observation des faits langagiers montre que les phénomènes d'ambiguïté et de sous-détermination sont parties intégrantes de la langue et du discours, comme en témoignent les travaux de bon nombre de chercheurs (entre autres Crystal & Davy 1975, Le Goffic 1981, 1982, Fuchs & Le Goffic 1983, Fuchs 1996, Channell 1994, Jucker *et al.* 2003, Kaltenböck *et al.* 2010, Bat-Zeev Shyldkrot *et al.* 2010, 2014, 2016).

Nous proposons donc d'approfondir la question du « flou » qui peut concerner la référence aux objets-de-discours, susceptible de poser problème aux théories de la référence existantes. Après une caractérisation de la notion de sous-détermination et une délimitation de l'étendue des faits qui nous intéressent (§2), nous prenons soin de la distinguer de phénomènes apparentés (§3), puis nous en dégageons les circonstances et modalités d'apparition et évoquons quelques hypothèses sur son traitement cognitif (§4). Enfin, nous présentons un inventaire des principales ressources linguistiques à la disposition des locuteurs pour exprimer la sous-détermination en français (§5).



## 2. Caractéristiques de la sous-détermination

En matière de référence, les approches traditionnelles supposent que l'identification des référents est l'un des critères d'une communication réussie. Néanmoins, les échanges conversationnels ordinaires montrent plutôt que les locuteurs manient le degré de détermination référentielle en fonction de leurs objectifs discursifs. Ainsi, la manipulation de la granularité référentielle permet notamment de gérer le degré de saillance des objets, contribuant à la dynamique des représentations en M, laissant certains objets en retrait et en plaçant d'autres au centre de l'attention (Jucker *et al.* 2003). Elle permet en outre, comme on le verra, de créer des effets discursifs particuliers.

### 2.1 *Sous-détermination et pertinence*

Si dans certaines situations de sous-détermination, le locuteur ne dispose tout simplement pas d'informations précises à propos d'un référent évoqué (par simple ignorance, oubli, etc.), on peut supposer que dans d'autres situations, il dispose de connaissances plus précises mais qu'il renonce délibérément à les exploiter, pour des raisons de pertinence (Sperber & Wilson 1986) : en effet, il peut estimer les coûts d'une telle détermination trop élevés, que ce soit à l'encodage ou au décodage présumé, par rapport aux effets communicatifs escomptés. Inversement, l'interprète calcule de son côté, sur le même principe, uniquement un ensemble d'inférences rentables, renonçant à d'autres qui seraient pourtant possibles. C'est donc surtout sur un objectif de pertinence, plutôt que de vérité, que repose la variation du degré de détermination référentielle.

### 2.2 *Une notion scalaire*

La sous-détermination s'oppose à une situation de détermination référentielle, c'est-à-dire dans laquelle les objets évoqués sont identifiés et clairement opposables les uns par rapport aux autres. On parlera de sous-détermination lorsque surgissent des difficultés d'attribuer une valeur à une variable. Cela peut s'expliquer, au niveau des référents, par le manque d'attributs distinctifs (en particulier une dénomination, des propriétés typantes et discrètes). En parlant de sous-détermination, nous mettons en avant le caractère scalaire de la notion. S'il nous semble intuitif d'envisager une situation de détermination « complète », *i.e.* lorsqu'il est possible d'attribuer une valeur à un objet, il nous semble à l'inverse plus problématique de concevoir un objet totalement indéterminé. En effet, tout acte sémiotique apportant au moins une information minimale sur l'objet auquel il renvoie dans M, nous jugeons plus adéquat de recourir à une notion en adéquation avec ce point ; dans le cas d'un acte énonciatif, l'expression au moyen de laquelle le référent est introduit de même que la prédication dont il fait l'objet fournissent de fait des indications à son propos : elles le déterminent forcément d'une certaine manière, aussi minimale soit-elle. C'est pourquoi nous recourons à la notion de *sous-détermination* plutôt qu'à celle d'*indétermination*.

Il faut rappeler qu'en tant qu'analyste, nous n'avons accès qu'à des traces du discours effectivement produit par des interactants : il nous est donc impossible de mesurer fiablement le degré de sous-détermination des objets en M. Nous ne pouvons de ce fait que nous appuyer sur les indications partielles à disposition, comme les indices verbaux, les inférences qu'on peut en tirer et parfois les métadonnées disponibles.

### 2.3 Types de sous-détermination

Dans cette étude, nous nous intéressons exclusivement à la sous-détermination susceptible d'intervenir dans le processus d'unification d'une variable avec un objet-de-discours. De fait, nous laissons de côté les procédés de sous-détermination se situant à d'autres niveaux, comme ceux se rapportant à des quantités, comme illustrés ci-dessous :

(208) voilà alors j'aimerais te poser *quelques* questions sur ton | \_ | métier d'enseignant (ofrom)

(209) et moi j'ai travaillé dans le deuxième étage | \_ | qui dans lequel y avait euh trente-trois résidents | \_ | et puis on était *environ* huit aides-soignantes en temps normal (ofrom)

Dans l'exemple (208), le quantificateur *quelques* exprime la sous-détermination du nombre de questions à poser (on sait qu'il y en a plus qu'une, mais qu'elles sont jugées peu nombreuses). Le marqueur *environ* en (209) fonctionne quant à lui comme un approximateur (cf. *infra* §3.2) à l'égard du nombre d'aides-soignantes exprimée, conduisant d'emblée à en relativiser la précision.

De même, nous ne nous intéressons pas à la sous-détermination affectant les propriétés, comme ci-dessous :

(210) ouais pis c'est pas que je préfère lire que | \_ | regarder un film mais pour les films qui sont *assez* compliqués et pis euh | \_ | *assez* durs *comme ça* je trouve que c'est mieux un livre parce que sinon après c'est *un peu* | \_ | ça c'est des fois *un peu* violent *ou comme ça* (ofrom)

Les expressions *assez* et *un peu* indiquent ici aussi la nécessité de relativiser les qualités respectives évoquées. Quant aux occurrences de (*ou*) *comme ça*, c'est plutôt sur leur adéquation lexicale que semble porter la fonction mitigatrice (Béguelin & Corminboeuf à par.)

D'autre part, certaines expressions sont spécifiquement dédiées à l'introduction d'une variable sous-déterminée ou quelconque. C'est le cas d'un certain nombre de pronoms dits « indéfinis » (cf. *infra* §3.1) :

(211) y a aussi l'habitude parce que *quelqu'un* qui travaille d'une semaine à l'autre | \_ | de temps en temps chez lui | \_ | ou bien *quelqu'un* qui en fiche pas une et pis qui qui fait | \_ | que son travail à à la leçon ça s'entend parce qu'en général on retrouve un petit peu l'instrument | aux mêmes endroits qu'où on l'a laissé avant (ofrom)

Le pronom *quelqu'un* (mais aussi *quiconque*, *quelque chose*, *n'importe qui/quoi*, *quoi que ce soit*, *je ne sais qui/quoi*, *certains*, *d'aucuns*, etc.) manifeste à travers son signifié même la nature sous-déterminée de la variable introduite et l'inutilité de saturer celle-ci.

Bien que ces expressions soient au cœur du domaine de la sous-détermination, nous avons choisi de nous pencher uniquement dans la présente étude sur le cas des *pointeurs référentiels* (ou *désignateurs*), car ils ont ceci d'intéressant qu'ils donnent lieu à une forme de conflit entre l'instruction d'unification véhiculée et l'opération même d'unification.

Avant de distinguer les différents cas de figure de sous-détermination, précisons ce que nous entendons par « anaphore ». L'opération anaphorique canonique correspond à un type d'unification non problématique : un objet déterminé est introduit à un moment donné en M, puis une énonciation ultérieure instancie une variable à unifier avec celui-ci (ce processus reposant, comme on l'a vu, sur un calcul inférentiel à partir d'indices de divers ordres). En dehors de cette situation, on peut distinguer plusieurs types de dissension entre l'instruction d'unification de la variable et sa saturation.

Berrendonner (2016) recourt à la notion de *variables provisoires* pour désigner « une classe particulière de référents sous-spécifiés, qui ne satisfont pas à l'exigence conventionnelle de pertinence optimale, et dont l'introduction dans M crée par conséquent un état non stable du savoir partagé ». Pour le cas des pointeurs référentiels, on peut en relever un premier type, traditionnellement appelé « cataphore » :

(i) L'unification de la variable sous-spécifiée est différée :

(212) Vous *l'*avez tous remarqué... Comment d'ailleurs ne pas *le* remarquer ? Dans certaines zones d'Épinay-sur-Seine, nos rues, nos trottoirs et nos espaces publics sont occupés par d'importants chantiers. (circulaire communale, Épinay-sur-Seine, mai 2011)

Dans cet exemple, la variable sous-déterminée est introduite à travers deux occurrences pronominales, avant qu'une énonciation ultérieure vienne l'instancier. Le procédé de cataphore crée ainsi un état instable de M, perçu comme un effet stylistique d'attente (*ibid.*) Dans ce cas, l'instanciation de la variable provisoire est simplement différée.

Il existe un autre type de variable provisoire :

ii) L'objet faisant défaut pour l'identification demande à être inféré, mais le calcul à partir des indices contextuels « n'aboutit généralement qu'à cerner une classe floue de valeurs possibles, plus ou moins vraisemblables » (*ibid.*) :

(213) après on débarrasse après nous on a une pause de vingt minutes on fait la | pause café  
*tout ça* | \_ | après on se déba/ on arrive on débarrasse d/ les déjeuners (ofrom)

Dans ce cas, il est possible d'inférer d'autres objets en relation avec 'faire la pause café', par exemple {... discuter, fumer une cigarette, aller aux wc, faire des sudokus, etc.}. Cependant, dans le procédé de référence à l'œuvre ci-dessus (qui fera l'objet du chapitre IV *infra*), il y a de bonnes raisons de penser que l'inférence des valeurs potentielles par l'interprète soit court-circuitée par l'application du principe de pertinence : autant le locuteur, en utilisant un *tout ça* sur le mode allusif, juge trop coûteux de détailler davantage le référent en vue de ses objectifs, autant l'interlocuteur estime le traitement inférentiel cognitivement peu rentable. En fait, *tout ça* évoque ici un « faisceau d'objet » au sens de Grize (1990 : 78), c'est-à-dire un ensemble d'éléments hétérogènes (propriétés, relations et schèmes d'action) attachés à un objet donné (ici le référent '*pause café*'). Grize considère que toute introduction d'objet ouvre simultanément une *classe-objet*, autrement dit les éléments hétérogènes qui lui sont attachés et qui ont « affaire » les uns avec les autres. A cet égard, il faut noter que cette notion de *classe-objet* se distingue radicalement de la conception mathématique de « classe » impliquant des membres homogènes. Quoi qu'il en soit, la référence de *tout ça*, tout en s'apparentant allusivement à une classe-objet, demeure ici relativement sous-déterminée.

On peut enfin relever un dernier type de difficulté surgissant dans le processus d'unification d'une variable, qui se distingue des deux cas précédents par le fait qu'il n'implique pas de variable provisoire, du moins dans le même sens que précédemment :

iii) L'objet supposé manifeste se montre en fait présent dans M au moment du pointage, mais l'expression référentielle, typiquement une expression dite « résomptive », renvoie à une partie mal délimitée du réseau référentiel en construction :

(214) Bon moi j'ai travaillé mais vous me direz euh. Dans des conditions euh agréables euh j'avais mes horaires je me rapportais du travail à la maison. Tout ça était tellement souple que j'étais quand même beaucoup avec les enfants. Mais euh je sais pas si ils s'*en* sont tellement aperçus hein. (pfc)

Le clitique oblique *en*<sup>166</sup> renvoie bien à un objet préalablement introduit dans M, mais cet objet se montre composite, lui-même constitué d'individus hétérogènes et de relations entre eux, dont certain(e)s sont à inférer (*cf.* ici aussi la notion de classe-objet) : il renvoie ainsi à un sous-graphe de M, dont pour rappel, l'architecture peut être représentée sur le mode « gravitationnel » (*cf. supra* Ch.I §4.6.3 et §6.3.4). En tenant compte des indications sémantiques véhiculées par le verbe recteur (*s'apercevoir de qqc*), on peut interpréter *en* non seulement comme le temps passé par la locutrice avec ses enfants, mais peut-être aussi comme la flexibilité de travail dont elle bénéficiait, le privilège de sa situation, etc. Dans ce cas aussi, une part d'incertitude demeure, en particulier au niveau de la composition et des limites du sous-graphe concerné. Comme dans le cas précédent, tout porte à croire que le référent est traité en gros et que le calcul de la teneur détaillée ou des frontières de l'objet n'est pas jugé pertinent.

<sup>166</sup> On pourrait faire une analyse similaire pour le *tout ça* aussi présent.

Dans la suite de ce travail, nous nous concentrerons sur ces trois types de sous-détermination, pour rappel : i) le cas de mise en suspens de l'instanciation d'une variable ; (ii) le cas où l'objet avec lequel unifier la variable n'est pas présent et dont l'inférence, du fait de son incertitude, laisse l'objet sous-déterminé ; (iii) le cas où l'objet est manifestement présent, mais où la sous-détermination porte davantage sur sa composition et sa délimitation.



### 3. Notions voisines

Comme on a pu s'en apercevoir en restreignant le domaine d'observation, certaines notions se rapprochent de celle de sous-détermination, sans toutefois que leurs extensions respectives se recouvrent. C'est notamment le cas des notions de *référence indéfinie* (§3.1), d'*approximation*<sup>167</sup> (§3.2), d'*ambiguïté* (§3.3) et autres phénomènes apparentés à celle-ci comme l'*ambivalence* (§3.4) ou encore l'opposition *transparence vs opacité* (§3.5). Dans un sens courant, toutes évoquent en effet quelque chose de flou ou d'imprécis et s'opposent ensemble à l'idée d'univocité du langage. Il nous paraît dès lors important de nous situer plus précisément par rapport à ces notions.

#### 3.1 Référence indéfinie

La notion de *référence indéfinie* se rapproche de celle d'*indétermination référentielle*, en ce qu'elles renferment toutes deux l'idée qu'on ne peut associer une valeur stable à une variable. Sur ce point, les notions se recouvrent. Mais la notion d'indéfini est en outre exploitée pour cerner des catégories grammaticales traditionnelles, à savoir des « articles » (*un/e, des*) et autres déterminants (*quelque/s, plusieurs, aucun, chaque, nul, etc.*), ainsi que des pronoms du type *quelqu'un, quelque chose, chacun, personne, rien, on, etc.* A ces derniers, on peut encore ajouter d'autres classes dites *indéfinies* comme les adverbes *quelque part, nulle part, jamais, etc.* (Haspelmath 1997). Du fait de cette hétérogénéité, la classe des indéfinis a souvent été cataloguée de « fourre-tout » (entre autres Arrivé, Gadet & Galmiche 1984, Flaux & De Mulder 1997, Schnedecker *à par.*).

Notre enjeu étant d'étudier les pointeurs dont on ne parvient à unifier la variable malgré l'instruction qu'ils comportent (*cf. supra* §2.3), la notion d'indéfini est à différencier de notre objet d'étude. En effet, la plupart des formes dites indéfinies introduisent, au contraire, explicitement une variable sous-déterminée, sans instruction de saturation. Néanmoins, on verra que l'usage de certaines expressions étudiées dans les parties empiriques de ce travail<sup>168</sup> se rapproche à certains égards de l'emploi de l'une ou l'autre forme dite indéfinie. Dans ce cas, nous n'hésiterons évidemment pas à examiner de plus près leurs rapports (*cf. infra* Ch.V §3.4.2).

#### 3.2 Approximation

D'après le *TLFi*, une approximation est, au sens mathématique, une « opération par laquelle on tend à se rapprocher de plus en plus de la valeur réelle d'une quantité ou d'une grandeur sans y parvenir rigoureusement. » Dans la langue courante, elle représente une « évaluation approximative<sup>169</sup> d'un nombre, d'un chiffre, d'une distance, d'une date, etc. » et par extension de sens, simplement le « caractère d'une chose lorsqu'elle n'offre qu'une exactitude relative ».

---

<sup>167</sup> *Cf. infra* (§4.5.3) l'utilisation d'*et tout* et *tout ça* comme marqueurs d'approximation de paroles rapportées.

<sup>168</sup> En particulier l'emploi de *tout ça* ou du clitique sujet *ils* à valeur sous-déterminée.

<sup>169</sup> Noter la circularité de la définition.

Selon Bat-Zeev Shyldkrot *et al.* (2010 : 3), le point commun de ces significations est « l'approche évaluative ou imprécise et l'écart par rapport à une exactitude rigoureuse », comme l'illustrent les propos de Valéry ci-dessous :

(215) C'est faute de le savoir [le langage des dieux] que l'homme ou que l'être de l'homme, a créé ces *approximations* : les larmes, le sourire, le soupir, l'expression du regard, le baiser, l'embrassement, l'illumination du visage, le chant spontané, la danse... (Valéry, *Mauvaises pensées et autres*, 1942 < TLFi)

Les énoncés contenant des marqueurs d'approximation posent des problèmes aux logiciens, qui ne parviennent pas à leur attribuer une valeur de vérité dans le système binaire vrai-faux :

(216) Sam is *approximately* six feet tall (Sadock 1977 : 434).

Sadock (1977) évoque la solution consistant à recourir à la théorie des ensembles flous<sup>170</sup>, où la notion de vérité est appréhendée en termes scalaires, le but étant de déterminer par des algorithmes jusqu'à quelle distance l'estimation garantit la vérité de l'énoncé. L'auteur reste néanmoins sceptique envers cette approche vériconditionnelle<sup>171</sup> et privilégie une analyse pragmatique selon laquelle le rôle d'un marqueur d'approximation est simplement de relativiser le sens d'un énoncé :

the role of an approximator in the pragmatic theory is to trivialize the semantics of a sentence, to make it almost unfalsifiable, to hedge in a genuine sense. (*ibid.* : 437)

Autrement dit, il attribue une fonction mitigatrice à ce type de marqueur : une approximation ne s'évalue pas en termes de valeur de vérité car elle est d'emblée présentée comme « a purposely, and unabashedly inaccurate statement » (*ibid.* : 434). Il traite l'emploi d'approximateurs via la maxime conversationnelle de quantité (Grice 1975) enjoignant de n'en dire ni plus ni moins que requis.

Dans une approche pragmatico-interactionnelle, les marqueurs d'approximation ont ainsi été considérés comme réduisant l'engagement du locuteur à l'égard de son propos, lui permettant par la même occasion de ménager sa face dans l'échange. Les marqueurs d'approximation peuvent se répartir en deux classes (Vaguer 2010) : ceux qui s'appliquent à une quantité, la présentant comme non extensive (*environ, à peu près, presque, etc.*) (*cf. supra* l'exemple (209)) et ceux qui portent sur une dénomination (*plutôt, sorte de, genre de, espèce de, plus*

---

<sup>170</sup> Cf. Zadeh (1975: ix) : « We have been slow in coming to the realisation that much, perhaps most, of human cognition and interaction with the outside world involves constructs which are not sets in the classical sense, but rather 'fuzzy sets' (or subsets), that is, classes with unsharp boundaries in which the transition from membership to non-membership is gradual rather than abrupt. Indeed, it may be argued that the logic of human reasoning is not the classical two-valued or even multivalued logic but a logic with fuzzy truths, fuzzy connectives, and fuzzy rules of inference. » (cité par Channell 1994 : 200)

<sup>171</sup> Le traitement de l'approximation en termes de vérité/fausseté apparaît comme un réflexe, comme en témoigne ce commentaire de Guilbaud (1977 : 126) évaluant l'approximation sous l'angle du mensonge : « [...] An approximate value, as common sense says, is that which is not exact. Is it a lie then? I would not deny that newspapers sometimes contain lies. But there is not always so much malice. No! Talking and thinking by means of 'about', 'nearly' is a necessity ».

*exactement*, etc.), relativisant son adéquation. Au vu de (210), nous ajouterions une troisième classe de marqueurs, qui porte sur les propriétés.

Les marqueurs d'approximation ont été rapprochés des marqueurs de modalité épistémique (Le Querler 1996 : 74 *sqq.*) tels que *peut-être*, *paraît-il*, *je pense*, etc. qui réduisent le degré de certitude à l'égard d'un contenu asserté :

(217) ben voilà ben avec lui ben j'ai j'ai dû partir vraiment à zéro il savait pas tenir ses baguettes il avait pas de batterie *je crois* même | \_ | à la maison (ofrom)

En modalisant son énoncé par l'usage d'un verbe parenthétique, le locuteur en réduit la force illocutoire et par la même occasion, les risques d'objection(s) qu'il encourt contre son propos.

Les deux types d'expression (marqueurs modaux et marqueurs d'approximation) relèvent ainsi d'une stratégie de *mitigation* ou d'atténuation (Caffi 2007), au niveau du contenu d'un énoncé, au niveau de la relation entre un locuteur et son énonciation et au niveau interpersonnel. On voit par là que les faits qui relèvent de l'approximation (ou la modalité épistémique) ne recouvrent pas, a priori<sup>172</sup>, ceux que nous avons délimités pour notre étude de la sous-détermination, qui concernent les objets-de-discours (et non leur quantité ou leur attribut de dénomination). Cela se traduit notamment par l'examen de classes de marqueurs différentes : tandis que l'approximation est essentiellement appréhendée à travers le cas de modificateurs ou de quantificateurs, nous nous intéressons principalement à des marqueurs nominaux ou pronominaux.

### 3.3 *Ambiguïté*

Dans l'usage courant, on recourt fréquemment à la notion d'ambiguïté pour décrire une situation discursive confuse, vague, dont l'interprétation est incertaine (Fuchs 1996 : 13). Or, en linguistique, on s'efforce de mettre sous cette notion un contenu plus fonctionnel :

L'ambiguïté linguistique se caractérise comme une alternative entre plusieurs significations mutuellement exclusives associées à une même forme, au sein du système de la langue. (*ibid.*)

A partir de cette définition, Fuchs distingue l'ambiguïté virtuelle, où le contexte linguistique sélectionne l'interprétation adéquate parmi les significations potentielles concurrentes, de l'ambiguïté effective, où le destinataire est véritablement confronté à un choix insoluble. Ainsi, dans les énoncés suivants, le premier sera considéré comme un cas d'ambiguïté virtuelle, résolu par le contexte linguistique, tandis que dans le second, l'interprétation demeure bloquée jusqu'à la fin de la phrase :

(218) Les croque-morts déposèrent la bière dans le corbillard (*ibid.* : 12)

---

<sup>172</sup> Néanmoins, les deux phénomènes peuvent être à l'œuvre en même temps, cf. l'emploi *infra* (Ch.IV §4.5.3) de *tout ça* fonctionnant comme marqueur d'approximation de paroles rapportées tout en introduisant une représentation sous-déterminée des propos implicites.

(219) Comme il faisait très chaud le jour de l'enterrement, on sortit la bière (*ibid.*)

Selon Fuchs, le nom *bière*, dans l'exemple (218) s'interprète hors contexte avec la signification de 'cercueil', tandis que dans (219), il n'est pas possible de trancher entre ce même usage ou celui du substantif homonyme signifiant un type de boisson alcoolisée. Cette distinction est toutefois discutable au vu de ces exemples, la lecture proposée pour (218) n'excluant pas la possibilité de l'autre interprétation. A l'inverse, en (219), on peut rétorquer que la structure argumentative oriente vers une lecture préférentielle (la relation de cause à effet qu'on peut dégager entre la chaleur et le fait de boire de la bière). En fait, Channell (1994) relève à juste titre qu'en discours, les situations d'ambiguïté effective ne sont pas courantes : les participants d'une conversation se rendent rarement compte que plusieurs interprétations sont possibles. En effet, le contexte au sens large, *i.e.* linguistique – y compris la prosodie – situationnel, le savoir encyclopédique ou les inférences qu'il est possible de tirer, oriente généralement vers l'une des interprétations. D'ailleurs, Fuchs ramène l'ambiguïté à un phénomène propre au système de la langue, tandis qu'elle place la sous-détermination au niveau du discours (*ibid.* : 23).

Dans les deux cas néanmoins, on rencontre une situation d'« indétermination » à un moment donné : dans le cas de l'ambiguïté effective en discours, l'indétermination vient de l'impossibilité de choisir entre plusieurs significations elles-mêmes déterminées. Dans les cas que nous étudions, l'indétermination se situe dans la difficulté d'attribuer une valeur référentielle à la variable introduite. Fuchs (1996 : 24) recourt à une comparaison parlante pour distinguer les deux cas de figure : elle rapproche l'ambiguïté d'images susceptibles d'être perçues de plusieurs manières, mutuellement exclusives, telle la fameuse illusion d'optique amenant à reconnaître tantôt le portrait d'une vieille femme, tantôt celui d'une jeune fille. Quant à la sous-détermination, elle est illustrée par le cas d'images floues représentant une seule situation, mais aux contours indistincts. On pourrait être tenté de prolonger l'analogie de Fuchs en remarquant que l'image est généralement réalisée à dessein dans le premier cas, mais involontairement dans le second cas. Cependant, Fuchs relève nombre d'ambiguïtés « non contrôlées »<sup>173</sup> ; à l'inverse, la référence à une situation floue peut être délibérée et répondre à certains besoins, comme le remarque Wittgenstein (et comme on le verra d'ailleurs dans la suite de ce travail) :

Ja, kann man ein unscharfes Bild immer mit Vorteil durch ein scharfes ersetzen? Ist das unscharfe nicht oft gerade das, was wir brauchen?  
(Wittgenstein 1953 = 2011 : §71)

---

<sup>173</sup> Fuchs illustre cela par le titre « Marche silencieuse contre la violence à Bastia » (*Le Monde* < Fuchs 1996 : 75), où le journaliste n'est probablement pas conscient de l'ambiguïté virtuelle créée (qu'est-ce qui, de la marche ou de la violence a eu lieu à Bastia ?). Néanmoins, en contexte, cet énoncé n'est sans doute pas ambigu pour l'auteur (ni d'ailleurs pour le récepteur). On peut donc se demander si cela fait sens de parler d'ambiguïté non contrôlée en discours.

Nous nous relativiserons cependant dans la suite de ce travail l'idée de « substitution » sous-jacente évoquée par Wittgenstein, une représentation sous-déterminée n'étant pas (toujours) à considérer comme un substitut d'une représentation déterminée.

Mentionnons pour finir encore deux acceptions du terme « ambiguïté » relevées par Le Goffic (1983 : 77) : celle d'« incomplétude de la signification », notamment adoptée en analyse littéraire, en psychanalyse ou en sémiotique, où « le potentiel de signification d'un énoncé ou d'un texte n'est jamais épuisé ». Ainsi, une séquence donnée renfermerait un nombre d'interprétations infini. L'auteur évoque enfin une acception de l'ambiguïté située au niveau de l'échange discursif, appréhendée comme le décalage d'encodage et de décodage entre le locuteur et l'interlocuteur : on ne peut jamais être sûr qu'il y ait coïncidence entre le sens visé par l'un et celui interprété par l'autre. Dans ce cas, nous ne parlerons pas d'ambiguïté mais de méprise ou de dissension sur le contenu de M.

En somme, nous pouvons retenir de cette mise au point que l'ambiguïté en linguistique se distingue de la sous-détermination telle que nous l'entendons par le fait qu'elle se situe au niveau sémantique, et non au niveau référentiel.

### 3.4 *Ambivalence*

Le Goffic (1982) distingue l'ambivalence de l'ambiguïté pour les cas qui, tout en présentant une pluralité de significations, se prêtent mal à l'exclusivité stricte d'une interprétation au détriment de l'autre. Dans le cas de l'ambiguïté, l'activation d'une signification annule la possibilité d'activer l'autre (les autres) signification(s). Selon Le Goffic, dans certaines situations, il est pourtant difficile de distinguer entre deux interprétations exclusives. Il recourt alors à la notion d'ambivalence pour décrire des cas où des significations *a priori* distinctes sont simultanément en vigueur. Issu de la psychanalyse, le concept désigne « la présence simultanée de deux composantes de sens contraires » (p. 86), dont le domaine de prédilection est celui des sentiments.

Exploitée au niveau sémantique, la notion permet de faire coexister plusieurs sens en principe contradictoires, soit en les activant chacun dans leur totalité, soit en en faisant un produit mixte ou intermédiaire. L'exemple des syllepses peut servir à illustrer cela, comme dans *Il n'y a pas de différence entre un coiffeur et un peintre, ils peignent tous les deux.* (p. 93) : l'occurrence *peignent*, du fait de l'homonymie qu'elle manifeste, est susceptible d'activer simultanément le sens des verbes *peindre* et *peigner*. Le Goffic évoque encore l'exemple d'unités lexicales à l'origine homonymes dont les sens ont fusionné (le verbe *errer* < lat. *errare* et < lat. *iterare* (p. 94) ou enfin le cas de l'adverbe *encore*, dont les valeurs durative et répétitive peuvent se cumuler (*j'écris encore deux pages et je te rejoins*) (p. 97). Autrement dit, le terme d'ambivalence est employé dans des situations de plurivocité où le principe d'alternative ne s'avère pas adéquat. Fuchs (1996) parle d'ailleurs dans ces cas « d'indétermination » ou de « plurivocité sans ambiguïté » (p. 31), phénomène qu'elle décrit ainsi :

Plusieurs significations sont données à comprendre comme co-possibles ; [...] le contexte conduit à arrêter l'interprétation en-deçà d'une distinction tranchée entre plusieurs significations de même niveau : ce qui est donné à comprendre se situe à mi-chemin entre plusieurs significations, participe un peu de toutes, neutralise leurs différences. (p. 30)

On peut rapprocher ici la conséquence référentielle des indices sémantiques contradictoires de l'ambivalence, qui peuvent contribuer à évoquer des objets indiscrets (Ch.I §3.3). Néanmoins, la difficulté à attribuer une valeur référentielle ne résulte pas du défaut d'attributs distinctifs comme dans les cas de la sous-détermination, mais plutôt d'un amalgame de propriétés contradictoires.

### 3.5 *Interprétation transparente vs opaque*

On peut encore relever un cas particulier d'ambiguïté, dite parfois logique, distinguant une interprétation *transparente* d'une interprétation *opaque* (Quine 1960) :

(220) Nancy veut épouser *un Norvégien*. (Fuchs & Le Goffic 1983 : 112)

Dans son interprétation *transparente*, l'énoncé dit de Nancy qu'elle veut épouser un individu qui se trouve être norvégien ; dans une lecture *opaque*, on comprend que Nancy veut devenir femme-de-Norvégien. Ces deux significations donnent lieu, selon les cadres théoriques respectifs, à des formes logiques distinctes, de même qu'à des structures profondes différentes. On explique généralement la différence d'interprétation par le fait que le SN *un Norvégien* peut être attribué à des énonciateurs distincts, à savoir, le locuteur même de l'énoncé (1<sup>ère</sup> interprétation) ou le personnage de *Nancy* lui-même (2<sup>e</sup> interprétation). Dans le premier cas, on dit que l'expression est *transparente* puisqu'elle laisse transparaître le point de vue du locuteur (comme dans le célèbre *Œdipe veut épouser sa mère*) ; on dit qu'elle est *opaque* lorsque c'est l'individu concerné (ici Nancy) qui est responsable de l'emploi du SN (Charolles 2002 : 99).

Néanmoins, cette explication n'est pas totalement convaincante étant donné que l'ambiguïté du SN *un Norvégien* demeure lorsqu'il n'y a qu'un énonciateur, comme dans l'énoncé *je veux épouser un Norvégien*. Il faudrait plutôt rechercher la source du problème au niveau des types de verbes concernés et de leur capacité à ouvrir des champs d'interprétations variés pour les SN indéfinis du type *un N*. En effet, si l'on fait par exemple commuter le verbe *vouloir*+inf. avec le verbe *savoir*+*que P* (*Nancy sait qu'elle épouse un Norvégien*) l'ambiguïté disparaît. Malgré l'intérêt de la problématique, nous laissons aux sémanticiens le soin d'approfondir ce champ d'étude très spécifique.

## 4. Emergence et traitement de la sous-détermination

Dans cette section, nous tentons de cerner un certain nombre de facteurs déterminants dans l'émergence de la sous-détermination (§4.1), puis nous examinons le type de ressource linguistique principale dans l'expression de la sous-détermination, à savoir la sous-spécification sémantique (§4.2). Pour finir, nous présentons quelques pistes sur le traitement cognitif de ce type d'information (§4.3).

### 4.1 Motivations de la sous-détermination

On peut s'interroger sur les causes de l'introduction en M d'une représentation sous-déterminée. Crystal & Davy (1975 : 111-112) mettent au jour un certain nombre de facteurs pour expliquer le « vague lexical » dont ils constatent la fréquence dans les conversations familières et qu'ils jugent partie intégrante du genre discursif. Ils distinguent ainsi :

- a) La perte de mémoire
- b) L'absence de dénomination adéquate ou ignorance de la dénomination
- c) L'absence de pertinence de la précision étant donné le sujet de conversation
- d) Le choix délibéré d'un terme imprécis pour maintenir l'atmosphère décontractée

Il nous semble que ces paramètres intuitifs peuvent être répartis selon leur caractère accidentel (en gros, l'ignorance : a, b), ou stratégique (en lien avec la pertinence : c, d). Les facteurs proposés par les auteurs relèvent simplement d'un constat général sur l'émergence du vague lexical, mais ne sont ni présentés comme le fondement d'une typologie élaborée, ni véritablement argumentés, ni systématiquement illustrés. Il nous semble ainsi que cette liste pourrait servir de base à une typologie plus rigoureuse, couplées à quelques-uns des rendements mis au jour par Béguelin (1997a et b). Nous proposons donc d'examiner quelques exemples révélateurs d'autres aspects déterminants :

- (221) - Sous prétexte que j'avais laissé entrer quelqu'un qui ne faisait pas partie du personnel, votre ami Walter en l'occurrence, *ils* m'ont menacé de me licencier en invoquant une faute professionnelle grave.
- Mais qui ça « ils » ?
  - Ceux qui financent l'observatoire, notre gouvernement.
  - Enfin, Martyn, cette visite était parfaitement anodine, et puis Walter et moi sommes tous deux membres de l'Académie, cela n'a aucun sens ! (M. Levy, *La première nuit*, p. 239-240)

Le locuteur de la première réplique recourt au pronom sous-spécifié *ils* pour évoquer l'agent impliqué dans un procès, soit qu'il juge, à tort en l'occurrence, son identité évidente, laissant le soin à son interlocuteur de l'inférer (O2, cf. *supra* Ch.I §6.3.6), soit qu'il la juge non pertinente, prévenant d'éventuels coûts de décodage (O1). Dans tous les cas, il y a divergence entre les interlocuteurs sur le contenu courant de M, étant donné la demande d'identification

consécutive, à laquelle le locuteur satisfait, mettant un terme à l'état instable ainsi créé. Au vu du climat de crainte décrit (le rôle menaçant du référent), des inférences qu'on peut tirer à partir de l'identité révélée après coup du référent ainsi que du genre discursif, il y a de bonnes raisons de considérer que l'emploi de *ils* procède d'une stratégie d'évitement, visant à taire le nom d'un référent jugé néfaste (conformément à O5, pour rappel, la sauvegarde des faces)

Dans le même genre, l'exemple ci-dessous met en jeu la désignation d'un référent momentanément sous-déterminé, manifestement également au profit de O5 :

(222) [Quelques jours après l'annonce de la disparition imminente d'un hebdomadaire romand] Depuis quelques jours, tout le monde ne parle que de *ça* autour de moi. De la liste. La nôtre. Nous sommes comme les autres, et j'imagine que ce serait la même chose dans un autre boulot [...]. Mais piocher 37 noms, c'est tellement laid, ça doit être un pari que l'on a perdu, sans même savoir de quoi il s'agissait. (*Le Temps*, 08. février 2017, Chronique *Ma première fois*)

Ce qui motive le caractère d'abord sous-déterminé du référent désigné par *ça*, c'est son caractère tabou, qui s'explique en aval par la révélation de son identité, à savoir une liste de noms de personnes à licencier. Le pronom est visiblement employé en « deixis mémorielle<sup>174</sup> » (Fraser & Joly 1980), c'est-à-dire qu'il renvoie à un référent évident dans l'esprit du locuteur – d'où un effet potentiel d'empathie – mais non disponible en M au moment du pointage.

Dans l'exemple ci-dessous, la locutrice justifie son intérêt pour les journaux dès son plus jeune âge et désigne, dans un discours direct, un objet dont elle n'a, au moment de l'énonciation rapportée, pas connaissance :

(223) depuis toute petite tu sais il me semble que j'ai voulu m'intéresser euh | \_ | tout ce qu'y avait dans les journaux enfin je prenais le journal je le lisais | % | c'est quoi *ça* enfin (ofrom)

Ici, l'ignorance responsable de la sous-détermination ne se situe pas forcément au niveau lexical, seul niveau évoqué par Crystal & Davy (1975), mais probablement au niveau encyclopédique : c'est parce que la locutrice ne connaît pas, au moment du discours mimé, tel ou tel sujet, concept, fait, etc., évoqué dans les journaux qu'elle utilise les pointeurs *c'* et *ça* pour le désigner. D'ailleurs, l'emploi de ces désignateurs token-réflexifs surmarquent ici la discontinuité entre le discours citant et cité (Béguelin 1997b), permettant au destinataire d'adopter le point de vue de l'enfant qu'était la locutrice au moment des faits rapportés.

La sous-détermination en jeu dans l'exemple suivant peut s'expliquer par l'absence d'une catégorisation adéquate existante pour le référent mais aussi par nonchalance (O2) ou non pertinence d'une étiquette plus spécifiée (O1) :

---

<sup>174</sup> Cf. aussi la notion de *recognitional deixis* de Himmelmann (1996) ou celle d'*anadeixis de reconnaissance* de Cornish (2010a).

(224) donc on a donc on s'est déplacés avec nos valises pour trouver notre euh notre hôtel | \_  
| et pis euh sinon | \_ | après avoir déposé *les trucs* on a petit-déjeuné (ofrom)

Il est difficile de savoir si l'emploi du SN *les trucs* ne fait que rappeler l'objet 'les valises' antérieurement activé ou s'il englobe plus de types d'objets (toutes sortes d'affaires hétéroclites, par exemple). En tous les cas, le N « postiche »<sup>175</sup> a l'avantage de recouvrir potentiellement un ensemble non spécifié d'objets, contrairement à une autre forme lexicale plus spécifiée (cf. *infra* §5.1.3). De plus, son caractère passe-partout est particulièrement avantageux en termes de coûts d'encodage (O2), mais il signale également au destinataire l'inutilité d'en inférer davantage (O1).

On remarque ainsi que la sous-détermination peut être motivée par de nombreux facteurs susceptibles de se combiner entre eux, certains accidentels (méconnaissance du référent, d'une dénomination), d'autres stratégiques (économie, sauvegarde de la face, marquage d'une hétérogénéité énonciative ou autres effets de point de vue). Il va de soi que les conditions de production propres aux différents genres discursifs ont un impact sur la présence des marqueurs de sous-détermination, en particulier sur la gestion des causes accidentelles. On sait en général que les genres bénéficiant de conditions de préparation et supposant des normes visant à prévenir des coûts élevés de décodage évitent autant que possible ce que d'aucuns perçoivent comme des « accidents de performance » (cf. *supra* Ch.I §6.3.1). A l'inverse, les discours produits sur le vif reflètent l'état cognitif courant du locuteur, y compris toutes les « instabilités » qu'il suppose.

Ci-dessous, l'extrait d'oral spontané cumule les traces de sous-détermination, résultant aussi bien de causes accidentelles que stratégiques, qu'il est par ailleurs parfois difficiles de distinguer :

(225) le domaine là | électronique là c'était à la | \_ | #<sup>176</sup> | \_ | où y a le maintenant le | \_ |  
comment ça s'appelle | \_ | le centre euh | \_ | de recherche *et tout ça* | \_ | euh | \_ | ils ont  
démoli l'usine pour construire un autre bâtiment à la place | \_ | j'ai assisté à  
l'inauguration | de cette fabrique en quarante-sept (ofrom)

Le marqueur « spatial » *là* vise manifestement à situer dans un espace supposé connu un référent ('le domaine électronique'), dont la dénomination exacte fait défaut (comment *ça* s'appelle). Mais le marqueur convoque, plutôt qu'une situation géographique, l'espace intersubjectif des interlocuteurs ainsi sollicité (cf. *infra* §5.2.4). La lacune lexicale est visiblement aussi à l'origine de l'emploi d'*et tout ça*, consécutif à la tentative de formulation du locuteur (*le centre de recherche*). Outre l'incomplétude ou l'approximation lexicale, l'expression *et tout ça* contribue à évoquer grossièrement une classe-objet, *i.e.* un ensemble d'éléments hétéroclites rattachés au référent (*e.g.* son nom, sa localisation, sa fonction, etc.), mais dont le contenu détaillé est jugé non pertinent. A cela s'ajoute, au niveau interactionnel,

<sup>175</sup> Cf. Kleiber (1987) et Halmøy (2006) sur le mot *chose*.

<sup>176</sup> Le signe # représente une séquence anonymisée, remplaçant la plupart du temps un nom propre.

un effet de connivence, *et tout ça* servant d'appel à une représentation commune (qu'on pourrait gloser par *tu vois ce que je veux dire, je te passe les détails*). Pour finir, on observe l'usage d'un *ils* renvoyant à un agent sous-déterminé, dont l'identité n'est pas fournie par le contexte. Outre l'ignorance potentielle du locuteur à cet égard, on peut l'expliquer à nouveau par l'absence de pertinence de l'identification. En effet, à la faveur de cette sous-détermination, c'est le procès décrit, à savoir la démolition du bâtiment, qui est présenté comme l'information pertinente. On voit donc que les causes de l'émergence de la sous-détermination relèvent de stratégies discursives variées et cumulables, faisant à ce titre de celle-ci une véritable ressource communicationnelle.

Au vu de ces observations, nous proposons de classer les motivations de la sous-détermination de la manière suivante :

- a) facteurs accidentels : lacune lexicale, inexistence d'une catégorisation lexicale idoine, méconnaissance encyclopédique, culturelle ou perceptive du référent.
- b) facteurs stratégiques : i) en lien avec la pertinence : réduction des coûts de décodage (O1) ou d'encodage (O2), gestion de la structure informationnelle ; ii) en lien avec l'interaction : gestion des faces (O5), intersubjectivité, hétérogénéité énonciative et effets de point de vue.

Précisons que l'opposition entre a) et b) n'est que théorique, les marqueurs pouvant relever des deux types. En effet, on a vu ci-avant qu'un marqueur apparaissant en raison d'une instabilité cognitive accidentelle peut par la même occasion être doté de rendements pragmatiques et interactionnels.

#### **4.2 Sous-spécification sémantique**

Comme on l'a vu à travers les exemples précédents, les ressources linguistiques couramment exploitées dans l'émergence de la sous-détermination sont les expressions linguistiques sous-spécifiées. Précisons à cet égard que nous réservons la notion de sous-spécification à la description sémantique des expressions linguistiques et celle de sous-détermination à celle des objets-de-discours.

Rappelons qu'il est d'usage en sémantique de situer les unités sur des échelles de généralité/spécificité et d'y reconnaître un degré intermédiaire dit *de base* (ou *basique*) qui représente le niveau par défaut et intuitif pour les sujets parlants (Rosch *et al.* 1976), situé entre les niveaux subordonné (ou spécifique) et superordonné (général) (*cf. supra* Ch.II §2.4). Ainsi en va-t-il pour la série de noms *chien* (niveau de base), *teckel* (niveau subordonné, terme dit hyponyme) et *animal* (niveau superordonné, terme dit hyperonyme), qui manifestent entre eux des rapports d'inclusion sémantique : le signifié du lexème *animal* est inclus dans celui du lexème *chien*, lui-même inclus dans celui du lexème *teckel* (signifié le plus spécifié).

La prédisposition pour le niveau de base s'accorde bien, au niveau pragmatique, avec la maxime de quantité (Grice 1975) qui garantit un équilibre communicationnel issu d'un compromis entre le surplus d'informations d'une part et le déficit informationnel d'autre part.

Il va de soi que tout recours au niveau superordonné n'a pas pour effet de produire une représentation sous-déterminée en M. En effet, il arrive fréquemment qu'un hyperonyme rappelle un objet déjà déterminé, sans que celui-ci en perde pour autant ses propriétés distinctives :

(226) A cette époque, Amazone n'était qu'un enfant et il habitait près du port de Bélem, là où la neige la plus familière était le coton dont les navires étaient chargés. Il n'allait pas encore à l'école, il restait à la maison avec sa mère et commençait déjà à jouer sur un vieux piano qui traînait dans la cour et que son père avait récupéré Dieu sait où. *L'instrument sonnait faux, mais c'était déjà quelque chose de magique, un piano qui sonnait faux.* (M. Ferminé, *Amazone*)

Il en va de même pour les pronoms que contient l'extrait, qui sont pour leur part dépourvus de tête lexicale (et donc de traits lexicaux inhérents), par conséquent, moins spécifiés que les SN lexicaux du point de vue sémantique<sup>177</sup> ; leur emploi dans des situations de rappel d'objets déjà déterminés n'induit en soi aucune sous-détermination référentielle, comme on le voit ici.

Les expressions sous-spécifiées qui nous intéressent sont celles qui interviennent dans les cas de sous-détermination répertoriés *supra* (§2.3), *i.e.* les cas où la variable qu'elles introduisent n'est pas aisément unifiable avec un objet en M, celui-ci n'étant pas présent ni facilement inférable, ou alors sa délimitation demeurant indistincte.

### **4.3 Traitement cognitif de la sous-spécification**

Au niveau cognitif, on peut se demander comment le cerveau humain traite la sous-spécification sémantique. A l'instar de la linguistique, la psycholinguistique se fonde en général sur un postulat de traitement analytique et rigoureux du langage. Cependant, un certain nombre de chercheurs penchent pour une approche plus réaliste et pragmatique de l'interprétation du langage. On admet de plus en plus l'idée d'un traitement superficiel (*shallow processing*) dans certaines circonstances, pouvant se satisfaire d'interprétations sous-spécifiées (Sanford & Sturt 2002). En effet, un interprète peut se contenter d'un traitement superficiel si l'information explicitement communiquée demeure lacunaire ou si une interprétation complète ne vaut pas la peine d'être mise en œuvre. Il peut manipuler des représentations sous-déterminées en s'abstenant de s'engager sur certains aspects du sens. On parle alors de *minimal semantic commitment* (Frazier & Rayner 1990). L'idée est qu'un allocataire adaptera son traitement interprétatif en produisant des représentations « suffisamment bonnes » (*good enough representations*) pour l'objectif de la communication

---

<sup>177</sup> Rappelons que les pronoms ne sont pas totalement dépourvus de sens. Leurs traits morphologiques (genre, nombre, personne, cas, etc.), notamment, comportent des indications sémantiques. Voir *supra* Ch.II §2.

en cours, autrement dit, pour générer une réaction appropriée au stimulus (Ferreira *et al.* 2005). Le but n'est pas de nier l'existence de traitements analytiques approfondis, mais de doter le processus de compréhension d'une dimension plus heuristique lui permettant d'opérer plus rapidement (Towsend & Bever 2001). On reconnaît désormais en psycholinguistique que le système interprétatif calcule une signification plausible avant même que celle-ci soit confirmée par les informations explicites. Les concepts d'interprétation « suffisamment bonne » (Ferreira *et al.* 2005) ou de « traitement superficiel » (Sanford & Sturt 2002) illustrent le recours à un type d'heuristique visant à accéder à une signification en faisant le moins d'efforts possibles<sup>178</sup>. Diverses expériences psycholinguistiques mettant en scène des inconsistances sémantiques ou logiques, connues sous les noms de *Moses illusion* (Erickson & Mattson 1981), *'survivors' anomaly* (Barton & Sanford 1993), corroborent l'idée d'un traitement rapide et superficiel dans certaines conditions.

Il faut préciser que ces résultats portent sur des données de laboratoire, souvent obtenues à partir de stimuli visuels produits dans des conditions rigoureusement contrôlées et dépouillées de tout ancrage contextuel. Certes, c'est là une condition nécessaire à la conduite de toute étude expérimentale visant à vérifier telle ou telle hypothèse au moyen de tests statistiques appliqués aux données récoltées. Mais malgré une volonté louable de décrire de manière plus réaliste le traitement du sens par les interprètes, l'approche demeure paradoxalement éloignée des usages réels et des conditions accidentelles et stratégiques (*cf. supra* §4.1) qui favorisent la sous-détermination.

D'autre part, la perspective se concentre uniquement sur la tâche d'interprétation, et ne considère pas le processus de production, où des stratégies de simplification sont manifestement aussi à l'œuvre. Enfin, l'approche part implicitement du principe que le traitement superficiel dépend du seul niveau sémantique (*cf.* les inconsistances sémantiques proposées dans les stimuli respectifs<sup>179</sup>) ; elle n'envisage pas le fait qu'une telle interprétation puisse résulter de la nature même des référents, en tant que constructions cognitives potentiellement floues et indistinctes (Berrendonner 2014). Comme toujours, à défaut de pouvoir accéder directement aux représentations discursives des interlocuteurs d'une interaction (Grize 1993), nous sommes sur ce point réduits, en tant que linguistes, à faire des hypothèses à partir des indices lacunaires dont nous disposons. C'est néanmoins la tâche que nous nous efforçons de poursuivre dans ce travail.

---

<sup>178</sup> Ces concepts gagneraient à être rapprochés de la notion de pertinence de Sperber & Wilson (1986).

<sup>179</sup> *Cf.* « How many animals of each sort did Moses put on the ark ? » (Erickson & Mattson 1981) ; « After an aircrash, where should the survivors be buried ? » (Barton & Sanford 1993) ; « If you don't break the rules, then I'll tell the authorities » (Fillenbaum 1974) ; « The dog was bitten by the man. » (Ferreira *et al.* 2005)

## 5. Inventaire des moyens d'expression de la sous-détermination

Après avoir défini ce que nous entendons par *sous-détermination*, délimité l'objet d'étude et dégagé les conditions favorables et les mécanismes à l'œuvre dans son émergence, nous nous proposons d'examiner les ressources linguistiques à travers lesquelles elle se manifeste dans des genres variés. Cette section contient donc un inventaire des moyens d'expression à la disposition des usagers. Rappelons que nous nous intéressons exclusivement aux opérations de pointage référentiel (et non d'introduction) (*cf. supra* Ch.I §6.3.5) dans le cadre desquelles surgit un obstacle à l'unification de la variable introduite avec un objet valide en M : d'une part, l'objet visé peut être absent, soit momentanément (versé en M lors d'une étape ultérieure) (i), soit il doit être inféré de manière non fiable, étant de ce fait susceptible d'être incarné par toute une gamme de référents potentiels (ii) ; d'autre part l'objet peut être valide mais ses frontières et sa composition exhaustive demeurent incertaines (iii).

Notre inventaire est organisé selon un critère lexical, distinguant d'une part les ressources lexicales (bien que sous-spécifiées), dont la signification se compose de traits conceptuels inhérents (§5.1), d'autre part les ressources non lexicales (§5.2), dépourvues de ce type de traits et dont la signification peut être décrite comme fonctionnelle. Il faut noter que la distinction entre *lexèmes* et *morphèmes* (entendu grammaticaux) n'est de loin pas toujours claire, comme le souligne Martinet (1967), certains éléments relevant à la fois du lexique et de la grammaire (il donne l'exemple des prépositions *pour* ou *avec*). C'est, selon le linguiste, l'aspect fonctionnel qui fait basculer un élément du côté des morphèmes (grammaticaux) (*ibid.* : 112)<sup>180</sup>. Nous verrons que cette distinction a ses limites, notamment dans le cas des termes postiches (§5.1.3). Néanmoins, elle se montre suffisamment pertinente pour notre propos, qui est de fournir une vue d'ensemble sur les ressources disponibles.

### 5.1 Ressources lexicales

#### 5.1.1 Hyperonymes

Nous avons déjà évoqué *supra* (§4.2) la possibilité de classer hiérarchiquement les unités lexicales en termes d'inclusion sémantique et en distinguant différents degrés de spécification entre celles-ci. Ainsi, les hyperonymes ou termes superordonnés se caractérisent par une sous-spécification lexicale par rapport à leurs subordonnés, ou hyponymes. Leur sens est inclus dans celui de ces derniers, qui se montrent sémantiquement plus étoffés. A l'inverse, l'extension des hyperonymes est plus large que celle de leurs hyponymes : elle recouvre des entités moins homogènes que celle des hyponymes (*cf.* la différence *supra* entre *instrument* vs *piano*). L'exemple suivant met en jeu l'hyperonyme *acte* dans l'expression *passer à l'acte* :

---

<sup>180</sup> Néanmoins, il existerait selon Martinet un critère distributionnel : les lexèmes apparaîtraient à des places susceptibles d'être remplies par des classes ouvertes d'items, tandis que les morphèmes grammaticaux appartiendraient à des classes limitées (*ibid.* : 119).

(227) « Mathieu est passé à *l'acte*... le pire moment de ma vie » (titre d'un article, <http://www.lapresse.ca/la-tribune>, 15.11.2016)

La ritualisation du rendement euphémique de l'expression *passer à l'acte* fournit quelques pistes d'interprétation (e.g. la prédilection pour les domaines de la violence ou de la sexualité), mais il demeure à partir du titre impossible d'attribuer une valeur stable au référent, étant donné la diversité des candidats potentiels ({agression, meurtre, suicide, acte sexuel, radicalisation, etc.}). Outre son rendement euphémique, l'expression, par ailleurs imputée à une instance énonciative distincte de celle du journaliste, vise à créer un effet d'attente, auquel le contenu de l'article met fin en révélant le type d'acte en question, à savoir le suicide d'un adolescent.

Dans la référence aux humains, on peut mentionner l'emploi du SN *les gens*, le N *gens* englobant dans son extension toutes sortes de sous-catégories potentielles. Cappeau & Schnedecker (2014) font l'hypothèse d'une évolution des SN *les gens / des gens* vers un statut pronominal, au vu des réanalyses auxquelles ils donnent lieu, les intégrant au paradigme des pronoms indéfinis du français. Dans l'exemple ci-dessous, il n'est pas évident de décider si le SN renvoie à un ensemble hétérogène ({les gens qu'ils croisaient, quels qu'ils soient}) ou, à l'inverse, s'il convient d'inférer une communauté particulière à partir du contexte ({les élèves, les instituteurs}) :

(228) tandis que mon mari | \_ | qui avait neuf ans de plus que moi | \_ | entre eux et les parents ils parlaient le patois depuis tout petits | \_ | et quand ils ont été à l'école ils ont eu beaucoup de peine | \_ | et *les gens* ils se moquaient d'eux des fois | \_ | parce que ils avaient peine à parler français (ofrom)

Le SN ne semble toutefois pas requérir un traitement plus approfondi, laissant par là même le référent sous-déterminé.

### 5.1.2 N sous-spécifiés « capsules »

On peut distinguer des hyperonymes les « noms sous-spécifiés » (Legallois 2006, 2008) d'un autre type, appelés aussi noms « capsules » (*shell nouns* < Schmid 1997, 2000) ou encore « labels » (Francis 1994), parmi lesquels<sup>181</sup> on trouve pour le français des noms comme *idée*, *objectif*, *conseil*, *aspect*, *résultat*, *souhait*, *fait*, *chose*, *problème*, *raison*, *conséquence*<sup>182</sup>, etc. Contrairement aux hyperonymes, ces noms ne s'inscrivent pas au sein d'une hiérarchie lexicale telle que les degrés de Rosch *et al.* (1976) reposant sur des rapports d'inclusion sémantique et leur interprétation ne passe pas par la reconnaissance de traits prototypiques assurant à ceux-ci une certaine stabilité conceptuelle (Legallois : 2008).

<sup>181</sup> Pour un recensement sur un corpus journalistique, voir Legallois (2008).

<sup>182</sup> A noter que la plupart de ces termes sont polysémiques et qu'ils n'entrent pas dans cette catégorie dans tous leurs emplois (cf. le mot *chose* qui peut fonctionner comme N postiche, cf. *infra*).

Pour délimiter la catégorie des noms sous-spécifiés, Legallois (2006, 2008) exploite le critère d'occurrence dans des structures « spécificationnelles », *i.e.* des constructions du type *N (c')être de inf./que P*, comme dans *l'objectif est de/que...*<sup>183</sup> :

(229) Leur idée a été de dresser le « portrait-robot » du patron performant (*Libération* < Legallois 2008)

Ce type de constructions vise à spécifier le N en question à travers la P qu'il introduit, constituant à cet égard sa « réalisation lexicale » (Winter 1992). Ces constructions ne sont cependant pas celles qui nous intéressent, étant donné que la P infinitive fournit justement l'identité référentielle au SN *leur idée* au sein d'une même clause. Mais une part de sous-détermination peut surgir dans le cas où le SN renvoie par exemple, selon le type (iii), à un sous-graphe déjà présent en M (§6.3.4), comme ci-dessous :

(230) Nous avons encore tous et toutes en mémoire les images graves du tremblement de terre catastrophique de janvier en Haïti et nous savons que les habitants de ce pays auront besoin d'aide encore très longtemps. Dans **cet esprit**, le Département de Chimie de notre Université a pris l'initiative d'organiser deux éditions spéciales de son spectaculaire cours expérimental "Harry Potter et le tournoi magique" en faveur des victimes du tremblement de terre. (courrier électronique adressé à la communauté universitaire, 05.02.2010)

Le SN *cet esprit* renvoie à l'état mental – via sa composante lexicale – résultant des informations évoquées par le contexte préalable, à savoir, le souvenir des événements en question et la conscience du besoin d'aide qui en découle. Il embrasse et construit par anaphore résomptive sous cette « étiquette » (*cf.* le terme « label » de Francis 1994) un ensemble d'objets hétérogènes et de relations entre eux, dont les limites ne sont pas posées a priori. En outre, il engage un positionnement métadiscursif et argumentatif du locuteur, visible notamment dans son rôle de connecteur consécutif (*dans cet esprit*).

La capacité d'« encapsulation » (Conte 1996), d'« emballage » (Legallois 2008) résomptif, susceptible de véhiculer un point de vue (Schmid 2001) ou un aspect évaluatif (Legallois 2008) a été maintes fois relevée dans la littérature (Francis 1994, Conte 1996, Flowerdew 2003, Flowerdew & Forest 2015). L'une des spécificités du procédé référentiel relevée par Schmid (1997) est le caractère temporaire de la « réification » opérée, presque exclusivement dépendante du contexte, étant donné que l'identification s'affranchit dans ce cas du recours à des traits stables du signifié. C'est ce qui fait dire à Legallois (2008) que le locuteur, libéré des contraintes sémantiques, agit avec les N sous-spécifiés en « démiurge », opérant à sa guise des catégorisations d'objets, des « fictions » créées en discours « libre de catégoriser ce qu'il veut, comme il veut ».

---

<sup>183</sup> *Cf.* aussi les contextes d'apparition de Schmid (2000 : 22) pour les *shell nouns* en anglais : *N-that, N-to, N-wh, th-N, th-be-N* (nous reprenons sa notation).

A notre avis, cette vision des objets construits par le discours et catégorisés à leur guise par les sujets parlants vaut au-delà des emplois de noms sous-spécifiés ou autres métaphores : elle caractérise le fonctionnement de tous les procédés référentiels. Ainsi, l'emploi de tout SN référentiel évoque, au moyen de ses traits sémantiques inhérents (certes en partie influencés par le monde extérieur), une représentation discursive façonnée par le locuteur « démiurge ». Autrement dit, l'usage des N sous-spécifiés ne doit pas être vu comme un cas particulier de la créativité discursive du locuteur, comme le suggère Legallois (2008), mais il confirme au contraire le véritable rôle joué par le locuteur dans les opérations de référence en général.

La caractéristique des noms « capsules » réside à nos yeux dans la présence d'une catégorisation lexicale un peu particulière : conformément aux vertus « réifiantes » des N en général (Schmid 1997), elle confère à l'objet une certaine cohésion sous couvert d'un point de vue ; mais en étant sous-spécifiée, elle est par définition dans l'incapacité de délimiter clairement celui-ci. Il sera intéressant de mettre cette situation en perspective avec l'usage de ressources non lexicales ayant un fonctionnement semblablement résumptif (les pronoms *ça*, *le*, *en*, *y* (§5.2.2)), *i.e.* capables de synthétiser un même sous-graphe de M sans pour autant catégoriser celui-ci.

### 5.1.3 Termes postiches

Une catégorie de termes sous-spécifiés se distingue enfin des précédentes par le fait que ses membres ne s'inscrivent ni dans une hiérarchie lexicale, ni dans un fonctionnement d'« encapsulation ». Bien que d'apparence « lexicale » (ou « semi-lexicale », Mihatsch 2006), leur caractéristique réside dans leur emploi passe-partout en lieu et place de lexèmes plus spécifiés<sup>184</sup>, d'où la qualification de termes « postiches » ou « vicaires ». Parmi ceux-ci, on relève par exemple les N *chose*, *machin*, *truc*, *bidule*, *zinzin*, *schmilblick*<sup>185</sup>, etc. Les motivations principales de leur emploi semblent être la volonté de cryptage (par euphémisme par exemple), l'indifférence d'une spécification (le N fonctionnant comme *free choice item*), un déficit lexical, la complexité du concept qu'il « remplace » ou l'insignifiance du référent. Les paroles d'une chanson d'Henri Dès ci-dessous combinent ces diverses finalités, où un papa-expert est mis en scène dans les réponses qu'il donne invariablement à son enfant sur des questions d'ordre technique (*e.g. comment c'est fait dans un piano/ dans une fusée*, etc. ?) :

(231) C'est **le p'tit zinzin** qui passe par ici et qui va toucher **le p'tit machin**, et **le p'tit machin** qui repasse par là et qui fait marcher **le p'tit zinzin** (refrain de la chanson *Dis papa*, Henri Dès)

<sup>184</sup> Pour une description de ces noms postiches en lieu et place de noms propres, voir Kerbrat-Orecchioni (2010) et Schnedecker (2011).

<sup>185</sup> La liste est vraisemblablement restreinte, mais nous n'excluons pas des variantes de type régional, voire d'autres types de variantes (idiosyncrasiques par exemple). A noter également que certains items peuvent se combiner (dans divers ordres) : *truc-machin-chose*, *truc-bidule-chouette*, etc. Il faut encore ajouter qu'ils ne commutent pas aléatoirement et possèdent donc des conditions d'emploi variables. A cet égard, voir les analyses de Kleiber (1987), Halmøy (2006), Mihatsch (2006), Bidaud (2015).

L'effet humoristique de la chanson repose notamment sur le procédé parodique, auquel participent les termes postiches, visant à ridiculiser le discours pseudo-savant du père. Il repose surtout, à la fin, sur le caractère tabou de la dernière question du fils, qui interroge son père sur la manière de faire les bébés. La réponse reste toujours la même, mais l'embarras est alors bien perceptible à l'oreille (raclement de gorge, timbre et intensité de la voix, etc.) ! On peut considérer que les dénominations postiches sont employées en lieu et place, dans un premier temps, de termes techniques compliqués, dont le père ignore peut-être le nom ou dont ils soulignent la futilité ; dans un second temps, en lieu et place de termes tabous autour du mystère de la sexualité. Dans tous les cas, il reste difficile d'en inférer une identité ou une dénomination univoques, en raison de leur occultation délibérée.

## 5.2 *Ressources non lexicales*

### 5.2.1 *Anaphores zéro*

Il arrive que l'instruction référentielle soit exprimée au travers d'un « désignateur » zéro. C'est le cas par exemple avec des verbes transitifs dont le complément n'est pas réalisé « en surface » :

- (232) [la locutrice, 82 ans, raconte à sa petite-fille ses souvenirs de famille] je faisais des tartines pour euh | \_ | enfin pour toute l'équipe quoi | \_ | et pis à la récré ben ils venaient au/ | % | \_ | pis une fois que quand c'était passé que c'était le | \_ | quand | #<sup>186</sup> | allait arriver alors là | \_ | alors là | \_ | tout le monde fi/ filait parce qu'entre deux ils avaient fait des crasses | \_ | ils avaient cassé des fenêtres tu vois enfin | \_ | je t'*explique* Ø un peu en gros | \_ | t'es pas obligé de tout dire hein | \_ | enfin voilà quoi (ofrom)

Le régime direct du verbe *expliquer* n'a pas ici de réalisation lexicale. Outre l'implication sémantique d'un actant<sup>187</sup>, l'élément zéro donne l'instruction d'unifier sa variable avec un objet valide en M, ici un sous-graphe de M déjà introduit ou en cours d'introduction, dont la nature, aussi bien que les limites, demeurent largement sous-déterminées.

Schnedecker (2014) relève la disposition des anaphores zéro à apparaître dans les recettes de cuisine, aptes à renvoyer à des référents peu identifiables :

- (233) Mettre les légumes (pelés et en cubes), le whisky et le bouillon dans une casserole. Couvrir Ø d'eau et porter Ø à ébullition. Laisser cuire Ø à feu moyen pendant 30 minutes.  
Retirer les clous de girofle et les graines de cardamome verte.  
Passer la préparation au mixer hors du feu.  
Une fois mixée, mettre la préparation sur feu doux, ajouter les épices et le beurre.  
Laisser cuire Ø au minimum 15 min (Plus ça mijote, meilleur c'est !)  
Saler Ø et poivrer Ø à votre goût !

<sup>186</sup> Pour rappel, signe d'anonymisation.

<sup>187</sup> Cf. les emplois absolus du type *Il y a toute une population qui casse pour casser* (Noailly 1997).

Vous pouvez servir  $\emptyset$  avec des croûtons au beurre et des copeaux de parmesan.  
(texte 3 < *ibid.* 31)

Les formes zéro réfèrent régulièrement au résultat de mélanges ou autre processus propres au domaine en question (*ibid.*). Le fait que les anaphores zéro puissent renvoyer à des objets mal délimités remet notamment en cause – ou du moins nuance – l’idée qu’elles encodent une accessibilité élevée ou autre facilité de traitement (Ariel 1988, 1990).

## 5.2.2 Pronoms conjoints

### 5.2.2.1 Pronoms sujets : *ça, ils*

Le pronom clitique sujet *ça* (ou la forme *ce* devant le verbe *être*) possède une flexibilité référentielle remarquable lui permettant de renvoyer à toutes sortes d’objets (catégorisés ou non, animés ou non, etc.) dans des circonstances très diverses. Nous nous limitons ici à un aperçu de ces possibilités et renvoyons pour le reste *infra* (Ch.IV §2) et aux nombreux travaux consacrés à la question (Porquier 1972, Maillard 1989, 1994, Cadiot 1988, Corblin 1991, Carlier 1996, etc.). Le clitique *ça* s’utilise régulièrement avec des prédicats à agent animé pour mettre en évidence un procès « dépourvu d’agent assignable » (Maillard 1985). Il se rapproche en cela d’un emploi impersonnel (Maillard 1985, 1989, 1994) :

(234) Ici c’est Fribourg ! **Ça** jase au bord de la Gérine. Et ces remous ont pour effet de pimenter la campagne électorale à Marly. Leur cause : le refus de la coopérative locale CRIC-Print d’imprimer 6000 tracts de l’UDC marlinoise en vue des élections communales du 28 février. (*La Liberté*, 04.02.16, chapeau d’un article intitulé *Une entreprise refuse de travailler pour l’UDC*)

Maillard (1991) explique que l’emploi de *ça* au détriment d’un *on* indéfini est favorisé par la concurrence de l’emploi personnel de ce dernier ( $\approx$  *nous*). Plus que *on*, *ça* permet d’occulter l’agentivité du verbe, en plaçant le procès au centre de l’attention. Selon Maillard (1985), on peut considérer ce genre de construction comme un énoncé thétique, dans le sens où il pose l’existence d’un objet, qui devient dès lors disponible comme thème pour une prédication successive<sup>188</sup>. C’est d’ailleurs ce qu’il se passe par la suite, les ‘jaseries’ se voyant rappelées métaphoriquement via *ces remous*, comme thème du propos qui suit. Selon cette analyse, la variable introduite par *ça* demeure une variable sous-déterminée, *i.e.* dont l’instanciation n’est pas réalisée. On peut donc se demander, dans ce cas, si cela fait encore sens de parler de désignateur (*cf.* le cas du *il* dit impersonnel).

Outre sa productivité dans « l’impersonnalisation » de verbes à agent animé (Maillard 1994), l’emploi de *ça* est un bon moyen de rendre compte d’une référence indistincte liée à un contexte énonciatif spécifique :

---

<sup>188</sup> Une construction thétique s’oppose à une construction dite *catégorique* dont la composition est binaire, constituée d’un sujet et d’un prédicat, ou, en des termes informationnels, d’un thème et d’un rhème.

(235) [à propos d'un funiculaire muni de contrepoids à eaux usées] une fois j'ai | \_ | je l'ai pris pis y avait des Genevois qui étaient en | \_ | je sais pas genre en découverte ici pis ils font | \_ | mais | \_ | ça pue | \_ | qu'est-ce que cette odeur | \_ | c'est les égouts | \_ | c'est la merde | merde | pis après j'ai | \_ | j'ai dit ouais ben | \_ | ça marche aux égouts (ofrom)

Il faut tout d'abord noter que les deux occurrences de *ça* interviennent dans un discours direct et qu'elles donnent lieu, à la faveur de leur token-reflexivité, à une référence à interpréter en lien étroit avec la situation d'énonciation, celle rapportée par l'énonciateur citant. Elles s'inscrivent ainsi dans le jeu de mime du discours reproduit (Béguelin 1997b). La première occurrence de *ça* renvoie à l'émanation de la situation rapportée, dont l'origine est inconnue et difficilement descriptible pour les énonciateurs. S'ensuivent quelques hypothèses des énonciateurs cités visant à catégoriser l'odeur perçue au moment en question (*c'est les égouts, c'est la merde*).

La seconde occurrence de *ça*, sujet du verbe *marcher*, réfère à un objet manifestement établi en M dans le déroulement de l'énonciation reproduite, à savoir, le funiculaire. On peut dès lors se demander pourquoi la locutrice privilégie l'emploi de *ça* au détriment d'un *il*, supposant une classification déjà attribuée. On peut faire l'hypothèse que le prédicat s'applique au *type* dont on a évoqué un ressortissant perceptible, voire à sa classe extensionnelle. Ou alors, le pronom *ça* vise à englober plus que le véhicule, via sa capacité métonymique (Corblin 1987b), par exemple, l'ensemble complexe de son mécanisme : plus que le véhicule, c'est le dispositif en place dont on décrit le fonctionnement (*aux égouts*), qu'il est difficile d'envisager de manière délimitée.

Aux côtés de *ça*, on peut ajouter le clitique *ils*, régulièrement employé dans des constructions vouées à présenter un procès dont l'agent reste sous-déterminé :

(236) Le dernier plaisir de ce fossile vivant consistait à décortiquer une ou deux gambas. Il en mangeait d'ailleurs de moins en moins, car la cuisine à l'huile ne lui réussissait pas... Pauvre Tanabe! Bientôt tu seras accueilli au nirvana : **ils ont installé** à l'entrée un stand de gambas frites où tu te goinfreras à l'œil et là, pas trop d'huile... (Eric Faye, *Nagasaki*, p.17)

Contrairement à l'emploi traditionnellement décrit du clitique de 3<sup>e</sup> personne, l'usage de la 6<sup>e</sup> personne apparaît ici *ex nihilo*, sans réel signe précurseur de l'existence d'un référent préalable. Tout au plus peut-on inférer qu'il s'agit d'occupants dudit lieu (le nirvana), mais leur identité n'en est pas davantage inférable. Nous reviendrons largement sur le cas de ces expressions dédiées à la référence des agents sous-déterminés dans les deux derniers chapitres de ce travail.

### 5.2.2.2 Pronoms régimes : *le*, *y*, *en*

Un peu à la manière de *ça*, les clitiques régimes *le*, *y* et *en* sont susceptibles de renvoyer à des entités non classifiées, *i.e.* qui ne possèdent pas de dénomination conventionnelle, en

particulier des objets de type ‘procès’ évoqués par les constructions verbales. Les différentes formes correspondent à leur flexion en cas, respectivement à l’accusatif pour *le*, au locatif pour *y*<sup>189</sup> et à l’ablatif pour *en*<sup>190</sup>.

Le pointage peut opérer sur des objets plus complexes, comme on l’a déjà vu *supra* (§2.3, ex. (214)), référant sur le mode résomptif à un sous-graphe de M déjà valide en M mais dont la délimitation reste indistincte. Cela nuance la caractérisation souvent supplétive qui est faite de ce type de pronom<sup>191</sup>. Ainsi en va-t-il également de l’exemple suivant :

(237) je vais vous raconter ce qui m'est arrivé le le quatre en- avril quatre-vingt-huit + c'était pas sympa + j'arrivais de vacances c'était sympa pour moi mais pas pour + pas pour elle + j'étais parti seul + je sais pas comment je je v- + vais faire pour vous **le** raconter mais + mais je vais essayer quand même (crfp)

Cet exemple consiste en l’ouverture d’un épisode narratif explicitement présenté comme tel par le locuteur (*je vais vous raconter...*). A la fin de l’extrait, le locuteur interrompt le début de son histoire par un commentaire à fonction métadiscursive consistant à émettre des doutes sur ses compétences narratives. C’est à ce moment qu’apparaît le clitique *le* objet du verbe *raconter*. Dans une vision substitutive, le pronom *le* serait vraisemblablement considéré comme un représentant du complément *ce qui m’est arrivé le quatre avril quatre-vingt-huit*. Or, dans une approche constructiviste du discours, on interprétera de préférence le pronom *le* comme renvoyant à l’événement en cours de narration. Certes, il s’agit bien de l’objet annoncé via ledit complément, mais le pronom tient également compte de son évolution au fil du discours, en l’occurrence, des évaluations émises à son propos (*sympa pour moi, pas pour elle*) ainsi que de l’élaboration de la première étape narrative (*j’étais parti seul*). L’objet en question est donc en cours d’élaboration au moment où le locuteur y fait référence par anticipation. A cet égard, on peut considérer le fonctionnement du pronom comme anacataphorique (Kęsik 1989 : 79), non pas dans un sens textualiste, mais dans un sens cognitif : il donne l’instruction d’identifier sa variable avec un objet encore partiellement introduit en M partant, en attente de déterminations ultérieures (Johnsen 2008, 2014a).

### 5.2.2.3 Pronoms régimes dans les « aphorismes lexicalisés »

Certains pronoms clitics régimes ont la particularité d’apparaître dans des locutions verbales du genre *s’y mettre* (≈ commencer), *s’en sortir* (≈ réussir), *la fermer* (≈ se taire), *se les cailler* (≈ avoir froid), etc. (Béguelin 2014a). L’analyse accordée à ces lexies ne recourt pratiquement jamais à la notion de référence, probablement parce qu’on les considère comme des « idiomatismes » lexicalisés dont le sens serait compositionnellement inanalysable (*ibid.* :

<sup>189</sup> Le clitique *y* fonctionne également au datif : *je lui ressemble vs j’y ressemble*. D’autre part, le terme *locatif* n’a pas d’implication sémantique, *y* pouvant renvoyer à autre chose qu’un lieu. C’est une simple étiquette casuelle.

<sup>190</sup> Le clitique *en* fonctionne également à l’accusatif avec un sens partitif (e.g. *j’en ai bu*).

<sup>191</sup> Voir par ex. Riegel *et al.* (2009 : 369) : « La forme invariable *le* représente un groupe verbal (complément du verbe *faire*), une proposition ou un attribut [...] »

156). Béguelin examine néanmoins les circonstances favorisant l'émergence de ce genre d'« aphorismes<sup>192</sup> lexicalisés » qui proviennent, à l'origine, de situations référentielles. Elle relève à cet égard des cas intéressants de double analyse, où le clitique peut s'interpréter tantôt de manière référentielle, tantôt de manière « autarcique » (*ibid.* : 139, terme emprunté à Damourette & Pichon 1911-1946, VI). Nous avons également rencontré ce type de situation :

(238) il a mis tout de même une bonne cinquantaine d'années à se ruiner dans l'édition et encore il est pas ruiné il s'**en** est tiré assez bien (ctfp)

Dans cet exemple, le clitique locatif *en* peut renvoyer à la situation désastreuse évoquée (la ruine dans le milieu de l'édition) dont l'individu s'est tiré, sans toutefois que le référent soit bien circonscrit. Mais on peut aussi considérer qu'il s'intègre pleinement à la locution verbale, celle-ci signifiant d'un bloc l'aboutissement au succès de l'individu. On conviendra cependant que dans cette dernière interprétation, il est difficile d'omettre complètement l'état antérieur de la réussite.

Béguelin (2014a) propose également une double analyse du régime accusatif *en* ci-dessous :

(239) Dites bien comme moi, surtout, et n'**en** rajoutez pas.  
Moins vous **en** direz, mieux cela vaudra. (Frantext, Desnos, 1943 < Béguelin 2014a : 140)

Dans une interprétation référentielle, *en* peut évoquer en anaphore indirecte un référent comme des 'paroles', des 'propos', etc. qu'on peut inférer en prenant appui sur la première occurrence du verbe *dire*, comportant dans son signifié un objet du dire. Mais la locution *en rajouter* dans son ensemble peut approximativement équivaloir au sens d'*exagérer*. Il en va de même pour *en dire* qui peut signifier simplement *parler* ou *s'exprimer*.

Ces situations sont intéressantes par le fait que le choix entre interprétations référentielle ou « autarcique » n'est au final pas tranché et qu'il est à ce titre difficile de parler de réelle ambiguïté interprétative. On a plutôt l'impression d'avoir affaire à une référence « évanescence » (d'après le terme de Béguelin, *ibid.* : 152).

### 5.2.3 Pronoms disjoints

#### 5.2.3.1 Les démonstratifs *ça, ceci, cela*

Cette catégorie se distingue de la précédente par sa distribution syntaxique. On constate que *ça* y figure à nouveau, cette fois dans des positions non exclusivement sujet, contrairement à son emploi clitique. On a depuis longtemps relevé l'affinité de la série des démonstratifs *ça, ceci, cela* avec les entités non classifiées (Kleiber 1984, Maillard 1989, Corblin 1987) :

---

<sup>192</sup> Béguelin (2014a) emprunte le terme à Maillard (1974 : 56), l'adjectif *aphorique* s'opposant à *anaphorique/cataphorique* lorsque l'élément « est parfaitement clos sur lui-même et n'implique pas le texte ».

Les formes neutres *ce*, *ceci*, *cela* s'emploient quand on n'a aucun nom dans l'esprit (Bonnard 1950 : 77, cité par Corblin 1987b)

Dès lors, ils sont régulièrement exploités, comme les pronoms régimes *le*, *y*, *en*, sur le mode résomptif, renvoyant à des sous-graphes comportant nombre d'objets et de relations, certains implicites. A la différence de ceux-ci, ils présentent en outre un fonctionnement token-réflexif :

(240) oui alors | # | euh il pratique la natation | donc il est il est médaille d'or je dirais que il est très doué | \_ | à la natation | \_ | et **tout cela** ben ça demande aussi de l'entraînement ça demande de la présence | \_ | ça demande euh beaucoup de | \_ | de de motivation de la part euh des parents et | de la maman puisque c'est moi qui fais quand même **tout ça** (ofrom)

La locutrice répond à une question concernant le sport que pratique son fils (dont le nom est anonymisé via le signe #). Dans une position disloquée, *tout cela*, dont sont prédiquées les exigences (*ça demande...*), permet d'englober l'activité de son enfant ('la natation') de même que sa classe-objet : par exemple son talent, son niveau, etc. qu'on peut inférer de la séquence argumentative. Plus loin dans le discours, on relève aussi l'emploi d'un *tout ça* complément du verbe *faire*. Ici également, le SN renvoie globalement à l'ensemble des conditions et autres ingrédients concomitants implicites mis en œuvre par la mère pour la réussite de son fils. On peut envisager l'interprétation comme impliquant une « expansion métonymique » accompagnant le « déclasserement » opéré par *ça* et *cela* (Corblin 1987b).

Un autre exemple ci-dessous montre que *cela* se dote d'une capacité de synthèse remarquable ; les circonstances qu'il résume font l'objet d'un vaste développement dans le discours préalable :

(241) [...] le Mouvement de libération nationale (MLN) m'a proposé de partir pour Paris comme responsable régionale des Jeunes du MLN. J'étais logée avec cinq garçons, dans l'appartement de l'écrivain Robert Brasillach alors en prison pour Collaboration. Jamais je n'avais vu une telle bibliothèque ; moi qui adorais lire et qui avais toujours eu peur de manquer de lectures, j'étais éblouie par tous ces livres que je pouvais feuilleter sans contrainte. Puis avec Rivière et un autre camarade nous avons déménagé pour nous installer à Montmartre dans l'appartement de Le Vigan, acteur célèbre à cette époque, emprisonné lui aussi pour collaboration. Nous avions de l'amour plein le coeur et plein les yeux, nous parlions dans les congrès, nous écrivions dans les journaux. Nous avions l'impression d'être investis de grands pouvoirs et j'avais avec mes camarades l'espoir fou de transformer la vie. Nous allions changer le monde. Mais **cela** n'a pas duré très longtemps non plus : fin décembre 44, le MLN n'avait plus d'argent pour nous payer et mes parents sonnaient le tocsin pour que je revienne à Lyon reprendre mes études et ma vie de jeune fille convenable. (Domenach-Lallich, D., *Demain il fera beau : journal d'une adolescente* (novembre 1939-1944), 2001< Frantext)

Le pointage via *cela* permet ici de saisir grossièrement tout un sous-graphe déjà établi M, en tant qu'expérience ou épisode de la vie de la locutrice.

Signalons que les pronoms démonstratifs de ce type fonctionnent également volontiers de manière cataphorique, c'est-à-dire en pointant sur un référent encore non disponible en M. C'est l'un des emplois privilégiés de *ceci* « résomptif » (Kęsik 1989, Theissen 2008) :

(242) Donc, d'avarice, pas trace. Ah si, peut-être une petite chose : l'électricité. [...] Et, pour en finir sur ce point, **ceci**. Lorsque Jean Rouaud reçoit le prix Goncourt en 1990, nous sommes un certain nombre d'auteurs de la maison à recevoir une lettre de Jérôme Lindon nous informant que ce succès, à ses yeux, ne serait pas complet si nous n'y participions pas. Cette lettre est accompagnée d'un chèque dont je ne me rappelle plus le montant, mais dont je me souviens qu'il est tout à fait substantiel. (Echenoz, J., *Jérôme Lindon*, 2001< Frantext)

L'auteur de ce texte évoque l'absence d'avarice chez J. Lindon. Après avoir reconnu une exception qu'il développe (pour des raisons de place, nous avons éliminé ce passage), il propose une illustration de son appréciation via le pronom *ceci*. C'est alors toute une anecdote qui est relatée pour soutenir le propos initial. Au vu de l'orientation argumentative du texte, *ceci* est capable à lui seul – sans même faire l'objet d'une prédication (il fonctionne de manière syntaxiquement autonome) – de conférer au référent à venir, via sa token-réflexivité, le statut d'évidence, de preuve.

Ces pronoms apparaissent ainsi comme des moyens commodes de saisir « en gros » des sous-graphes soit établis en M, soit à venir, à savoir un type d'objets qui ne possède pas de dénomination conventionnelle. Comme ils partagent la distribution des SN, on peut les mettre en regard des noms sous-spécifiés (*supra* §5.1.2), qui permettent pour leur part au locuteur de catégoriser subjectivement le même genre d'entités. Alors que les vertus réifiantes (Schmid 1997) de la dénomination via un nom « coquille » s'accompagnent de l'attribution explicite d'un point de vue sur l'objet, l'opération référentielle via un pronom démonstratif se borne à réunir momentanément, sans le nommer, un sous-graphe concomitant à l'énonciation.

Il va de soi que les trois démonstratifs ne sont pas interchangeable en toutes circonstances. Notons d'abord que *ceci* et *cela* entrent régulièrement dans des locutions comme *ceci/cela dit*, *cela étant* fonctionnant comme connecteurs, ou alors en corrélation dans des expressions du genre *pour critiquer ceci cela* (cfpp) où ils se dotent d'une valeur postiche (cf. *supra* §5.1.3). D'autre part, les emplois semblent diverger en termes de registre et de fréquence. Selon Maillard (1989 : 34), *cela* est beaucoup plus rare que *ça* en français, aussi bien à l'oral (il relève dans son corpus oral un 1 *cela* pour 100 *ça*) qu'à l'écrit (il fait allusion aux statistiques du TLF). Dès lors, plutôt que de considérer comme les grammaires que *ça* est une variante familière de *cela*, Maillard juge plus adéquat d'inverser la perspective en regardant *cela* comme une variante « guindée » de *ça* là où ils commutent, autrement dit, comme le résultat

d'une hypercorrection donnant « l'impression d'être "à côté de la langue" », faisant « figure de traduction maladroite ou de récupération petite-bourgeoise du parler commun » (*ibid.*).

Il convient néanmoins de relativiser les chiffres donnés par Maillard, qui comprennent les emplois à la fois conjoints et disjoints de *ça*, alors que *cela* est exclusivement disjoint (l'auteur signale d'ailleurs que *cela* ne fonctionne jamais avec une valeur impersonnelle comme *ça* dans *ça bruine ce matin*). Concernant *ceci* et *cela*, Maillard (*ibid.* : 35) invoque à nouveau une différence sensible de fréquence à l'oral (1 *ceci* pour 10 *cela*). Comme pour *cela*, il situe l'emploi de *ceci* dans des contextes de français surveillé (*ibid.* : 35). Pour Corblin (1987b), *ceci* présente des contraintes d'emploi plus grandes que *ça* et *cela* ; bien que partageant avec ces derniers l'absence d'un trait de classification, il s'en distinguerait par la nécessité de référer à un objet délimité. Dans OFROM, une requête rapide<sup>193</sup> fournit 24 *cela* (2 en locutions) pour 10 *ceci* (8 en locutions). La même requête dans CFPP livre 33 *cela* (9 en locutions) pour 32 *ceci* (31 en locutions, principalement dans *ceci dit*). Pour l'écrit, une recherche sur Frantext de tous les textes postérieurs à 1990 donne 20'000 *cela* (19'143 en excluant *cela dit* et *cela étant*) pour 1544 *ceci* (1470 sans *ceci dit* et *ceci étant*). Une étude plus approfondie de la distribution de ces pronoms, tenant compte des spécificités syntaxiques (conjointes ou disjointes) et diaphasiques des expressions mériterait d'être menée sur un corpus à plus grande échelle pour compléter les résultats de Maillard.

### 5.2.3.2 Les pronoms l'un... (l'autre), les uns... les autres

L'identité des référents respectivement désignés par les formes *l'un...* (*l'autre*) et *les uns...* *les autres* n'est pas toujours aisément inférable. C'est le cas ci-dessous avec l'emploi de *l'un*, utilisé seul<sup>194</sup> :

(243) Dans la cuisine, il trouva trois de ses compagnons de beuverie, préparant du café et des toasts, et *l'un* même, plus intrépide, un petit déjeuner complet, tous arborant des traits décomposés et des yeux anormalement brillants. (J. Coe, *La maison du sommeil* < Schnedecker 2003 : 99)

La forme pronominale *l'un*, comporte des instructions en apparence contradictoires dans sa constitution morphologique qui combine « une forme associée habituellement à l'indéfinitude, *un* » et un déterminant défini (Schnedecker 2000). Elle requiert d'instancier sa variable avec un objet, ici présenté comme quelconque, extrait de la classe préalablement activée (*trois de ses compagnons de beuverie*).

Dans l'exemple suivant, l'emploi des corrélat *les uns... les autres* relève manifestement d'une stratégie de cryptage et de respect du politiquement correct :

---

<sup>193</sup> Au 15.02.2016.

<sup>194</sup> Cela semble beaucoup moins fréquent avec *les uns* (Schnedecker 2003 : 99), presque toujours suivi d'un corrélat, en général *les autres*.

(244) [le ministre belge Didier Gossuin s'exprime sur un congrès organisé par son parti sur le thème de l'éthique en politique, suite à une affaire de conflits d'intérêts] Je n'irai pas à ce congrès car j'aurais le sentiment de devoir sauver la face et crier en chœur « nous ne sommes pas tous des pourris ».

Quelques crapules, issues du sérail, nous ont déshonorés et ont contribué à miner notre démocratie. *Les uns*, avec leurs titres et fonctions, ont frayé sans vergogne avec les pires mafieux, corrupteurs et corrompus de notre société. *Les autres* ont usé de leurs mandats pour extorquer à leur seul profit la collectivité et le citoyen. Pour *les uns* comme *les autres*, la seule motivation fut l'enrichissement sans cause. (12.02.2017, <http://www.lesoir.be>)

Le locuteur, dans un « coup de gueule », introduit un ensemble explicitement sous-déterminé (*quelques crapules*), qu'il partitionne en deux sous-groupes via *les uns* et *les autres*, dont sont prédiquées certaines propriétés mais dont l'identité reste délibérément implicite.

#### 5.2.4 L'adverbe là

Nous terminons l'inventaire par l'emploi d'une forme sous-spécifiée, à l'origine dédiée à l'expression de l'espace, à savoir l'adverbe *là* (cf. *supra* (225)). L'adverbe *là* constitue en quelque sorte le pendant tonique des pronoms régimes locatif *y* et ablatif *en*. Dans les grammaires, on oppose *là* à *ici*, relation tantôt envisagée comme équipollente (éloignement vs proximité), privative (en faveur du trait de proximité pour *ici*,) ou encore synonymique (Smith 1995). Les linguistes admettent généralement la seconde hypothèse, à savoir celle qui considère *là* comme le terme non marqué de l'opposition (Barbérís 1987 : 36)<sup>195</sup>. On sait qu'*ici* et *là* peuvent s'affranchir d'un sens purement spatial<sup>196</sup> pour se doter d'un trait en lien avec l'énonciation ou la subjectivité, où *ici* marque l'engagement locuteur (Smith 1995) ; à l'inverse, *là* est au « chômage déictique » : « sa référence spatiale est floue, voire inexistante, et il ne sert plus qu'à indiquer une absence d'engagement (ou le peu d'engagement) de la part du locuteur » (*ibid.* : 52). Dans le même ordre d'idée, Barbérís (1998 : 30) attribue à *ici* un mouvement « centripète », « d'appropriation purement individuelle de l'espace » saisi par le locuteur sur le vif, tandis qu'elle associe à *là* un mouvement d'expansion « centrifuge », engageant une « relation interpersonnelle » évoquant un espace plus vague et déjà partagé. A ses yeux, ces caractéristiques respectives mettent en lumière une différence d'interprétation entre les énoncés suivants, malgré leur ressemblance :

(245) pasque : depuis le temps que chuis *là* :: (dans le quartier) euh :: / sans les connaître (les habitants) je les connais d' vue : / (Boulanger < Barbérís 1998 : 31, les contenus entre parenthèses sont des explicitations de l'auteur)

---

<sup>195</sup> Smith (1995) montre qu'aucune hypothèse n'est complètement satisfaisante, les données révélant des contextes respectivement communs aux deux adverbes, des contextes où *ici* ne se laisse pas substituer par *là* et des contextes incompatibles.

<sup>196</sup> A cet égard, Smith (1995 : 52) explique l'émergence de *là-bas* dans le système pour récupérer la composante spatiale d'éloignement désormais absente pour *là*. Sur le sens uniquement spatial de *là-bas*, voir Brault (2001).

(246) comme y a vingt ans que j' suis *ici* (dans le quartier) j'ai du mal à faire du changement hein ? (rire) (Boulangier < *ibid.*)

Dans le premier énoncé, la locutrice envisage son lieu de résidence dans sa relation aux autres, à l'environnement ambiant, alors que dans le second énoncé, elle le présente comme privé, comme « son espace égotique » (*ibid.*)

D'après nos observations, *là* manifeste davantage de possibilités référentielles qu'*ici*. Ainsi, il est capable de fonctionner de manière résomptive :

(247) Mais aussi par la gestion intelligente d'un personnage nouveau, le chœur, représentant le peuple de Marseille, qui fait plus que commenter l'action, qui y participe. Grâce aussi aux magnifiques décors de Dominique Pichou et aux costumes de Christian Gasc, l'époque est évoquée sans être reconstituée à l'identique. Ce sont sans doute *là* les clés de cette réussite visuelle. (<http://www.concertonet.com>, 09.04.2007 < Corpus français, Université de Leipzig)

Cet extrait provient d'une critique d'un opéra-comique dans laquelle la journaliste énumère les qualités de la pièce, à divers niveaux (structure, musique, scénario, personnages, décors, etc.). Difficile dès lors de cerner clairement l'étendue de l'adverbe *là*. Situe-t-il la réussite uniquement dans le rôle des décors tout juste mentionnés ? Comprend-il la fonction originale du chœur ? Recouvre-t-il l'ensemble des aspects abordés jusqu'à ce point (ou davantage ?) ? Il semble vain de s'orienter exclusivement vers l'une de ces solutions. L'interprète se contentera vraisemblablement d'une représentation globale et positive des divers facteurs de réussite évoqués.

Dans les genres improvisés, il arrive fréquemment que *là* ne fonctionne plus comme circonstant mais comme simple particule à fonction phatique (Barbérís 1987, Smith 1995) en particulier lorsqu'il intervient en fin de syntagme. Barbérís (1987) appelle cet emploi « *là* de clôture » :

(248) non j'ai j'ai des nouvelles de temps en temps pis | \_ | ah oui pis tu sais ça s'est réglé l'histoire *là* tu sais que | \_ | tu sais comme quoi j'avais pas répondu au message | message tu sais | ah c'est vrai | ouais on a parlé euh ben c'était euh | \_ | \_ | ben quand t'avais pas pu venir à cause du souper justement (ofrom)

Dans cet exemple, *là* ne redonde pas par simple coréférence sur le référent auparavant activé, en l'occurrence l'« histoire ». Son emploi s'interprète de manière éminemment intersubjective : il convoque un espace commun à l'expérience des interlocuteurs. A cette fin, la locutrice « mime » (*ibid.* : 37) l'existence d'un univers partagé à reconstituer, elle fait appel à la mémoire de son interlocutrice et recherche de sa part un signe d'approbation (*cf.* aussi le *tu sais*), qui finit d'ailleurs par arriver à travers *ah c'est vrai* de l'interlocutrice. D'après Smith (1995), c'est la sous-spécification de *là* qui lui permet d'acquiescer un tel fonctionnement discursif.

## 6. Conclusion

L'objectif de ce chapitre était de montrer que la sous-détermination est tout à fait courante dans les échanges, même lorsqu'elle n'est pas annoncée comme telle (par opposition aux expressions dites « indéfinies ») : l'emploi de certains désignateurs révèle qu'il est parfaitement possible de référer à un objet qui n'a pas de propriétés distinctives et que cette situation répond à des motivations discursives diverses, accidentelles ou stratégiques, qui se laissent en tout cas difficilement expliquer par une quête de vérité.

A travers l'inventaire des moyens linguistiques entrepris, nous avons distingué les ressources lexicales des ressources non lexicales. Ces dernières n'impliquant pas, contrairement aux premières, un acte de dénomination explicite, elles nous semblent susceptibles d'aller plus loin dans le phénomène de sous-détermination. Pour cette raison notamment, elles se montrent particulièrement fréquentes dans les contextes de production spontanée. Nous souhaitons par la suite porter une attention particulière à ces contextes, dans lesquels les locuteurs produisent leur discours sur le vif et sans possibilité de le modifier, avec tous les phénomènes d'instabilité en M que cela suppose. Ces conditions nous apparaissent en effet les mieux à même de refléter les comportements effectifs des usagers, affranchis de certaines contraintes propres à des genres dominants plus « consensuels » où le critère du coût de décodage est beaucoup plus important.

Dans les limites de ce travail, il nous est impossible d'examiner dans le détail le fonctionnement de l'ensemble des formes non lexicales abordées ci-dessus. Nous avons décidé de consacrer une première étude de cas au fonctionnement du syntagme *tout ça*, particulièrement prolifique dans les contextes d'oral en configuration d'énumération, où il reflète une gamme remarquable de stratégies discursives. La seconde étude porte sur l'emploi du pronom clitique de 6<sup>e</sup> personne *ils* exprimant un actant sous-déterminé, qui se révèle difficilement conciliable avec les descriptions en vigueur de la 3<sup>e</sup> personne en termes d'anaphore et de continuité référentielle. En sélectionnant ces champs d'étude particuliers, nous sommes consciente de laisser de côté un certain nombre de chantiers ouverts qui mériteraient une attention semblable et une exploration d'envergure sur la base de données authentiques. C'est notamment le cas de la référence des clitiques régimes *le*, *en*, *y*, où de l'adverbe *là*. Nous avons néanmoins l'intention de remettre ces questions à l'ordre du jour dans le cadre de projets ultérieurs.



## **Deuxième Partie**

### *Manifestations de la sous-détermination : deux études empiriques*



## *Avant-propos : la constitution des données*

Le choix des deux objets d'étude (*tout ça* et *ils*) de cette seconde partie tient avant tout à l'intérêt qu'ils ont suscité chez nous en tant qu'usagère du français en immersion dans les échanges ordinaires. En effet, nous avons été frappée au quotidien par la productivité de ces deux procédés de sous-détermination, productivité confirmée par les données examinées par la suite. Il nous a dès lors paru pertinent d'explorer plus avant ces ressources pour corroborer l'impression que celles-ci n'étaient pas seulement le fait d'une ignorance du locuteur, mais qu'elles participaient pleinement à l'organisation du discours.

La première étude, sur *tout ça* (Ch. IV), complète les travaux déjà abondants sur le pronom *ça* (e.g. Porquier 1972, Maillard 1985, 1989, 1994, Cadiot 1988, Corblin 1991, Boivin 1992, Carlier 1996, Guerin 2006, Sales 2008, etc.) et peut être mise en perspective avec les travaux sur les particules d'extension de liste (e.g. Ward & Birner 1993, Overstreet 1999, Ferré 2011, Secova 2014). Quant à la seconde étude, sur *ils* (Ch. V), elle interroge la fonction anaphorique que l'on associe invariablement au pronom de 3<sup>e</sup> personne et suggère de rapprocher certaines constructions avec *ils* d'autres procédés de sous-détermination de l'agent, tels que l'usage du *on* dit indéfini (e.g. Muller 1970, François 1984, Viollet 1988, Leeman 1991, Rabatel 2001, Blanche-Benveniste 2003, Guerin 2006, Fløttum *et al.* 2007, Béguelin 2014c, etc.) ou encore la diathèse passive (e.g. Berrendonner 2000, Muller 2000, Blevins 2003, Creissels 2006, etc.).

Concernant les données analysées, hormis celles issues de travaux antérieurs que nous discutons dans les parties préliminaires, nous nous sommes principalement tournée vers des corpus de français parlé en accès libre (cf. *supra Introduction générale*), à savoir la base de données OFROM<sup>197</sup> (env. 69h de français parlé en Suisse romande, voir Avanzi, Béguelin & Diémoz 2012-2015), la base PFC<sup>198</sup> (une trentaine de points d'enquête dans toute la francophonie, voir Durand *et al.* 2002, 2009) ainsi que le corpus CFPP2000<sup>199</sup> (environ 48h de français parlé à Paris, voir Branca-Rosoff, Fleury, Lefeuvre, Pires 2012). Nous avons conservé, pour les exemples retenus, les conventions de transcription propres à chaque corpus, une tâche d'uniformisation s'avérant trop laborieuse. Si les moteurs de chaque base de données ont permis d'extraire aisément les séquences concernées, un traitement manuel s'est néanmoins révélé indispensable par la suite, conformément à nos principes d'analyse. En effet, en plus des critères purement segmentaux, nous accordons une large place aux facteurs contextuels. Pour ce faire, nous avons régulièrement eu recours à la lecture simultanée

---

<sup>197</sup> <http://www.unine.ch/ofrom>.

<sup>198</sup> <http://www.projet-pfc.net>.

<sup>199</sup> <http://cfpp2000.univ-paris3.fr>.

texte/son d'enregistrements intégraux<sup>200</sup> à disposition dans les bases, afin de tirer parti d'un empan textuel maximal pour chaque occurrence examinée.

Dans l'impossibilité de traiter de la sorte les dizaines de milliers de résultats, nous avons orienté notre recherche sur des critères qualitatifs en privilégiant les exemples qui présentaient une réelle valeur heuristique. Bien que nous fournissions ponctuellement quelques indications chiffrées, nous ne prétendons donc pas à un panorama exhaustif des items étudiés.

D'autres raisons nous poussent à renoncer à la délimitation d'un sous-corpus : d'une part, les trois bases de données exploitées sont des corpus ouverts, c'est-à-dire qu'elles sont régulièrement alimentées de nouvelles données, dont nous n'avons pas voulu nous priver. D'autre part, à l'heure actuelle, malgré des projets de grande envergure en cours dans ce sens<sup>201</sup>, il n'existe pas encore de vaste corpus de français échantillonné et équilibré disponible qui permette de mener à bien des analyses quantitatives exhaustives<sup>202</sup> (Benzitoun & Cappeau 2010).

Signalons pour terminer que nous avons relevé ponctuellement des occurrences dignes d'intérêt au gré de nos conversations ou lectures ou puisées dans d'autres genres discursifs (SMS, Frantext, etc.), que nous avons jugé bon d'intégrer aux analyses, pour établir le cas échéant des parallèles révélateurs. Cela nous donnera l'occasion de dégager quelques différences de conditions de production entre les genres de l'« immédiat » et ceux de la « distance » communicative (Koch & Oesterreicher 1985).

---

<sup>200</sup> Par « intégralité », nous entendons la totalité de l'extrait mis à disposition en ligne et non l'intégralité de l'enregistrement d'origine, dont souvent seule une partie est audible et a été transcrite.

<sup>201</sup> Par exemple le projet CRFC (Siepmann, Bürgel & Diwersy 2016) pour le français en général, consultable à partir de fin 2018, et le projet ORFEO pour le français parlé (<http://www.projet-orfeo.fr>), disponible en principe courant 2017.

<sup>202</sup> A noter que la base PFC comporte régulièrement des problèmes de doublons dans les résultats.

## Chapitre IV

### Fonctionnement référentiel et pragmatique de *tout ça*

- *Sacré Valentin. Alors raconte-moi. Le mariage. Le commerce. Tout ça. Ça gaze ? Ça boume ?*
- *Ça va, ça va.*
- *C'est vague : ça va. (Queneau R., *Le dimanche de la vie*, 1951)*



## 1. Introduction

Outre le caractère référentiellement évasif du syntagme *tout ça*, celui-ci à la particularité de manifester des propriétés sémantiques en apparence contradictoires : *tout*, marquant une forme d'exhaustivité, associé à *ça*, pronom largement sous-spécifié. Ce paradoxe, que nous tenterons de résoudre, soulève la question de la possibilité d'envisager le parcours exhaustif d'un ensemble sous-déterminé. Cette problématique sera en particulier étudiée à travers la disposition du marqueur à apparaître dans les contextes énumératifs.

Nous débutons par une synthèse des emplois du pronom *ça*, sur la base des nombreux travaux antérieurs sur le sujet et de nos propres observations : nous illustrons ainsi sa maniabilité, qui se reflète à travers la gamme de référents auxquels il est susceptible de s'appliquer (§2). Nous étendons le champ d'étude à *tout ça*, dont l'inconsistance sémantique évoquée ne se révèle qu'apparente, en répertoriant ses différents contextes d'emploi (§3). En clôture de liste, *tout ça* se distingue des autres emplois par son affinité avec les contextes d'oral spontané : nous nous penchons donc pour finir sur son fonctionnement référentiel dans ce genre très « immédiat » et sur son aptitude à être au service d'un éventail de rendements pragmatiques (§4).



## 2. Les emplois de *ça*

Le pronom *ça* appartient morphologiquement à la catégorie des démonstratifs, par définition token-réflexifs et possédant, aux côtés de *ceci*, *cela*, *celui-ci*, *celui-là*, une distribution de pronom disjoint (tonique). En même temps, il fonctionne aussi comme un pronom clitique sujet et entre à ce titre dans le paradigme des pronoms conjoints de 3<sup>e</sup> personne aux côtés de *il/elle/on*. Il importe donc de distinguer deux formes homonymes du pronom *ça* (Creissels 1995, Sales 2008), l'une conjointe, où *ça* fonctionne exclusivement comme sujet, l'autre disjointe, où *ça* peut occuper toutes les positions dévolues aux SN (*cf. supra* Ch.II §2). Comme tout pronom, *ça* se caractérise par l'absence de contenu lexical et donc par un sens particulièrement sous-spécifié, davantage encore que le pronom conjoint *il/elle*, car il s'en distingue par l'absence de tout trait de catégorisation (Corblin 1991, Kleiber 1994b, Carlier 1996, Guillot 2006). Au niveau sémantico-référentiel, ces caractéristiques rendent son emploi très peu contraint et lui permettent de se comporter de manière particulièrement polyvalente : on lui impute la capacité de référer à « n'importe quel type de segment de réalité » (Carlier 1996), que les objets désignés prennent leur source dans la situation immédiate ou dans le contexte verbal.

Mais son fonctionnement va au-delà : certains attribuent également à *ça* clitique un emploi impersonnel (Maillard 1989, Guillot 2006), comme dans *ça craint dans ce quartier*. Dans les études qui l'envisagent, selon nous à tort<sup>203</sup>, comme un « substitut » (Porquier 1972, Dubois 1965, Maillard 1989), il pose inévitablement des difficultés de délimitation et de restitution référentielle. En effet, on relève sa capacité à mettre en œuvre des références diversement caractérisées comme *indistinctes* (Corblin 1991), *vagues* (Bartning 2006), *propositionnelles* (Cadiot 1988) ou *indéfinissables* (Carlier 1996), parfois accompagnées d'un effet métonymique (Corblin 1991, Willems 1998). Sa souplesse référentielle lui vaut en particulier une grande productivité dans les désignations qui posent en général des difficultés de description : la référence à des objets de type procès (que la tradition sémantique sous-catégorise en événements, faits, propositions, situations, etc. *cf. supra* Ch.I §3.2), la référence générique, la deixis textuelle (*cf. supra* Ch.I §4.3), etc., dont nous fournirons des illustrations.

De nombreux travaux sont consacrés à la description de *ça*, qu'il s'agisse de monographies (Maillard 1989, Boivin 1992, Guerin 2006, Sales 2008) ou d'articles scientifiques (Porquier 1972, Cadiot 1988, Corblin 1991, Maillard 1994, Carlier 1996, Guerin 2007, etc.). A partir des résultats de ces nombreuses recherches, nous proposons une vue d'ensemble des principaux emplois de *ça* (y compris la variante-combinatoire ou stylistique – *ce* de la forme clitique, Maillard 1989 : 32) pour illustrer sa flexibilité référentielle à la faveur de sa sous-spécification. Cet inventaire ne se veut pas exhaustif<sup>204</sup>, aussi laissons-nous de côté un certain nombre d'emplois dans des contextes spécifiques où le pronom est (quasiment) référentiellement démotivé et qui nécessiteraient chacun un examen particulier : les clivées

<sup>203</sup> *Cf.* la critique de la conception substitutive *supra* Ch.II §3.1.

<sup>204</sup> Pour un inventaire de tous les emplois attestés, nous renvoyons aux monographies indiquées ci-dessus.

(*c'est ... qui/que*), le dispositif temporel *ça fait ... que*, les pseudo-clivées *ce qui/que... c'est*, les redoublements de CV (*ça... de Vinf/que/si CV*, eg. « ça fait du bien de parler »). Nous écartons également les cas où *ça* ne fonctionne pas comme un pronom mais se rapproche d'une particule ou d'un adverbe d'énonciation (Sales 2008)<sup>205</sup> et peut se combiner à des éléments exclamatifs comme dans *ça alors! ça par exemple, ah ça !*, etc. En effet, la prise en compte de tous ces éléments dépasserait l'enjeu de ce chapitre.

## 2.1 *ça conjoint*

### 2.1.1 *ça* pointe sur un objet catégorisé

Malgré son défaut de trait de catégorisation, *ça* peut servir à référer à un objet qui a déjà été catégorisé par ailleurs dans le discours. Nous nous intéressons ici à deux situations (qui ne s'excluent pas mutuellement) : d'abord le cas où l'objet a été préalablement catégorisé via un SN détaché, que vient « redoubler » le clitique *ça* ; ensuite le cas où *ça* est impliqué dans la référence humaine.

#### 2.1.1.1 *ça* redouble un sujet

La situation de redoublement est illustrée par les exemples suivants :

(249) [à propos du besoin en eau des fleurs] ce brouillard **ça** mouille pas mal et pis | ouais | pis | pis maintenant ben | \_ | \_ | la terre était aussi pas assez sèche (ofrom)

(250) [le locuteur raconte les difficultés rencontrées en tant que pilote d'avion] à Londres y avait vraiment un vent très très fort tempétueux en | plus de de travers qui était pas dans dans l'axe de la piste | donc c'était très difficile d'atterrir des des choses comme ça la neige **ça** peut être difficile mais surtout les orages et le vent je dirais (ofrom)

A la différence d'un pronom personnel *il/elle* qui présente l'objet de manière individuée, on remarque avec *ça* une désignation plus floue du référent, donné à voir plutôt comme un phénomène (ici météorologique) diffus, non comptable, que comme un individu bien délimité. Dans l'exemple (250), *ça* peut renvoyer de manière plus générale à l'expérience du locuteur dans laquelle l'objet mentionné entre à titre d'ingrédient, glosable (non exhaustivement) par 'piloter par temps de neige', 'atterrir par temps de neige/sur la neige', etc. (Cadiot 1991). Outre le SN détaché, c'est surtout la prédication attribuée à *ça* qui joue un rôle considérable dans la manière d'envisager le référent. Il est à remarquer que le sujet lexical peut être instancié à droite du verbe, autrement dit que la dénomination peut être ultérieure à la désignation pronominale par *ça* :

(251) si si ils annoncent encore un petit peu beau | \_ | \_ | **ça** doit commencer la foire euh la foire d'automne à Bâle (ofrom)

<sup>205</sup> Sales relève des zones de recouvrement entre *ça* pronominal et *ça* adverbial, où l'on peut hésiter entre les deux analyses : Qui *ça*, « lui » ? (*ibid.* : 229)

Dans ce cas aussi, *ça* permet, au contraire du pronom *elle*, de présenter l'objet sous une forme plutôt globale (Cadiot 1991) qu'individuée, malgré l'existence d'un nom tout désigné identifiant l'événement en question (*la foire d'automne à Bâle*). Une autre interprétation possible est que l'étiquette lexicale n'est pas encore disponible (lacune lexicale) pour le locuteur au moment de la désignation par *ça* (cf. le *euh* d'hésitation). En effet, on a déjà vu *supra* (Ch.III §4.1) que *ça* est particulièrement propice en situation de recherche lexicale.

Le pronom conjoint *ce/ça* peut procéder à des références d'ordre métalinguistique, comme ci-dessous où *c'* renvoie dans un premier temps au signe *typewriter* :

(252) premièrement j'ai conçu un site internet en fait qui s'appelle typewriters of art [...] typewriter **c'**est non **c'**est pas un choix qui est anodin | \_ | le typewriter **c'**est une figure du street workout vous verrez tout à l'heure | \_ | et en même **ça** fait référence euh au processus ethnographique et à l'activité d'écriture (ofrom)

En effet, le choix en question concerne l'occurrence autonymique *typewriter* (Authier-Revuz 2003). Mais la deuxième occurrence du pronom (*c'est une figure...*) pointe quant à elle sur le type (Berrendonner 2002) introduit par le SN défini *le typewriter*. Le prédicat attribué à *c'*, consiste à fournir une caractérisation (*est une figure du street workout*) au type en question. Il sert ainsi de relais entre l'objet et une « définition » à caractère identificatoire (Burston & Monville-Burston 1981). On peut enfin remarquer que le *ça* qui prolonge ultérieurement la référence (*ça fait référence*) ramène la référence au signe, au vu de la prédication qui lui est assignée : *c'* est bien au nom *typewriter* qu'on attribue la propriété de faire référence à tel et tel aspect. Cet exemple illustre la flexibilité du pronom *ça*, qui se révèle totalement adapté aux situations de reformatage d'objet (ici entre le type et le signe).

#### 2.1.1.2 *ça* et la référence humaine

Contrairement à certaines idées reçues, *ça* n'est pas réservé à la référence [- animé] :

(253) si je regarde avec mon frère | \_ | lui il avait une phase | \_ | où il parlait seulement français | \_ | pis les Suisses allemands **ça ça** existait pas pour lui | et maintenant il a la phase seulement suisse allemand (ofrom)

Dans cet exemple, *ça*<sup>206</sup> ne renvoie pas aux usagers suisses-allemands pris dans leur individualité comme pourrait l'induire un *ils* distributif, mais il engage plutôt une interprétation collective de l'ensemble évoqué (*les Suisses allemands*) (Carlier 1996), voire une référence au type auquel cette classe ressortit. Cette affinité avec la généralité est souvent observée dans les études sur *ça*, où le SN détaché peut prendre diverses formes (*un zizi, ça sert à faire pipi debout, la pie/les pies ça jacasse* < Maillard 1989, Carlier 1996).

<sup>206</sup> L'écoute de l'enregistrement confirme que le redoublement est dû à un phénomène de bribe involontaire plutôt qu'à une véritable structure de détachement à gauche d'un premier *ça* tonique, redoublé d'un *ça* clitique. C'est pourquoi nous considérons qu'il y a simple piétinement d'un *ça* conjoint sur la position sujet.

Mais *ça* impliqué dans la référence humaine n'apparaît pas exclusivement dans les contextes génériques. Ci-dessous, *c'est* est utilisé pour pourvoir un individu humain spécifiquement nommé (cf. le #) d'un attribut typant « en le rangeant dans la classe des objets ou des êtres qui partagent ce type » (Boone 1987 : 96)<sup>207</sup> :

(254) | j'ai connu | # | \_ | qui était euh qui donnait des des cours de batterie **c'est** un batteur donc (ofrom)

Pourquoi recourir à *ce* lorsque l'individu en question est déjà par ailleurs identifié (via le nom propre masqué) ? Une alternative serait l'emploi d'une structure *il est N (il est batteur)* ; néanmoins, celle-ci n'exprimerait plus un processus de classification, mais « l'attribution au sujet des propriétés définitoires du type N1 » (Riegel 1985 : 198). Pour Cadiot (1988 : 177), la structure en *c'est un N* invoque « la mise en œuvre d'un décrochement entre une référence assurée et une référence à effectuer ». Le pronom conjoint *ce* apparaît comme l'instrument de relais idéal pour faire entrer l'individu dans une catégorie donnée (*un batteur*) (cf. aussi *supra* (252)).

L'exemple suivant illustre un cas de référence humaine où l'on peut hésiter entre deux interprétations :

(255) [le locuteur répond à une question sur la taille de son neveu de 2 mois] il est il est il non non donc pour son âge il est assez grand justement | \_ | bon il est né à cinquante centimètres ce qui est pas mal | \_ | pis là je sais pas combien il fait mais euh **ça** grandit à vue d'œil là on a fait le weekend de Pâques en Provence | \_ | ouais | \_ | et je l'ai pas revu depuis toute la semaine j'ai vu il a changé quoi en une semaine (ofrom)

On peut d'une part considérer que la prédication *ça grandit* concerne une classe prise dans son ensemble, celle des nourrissons, exprimant ainsi un commentaire à caractère générique. D'autre part, on peut interpréter le pronom comme renvoyant spécifiquement à l'enfant en question<sup>208</sup>, mais avec une appréhension particulière : l'aspect token-réflexif, simulant la présence effective du référent, combiné au trait non individué de *ça* (mise en avant du processus plutôt que de l'individu ?, marquage du caractère « inachevé » du bébé ? etc.) contribuent à produire un discours empreint de subjectivité (attendrissement).

On peut noter à cet égard que les grammaires ne relèvent souvent que l'effet péjoratif de *ça* pour la référence humaine (Riegel *et al.* 2009 : 377), comme dans :

(256) **ça** court partout, **ça** cavale avec tous les gars du coin, **ça** s'fait engrosser à tour de bras ! (Maillard 1989 : 60)

Mais les divers exemples de cette section montrent bien que ce n'est qu'un effet possible.

---

<sup>207</sup> La différence avec *il est batteur* semble être qu'avec ce dernier, on attribue simplement une propriété au référent en question, mais non une catégorisation à visée identificatoire (cf. Burston & Monville-Burston 1981).

<sup>208</sup> Cf. l'emploi de *it* en anglais ou *es* en allemand pour désigner un bébé dans des contextes spécifiques.

### 2.1.2 *ça* pointe sur un objet non catégorisé :

Le pronom *ça*, par définition, est particulièrement adapté à la référence aux objets auxquels on n'a pas explicitement octroyé de classification nominale. Nous en présentons ci-dessous les types de référence majeurs.

#### 2.1.2.1 *ça se rapporte à l'expérience du locuteur*

Ces objets peuvent avoir pour origine la situation environnante, comme l'incident rapporté ci-dessous par le locuteur :

(257) [un plat vient de glisser des mains d'un essuyeur de vaisselle] **ça** a glissé (Maillard 1989 : 68)

Le clitique *ça* se révèle dans cette situation particulièrement économique : il est utilisé par défaut et n'implique pas, contrairement à *il*, d'étiquette lexicale sous-jacente. Il révèle à ce titre que le locuteur n'a pas assigné à l'objet un attribut de dénomination particulier.

Dans le domaine des sens, *ça/ce* peut renvoyer à une cible diffuse, se mêlant à l'expérience sensorielle proprement dite :

(258) Mmh **c'**est bon (sur un bavoir pour bébé)

Pour référer indistinctement à (un aspect de) la situation courante, c'est à *ça* qu'on recourt :

(259) **Ça** suffit / **ça** commence à bien faire (*ibid.* : 71)

A cet égard, il devient difficile de distinguer le référent sur lequel opère la prédication : l'énoncé dans son ensemble exprime globalement l'exaspération du locuteur vis-à-vis d'une situation.

#### 2.1.2.2 *ça* « impersonnel »

De la même manière, il apparaît impossible de restituer l'agent du procès invoqué dans les contextes ci-après, *i.e.* impliquant les domaines respectivement physiologique, climatique ou interactionnel :

(260) [à propos de processus abdominaux involontaires] **Ça** gargouille (*ibid.* )

(261) **Ça** caille, ce matin !/ **ça** pleut (*ibid.*)

(262) **Ça** boume/**ça** baigne /**ça** gaze (*ibid.*)

Maillard (1985, 1989, 1994a) propose de traiter ces cas comme des impersonnels, ou plutôt, comme des SV « asubjectaux », au vu de l'impossibilité d'y cliver le pronom *ça*, (*c'est ça qui boume*) ou d'instancier la place d'argument sujet (par exemple par le pronom tonique *cela*, cf. Sales 2008 : 115). A ce titre, ils contribuent sémantiquement à une « baisse d'agentivité et à une déthématisation de l'agent » : le procès se voit dès lors « mis en vedette au détriment de

ses participants » (Maillard 1994a : 5). Corblin (1991) leur refuse néanmoins le statut d'impersonnel, considérant que *ça* n'y est pas complètement explétif mais qu'il possède tout de même un contenu référentiel, aussi indistinct soit-il. Il note à cet égard la capacité d'« expansion métonymique » des constructions en *ça+V* :

(263) **Ça** dort, là-dedans. (*ibid.* : 46)

Dans cet énoncé, le verbe sélectionne un trait humain imputé de ce fait à un référent indistinct : « le sommeil semble diffusé dans la situation, et n'être plus une propriété d'éléments discrets et classifiés » (*ibid.*). Selon Maillard, c'est au final chaque emploi verbal qui est responsable d'une lecture référentielle ou non du pronom :

C'est **le sens du verbe en contexte** qui vide ou non de sa capacité référentielle le morphème préverbal, qu'il s'agisse du **es** allemand, du **it** anglais, du **il** ou **ça** français. A priori, tous ces morphèmes sont présumés référentiels – exophoriques ou endophoriques – et seul tel prédicat particulier a le pouvoir de les réduire à l'état de formes postiches. (Maillard 1994a : 4)

Il faut reconnaître que la limite est ténue entre *ça* à référence indistincte et *ça* aréférentiel. S'il est clair que pour (261), *ça* ne recouvre aucun agent, on peut à l'inverse admettre qu'il renvoie à l'origine (imprécise) des gargouillements en (260). Ou plutôt, on pourrait considérer dans ce dernier cas que l'exemple fait l'objet de deux analyses concurrentes, à savoir une interprétation référentielle indistincte et une interprétation aréférentielle, où *ça* est employé comme forme postiche pour mettre en avant un procès sans agent. Maillard remarque d'ailleurs que *ça*, contrairement à *il* impersonnel, permet de créer de nouveaux prédicats asubjectaux :

[...] *ça* ne cesse de s'accoupler avec des formes verbales nouvelles (*ça boume, ça cartonne, ça bouchonne...*) tandis que *il* est aujourd'hui frappé de stérilité. (1989 : 37)

### 2.1.2.3 *ça* renvoie à un procès ou à un objet « indiscret »

Parmi les références à des objets non catégorisés figure évidemment en bonne place le renvoi aux procès, que la langue peut exprimer au moyen de constructions verbales. En effet, on a déjà vu (*supra* Ch.I §3.4) que ce type d'objets avait en commun la caractéristique de ne pas posséder de dénomination usuelle. Le clitique *ça* représente un bon moyen de référer à des objets de ce genre, et d'ailleurs le seul moyen de « redoubler » en position sujet<sup>209</sup> une construction infinitive détachée<sup>210</sup> (Cadiot 1988) :

(264) ah ouais parce qu'à | qu'à l'époque euh | \_ | marier<sup>211</sup> un étudiant ben ***ça se faisait pas*** hein enfin | \_ | à part justement si on était obligés (ofrom)

<sup>209</sup> En fait peut-être pas tout à fait, cf. la structure inversée en *-il* avec infinitif du type *marier un étudiant se faisait-il à l'époque ?*

<sup>210</sup> En position d'objet, on trouve également *le, en, y*, cf. *supra* Ch. III §5.2.2.

<sup>211</sup> Ici, d'après le contexte, au sens d'*épouser*, et non au sens de *donner un époux* (à *qqn*) ou de *consacrer une union*.

L'objet visé par *ça* n'est cependant pas toujours aussi clairement identifiable à travers le contexte linguistique. Ci-dessous, *ça* renvoie vraisemblablement à la tâche de répondre au téléphone sollicitée par le locuteur, ou, éventuellement à sa demande de faire, autrement dit à l'acte de langage rapporté (*cf.* la notion de deixis textuelle, *supra* Ch.III §4.3) :

(265) ouais il serait capable de tout faire si il s'investissait un peu plus | \_ | non c'est juste que | \_ | genre euh y a le téléphone qui sonne je lui dis ben prends | \_ | pis euh | \_ | **ça lui fait chier** ou bien euh | \_ | genre on doit aller faire un dépannage | \_ | il me dit ouais mais je sais pas comment faire (ofrom)

Le pronom *ça* peut également se rapporter à un sous-graphe de plus grande « envergure » en M, comme une série de péripéties :

(266) ben moi j'étais au stade j'ai vu j'étais juste derrière celui qui a jeté la torche | \_ | alors tu te dis super tu sais le gars super intelligent qui a tout compris | et allez tiens si on jetait une torse une torche sur le gardien adverse hein | \_ | et pis donc il balance sa torche après | \_ | comme c'est un grand courageux il va se planquer dans les gradins pis **ça a provoqué une bagarre** | entre les groupes et les ex/ les ex groupe (ofrom)

Le pronom fonctionne ici de manière résomptive<sup>212</sup> (Maillard 1989 : 61), résumant toute une séquence narrative, à propos de laquelle la locutrice prédique la conséquence (*ça a provoqué une bagarre*). La succession des énoncés peut donner l'impression que leurs contenus sont mis bout à bout dans M, ainsi engrangés au fil du discours et potentiellement disponibles pour les opérations de pointage. Mais il ne faut pas oublier que les procès ne font pas que s'accumuler : ils se combinent par diverses relations (temporelles, causales, oppositives, etc.) et peuvent aboutir à des sous-graphes inanalysables, des représentations au contenu partiellement implicite et aux contours vagues émergeant de multiples inputs.

Ci-dessous, un locuteur raconte son séjour aux États-Unis et son intégration difficile, faute de moyens linguistiques suffisants :

(267) c'est vrai que pendant | \_ | %<sup>213</sup> | \_ | un bon mois | \_ | j'ai pas vraiment compris ce qu'on me racontait | \_ | alors c'est ça peut paraître long | surtout que j'étais pas en cours en même temps que les gens | \_ | et j'étais là pour faire mon travail de diplôme donc euh | du coup euh j'avais déjà je voyais déjà peu de monde | \_ | euh la personne que je voyais le plus c'était le concierge | de l'institut | \_ | qui n'avait qu'une dent | \_ | et qui parlait euh | \_ | **c'était un peu particulier** ouais | \_ | alors je l'ai pas compris lui pendant | pendant un an et demi au moins il me parlait tous les jours euh | il venait fumer sa clope avec moi et pendant ouais | pendant un an et demi j'ai pas compris ce qu'il m'a raconté | \_ | **c'était sympa** (ofrom)

<sup>212</sup> Néanmoins, nous nous distançons de l'approche segmentale de Maillard, pour lequel ce sont des « fragments discursifs » qui sont « repris par *ça* » (*ibid.* : 61).

<sup>213</sup> Le signe % signale dans ofrom un élément inaudible.

Dans les deux prédictions en *c'était*, le pronom *ce* renvoie vraisemblablement au résultat global des éléments rapportés par le locuteur<sup>214</sup>, dont on peut inférer l'incongruité en mobilisant des compétences d'interprétation adéquates : il capte la conjonction de divers éléments, à savoir le défaut de ressources linguistiques du locuteur, sa solitude, sa rencontre avec un personnage atypique qui deviendra son compagnon malgré la barrière de la langue, etc. Le pronom *ce* permet ainsi d'évoquer grossièrement le bilan de ces ingrédients divers, qu'on ne peut ramener à une simple addition d'éléments.

#### 2.1.2.4 *ça renvoie à une quantité*

Le pronom conjoint *ça* se montre encore productif en contexte d'« estimation quantitative » (Maillard 1989 : 70), à travers des formules comme *ça fait...*, *ça donne...*, *c'est...* :

(268) les gens qui sont vraiment ce qu'on appelle des | \_ | pratiquants c'est dix-huit pour cent  
| \_ | et puis y a ce qu'on appelle les laïcs **ça fait** environ | treize quatorze pour cent  
(ofrom)

En (268), *ce* et *ça* s'interprètent comme le résultat d'un calcul, en l'occurrence, des proportions.

## 2.2 *ça disjoint*

Nous nous penchons dans cette section sur deux situations : la référence du pronom *ça* en position de complément, dont le fonctionnement référentiel est en partie semblable aux emplois référentiels de *ça* conjoint. Et nous envisageons ensuite le fonctionnement référentiel de *ça* en position détachée, où il apparaît tirer totalement parti de sa valeur token-réflexive, mise au service d'effets d'intensité ou de contraste.

### 2.2.1 *ça complément*

Dans la mesure où le fonctionnement référentiel de *ça* complément n'est pas si différent du *ça* clitique référentiel, nous nous limitons à illustrer quelques situations caractéristiques.

Outre son affinité avec les objets de type procès, *ça* complément peut également référer résomptivement à un objet indiscret, constitué de nombreux ingrédients hétérogènes :

(269) [à propos de la conduite automobile en Australie] donc euh la voiture c'est c'est quelque chose d'assez indispensable | \_ | d'ailleurs euh ils peuvent | \_ | avoir leur permis | \_ | à seize ans | \_ | mais ça pose d'énormes problèmes [...] ils ont ils avaient tendance à tous avoir des voitures de sport | \_ | et y avait enfin | \_ | un taux de mortalité dingue [...] avec les kangourous qui traversent la route rien à voir euh ça aide pas non plus | \_ | t'as meilleur temps<sup>215</sup> de pas aller trop vite | \_ | mais alors maintenant ils ont instauré des nouvelles lois | genre euh ils ont pas le droit d'avoir de trop grosses

<sup>214</sup> A noter une autre interprétation possible pour le premier *ce*, où celui-ci peut renvoyer à la manière de parler du concierge (*a fortiori* du fait qu'il n'avait qu'une dent).

<sup>215</sup> Locution utilisée en Suisse romande signifiant « avoir intérêt à ».

cylindrées | \_ | heureusement | \_ | euh | \_ | ils ont pas le droit d'être plus que deux dans une voiture après neuf heures le soir | \_ | pour pas qu'ils euh | \_ | enfin pas qu'ils aillent faire la foire euh | \_ | à plusieurs remplir une voiture y en a un qui | \_ | conduit qu'il boive ou qu'il boive pas euh | \_ | à seize ans tu rentres à une heure du matin t'es fatigué euh | t'as les copains qui te poussent à aller vite enfin c'est aussi le truc à faire des conneries | \_ | mais euh | \_ | bon voilà ils ont mis des restrictions par rapport à **ça** (ofrom)

On interprète *ça*, dans ce contexte récapitulatif (noter le *bon voilà*), comme renvoyant globalement à la situation problématique des jeunes conducteurs en Australie dont le détail a été préalablement rapporté.

Le pronom complément peut référer à un objet dont la dénomination lexicale fait momentanément défaut (cf. aussi (225)) :

(270) quand on va dans un musée | y a beaucoup de choses tandis que là | ouais | là | \_ | quand je vais voir justement ces | \_ | euh comment on appelle **ça** / les vernissages | oui | comme ça | \_ | mais | oui | bon ils sont | \_ | souvent plusieurs | \_ | ouais | y a | \_ | très très peu de choses (ofrom)

On rencontre également des *ça* compléments portant réflexivement sur l'énonciation en cours, comme dans l'exemple suivant où il réfère à l'un de ses aspects para-verbaux, à savoir la hauteur de la voix de la locutrice :

(271) le petit garçon qui chantait en soprano il est reparti de chez moi euh | \_ | avec une voix comme **ça** | \_ | et pis vraiment euh et il était très content (ofrom)

Le pronom *ça* apparaît dans une structure comparative où la locutrice, enseignante de chant, contrefait sa voix en en réduisant sensiblement la fréquence, sur le segment *avec une voix comme ça* pour imiter celle de l'élève devenu grand. A noter la productivité de la séquence *comme ça* à l'oral spontané, qui peut fonctionner comme l'expression de la simple approximation (quantitative, comme ci-dessous, ou lexicale) (Béguelin 2016, Corminboeuf 2016) :

(272) et puis ben voilà ça fait quoi une dizaine d'années douze ans **comme ça** ben | \_ | je suis partie là-dedans (ofrom < *ibid.*)

Dans cet exemple, il n'y a pas lieu de restituer analytiquement les éléments de la comparaison<sup>216</sup> : le syntagme *comme ça* sert simplement à mitiger la durée indiquée, à la manière d'un adverbe d'approximation (e.g. *environ*, *à peu près*, etc.).

Dans le domaine de l'estimation quantitative, Maillard (1989 : 70) rapporte, dans des contextes commerciaux, l'usage d'expressions confirmatives comme *c'est (bien) ça*, où *ça*

---

<sup>216</sup> Sur l'emploi référentiellement « démotivé » de *ça* dans les expressions *gros comme ça*, *haut comme ça*, etc., voir Béguelin (2016).

disjoint renvoie au montant demandé. L'expression peut également servir à « l'appréciation qualitative », par exemple d'une expérience perceptive :

(273) [un chef de chœur à ses choristes] c'est pas encore tout à fait **ça** (*ibid.*)

Le pronom clitique *c'* renvoie au résultat effectif de l'interprétation, tandis que le pronom attribut désigne le résultat souhaité, entre lesquels le locuteur signale l'écart qualitatif.

### 2.2.2 **ça détaché**

Le pronom *ça* disjoint intervient régulièrement en position détachée, où il se voit redoublé par un *ça* clitique en position sujet ou un autre pronom en position objet<sup>217</sup> :

(274) il faut euh il faut pas regarder les hommes | \_ | ou alors si on les regarde il faut pas les voir il faut avoir le regard qui passe à travers | \_ | **ça ça marche** assez bien mais | \_ | on n'a pas forcément l'habitude en Suisse (ofrom)

(275) peut-être que je faisais pas bien la vaisselle c'est un truc que je détestais faire | \_ | mais euh | \_ | mais à la base j'étais pas là pour ça j'étais là pour m'occuper de sa fille pis **ça** je le faisais très bien (ofrom)

Dans ces exemples, *ça* renvoie respectivement à un comportement (274) et à une activité (275). En tant que pronom accentué, il exploite ici pleinement sa valeur token-réflexive, responsable d'un véritable geste ostensif éventuellement accompagné d'un effet contrastif (277) (*cf. supra* sur les pronoms disjoints Ch.I §4.4), qui apparaît moins manifeste si le clitique est employé seul :

(276) ou alors si on les regarde il faut pas les voir il faut avoir le regard qui passe à travers | \_ | **ça** marche assez bien (exemple modifié à partir de (274))

(277) j'étais pas là pour ça j'étais là pour m'occuper de sa fille pis je **le** faisais très bien (exemple modifié à partir de (275))

Cet acte d'ostension marqué ne présente pas de contraintes référentielles particulières. Outre la référence aux procès illustrée ci-dessus, *ça* détaché renvoie volontiers aussi à un objet indiscret, ci-dessous, une étape d'un exposé scientifique, comme l'illustre cet exemple de Maillard (1989) :

(278) [...] alors **ça**, c'étaient nos constatations sur le terrain socio-linguistique (*ibid.* : 61)

Il est également compatible, dans cette position, avec la référence humaine :

(279) Yves Montand, **ça** c'était un chanteur ! (Leeman 1994 : 146)

---

<sup>217</sup> Nous n'abordons pas la question du statut syntaxique de ces structures détachées ou disloquées, qui, selon l'hypothèse de Berrendonner (2008), ne relèvent pas toutes du même type. Il nous suffit, pour notre propos, de comparer la présence à l'absence d'un redoublement pour mettre en évidence les différences référentielles.

(280) et pis y avait donc euh les enfants y avait | #<sup>218</sup> | \_ | la fille | \_ | et pis | # | \_ | pis **ça**  
c'était des enfants que monsieur | # | avait adoptés qui étaient à son frère

On voit à nouveau que lorsqu'on veut ranger un référent dans une catégorie ad hoc, c'est à *ça* qu'on recourt (cf. *supra* §2.1.1).

A la faveur de sa nature foncièrement ostensive, *ça* détaché est productif dans la désignation d'éléments de la situation ambiante (cf. avec les emplois clitiques *supra* (235), (258)), comme ci-dessous où la locutrice est en train de montrer les photographies d'un mariage à son allocutaire :

(281) **ça** c'est dans l'église des Invalides, **ça** c'est là à la maison ici. (pfc)

La désignation peut concerner un aspect de l'énonciation elle-même :

(282) L1 : je me suis souvent posé la question | est-ce qu'y a ce est-ce qu'on rencontre ce  
phénomène est-ce que par exemple tu lui montres un exercice | \_ | lui il aura tendance  
à il va peut-être ne pas comprendre ce qu'il doit faire corporellement mais il va essayer  
de | copier un petit peu euh le son qu'il a entendu | \_ | euh est-ce que |  
L2 : alors **ça** c'est c'est un c'est une question absolument pertinente (ofrom)

La cible de *ça*, au vu de sa prédication, constitue autant l'objet du dire que l'acte locutoire lui-même.

L'extrait suivant, qui implique un discours rapporté direct, montre que la limite traditionnelle tracée entre sources verbale et situationnelle est problématique (cf. *supra* l'exemple (180)) :

(283) [la locutrice raconte la façon dont elle a illustré sous forme de schéma les problèmes  
que se créait une amie] j'ai dit parce que tu t'es mis des bâtons dans les roues toute  
seule | \_ | oui |tu t'es fait une montagne pour rien | \_ | et pis là elle me regarde elle me  
dit oui | \_ | et j'ai dit mais **ça** | \_ | **ça** là cette montagne ce mur fais | \_ | une croix dessus  
(ofrom)

La répétition disjointe de *ça* désigne un objet déjà préalablement introduit par des moyens verbaux (*des bâtons dans les roues, une montagne*). Mais dans le discours direct reproduit, l'objet incriminé est convoqué gestuellement, il fait l'objet de six pointages successifs (*ça, ça, là, cette montagne, ce mur, dessus*) chacun accompagné de proéminences prosodiques produites avec une scansion rythmique remarquable. L'acte d'ostension est plus qu'insistant, se manifestant à plusieurs niveaux (parole, gestes, dessins). D'ailleurs, l'enregistrement capte le bruit répété (sur *croix* et *dessus*) de ce qui pourrait être le choc d'un doigt ou autre objet sur une surface. La locutrice est véritablement en train de mimer cette énonciation antérieure (cf. aussi (235) *supra*). L'interaction de ces divers paramètres contextuels (verbaux, gestuels, iconographiques) ainsi que l'imbrication des énonciations hôte et rapportée illustrent ainsi bien les difficultés qu'il y a à vouloir distinguer leur nature à tout prix.

---

<sup>218</sup> Pour rappel, signe d'anonymisation dans OFROM.

En somme, ce qui nous paraît significativement différent entre les structures à redoublement avec *ça* détaché et les structures sans redoublement<sup>219</sup>, c'est l'acte de confirmation référentielle qu'opère le geste ostensif avec *ça* détaché sur la validité du référent en M, là où la présence du seul clitique peut consister en une référence par défaut : *ça* détaché contribue ainsi à augmenter la saillance cognitive de l'objet désigné.

### 2.3 *Bilan sur les emplois de ça*

Pour la majorité des emplois de *ça*, la nature conjointe ou disjointe du pronom n'a pas de conséquence sur le fonctionnement référentiel : au vu de sa sous-spécification et du défaut de trait de catégorisation référentielle, *ça* s'illustre dans un éventail de références, que les objets soient dépourvus de dénomination *ad hoc*, appréhendés de manière non strictement individuée ou qu'ils se montrent ouverts à de nouvelles catégorisations. On peut néanmoins relever deux aspects sur lesquels pronom disjoint et pronom conjoint divergent : sauf dans des lexies particulières à pointeur démotivé, le pronom disjoint exploite toute sa potentialité référentielle ; à cet égard, en particulier en position détachée, il met en œuvre un véritable geste ostensif qui contribue à consolider la place cognitive du référent dans la mémoire discursive (*supra* §2.2.2). D'autre part, l'occurrence de *ça* dans les structures dites « impersonnelles » concerne exclusivement *ça* conjoint (*supra* §2.1.2). Dans ces emplois, la valeur référentielle de *ça* s'estompe à la faveur d'une mise en avant du procès exprimé par le verbe.

---

<sup>219</sup> A noter qu'il ne fonctionne pas, par définition, comme sujet détaché d'une construction impersonnelle ou « asubjectale ».

### 3. Les emplois de *tout ça*

#### 3.1 *Le quantificateur tout*

Dans le syntagme *tout ça*, *tout* fonctionne comme prédéterminant (Grevisse & Goosse 2011 : §638) devant *ça* disjoint, celui-ci occupant une position de SN :

*toute la compagnie*  
*tout le reste*  
*toutes les choses*  
*tout ça*

Le quantificateur *tout* exprime la totalité ou l'exhaustivité, qu'il s'agisse d'une « absence d'exception » (= lat. *omnis*) lorsque *tout* s'applique à une classe ou du caractère « entier » (= lat. *totus*) quand il concerne un individu (Grevisse & Goosse 2011 : §638). Pour rendre compte de la première conception (*omnis*), on parle généralement de quantification universelle. Kleiber (1998) formule deux conditions à l'égard de l'appréhension de la totalité par *tout* :

[...] *tout*, parce qu'il indique la totalité, nécessite d'une part, un domaine de quantification borné et d'autre part une structuration interne partitive de ce domaine. (*ibid.* : 90)

Aux yeux de Kleiber, envisager soit une entité dans sa totalité, soit des entités dans leur totalité, implique la délimitation de l'ensemble envisagé :

(284) J'ai mangé toute la saucisse (Kleiber 2011 : 143)

(285) J'ai mangé toutes les saucisses (*ibid.*)

(286) J'ai mangé toute une saucisse (*ibid.*)

(287) \*J'ai mangé toute de la saucisse (*ibid.*)

Le dernier exemple montre que le quantificateur est incompatible avec le déterminant partitif exprimant le caractère non délimité du référent. D'autre part, la totalité implique que l'ensemble se compose de parties, à savoir les occurrences de N pour une classe<sup>220</sup> (285) ou des subdivisions internes pour un individu singulier (284), (286). Enfin, Kleiber ajoute que contrairement à *tout N distributif* (e.g. dans *tout homme est mortel*), *tout SN* donne lieu à une saisie externe de la totalité, d'où une exhaustivité perçue de manière collective et non distributive<sup>221</sup>.

Il nous semble que quelques aspects de cette conception méritent d'être nuancés. La plupart des études sur *tout SN* se concentrent essentiellement sur son occurrence avec des SN définis

<sup>220</sup> Encore que l'on puisse se demander si les éléments d'un ensemble en constituent les « parties ».

<sup>221</sup> Une lecture distributive n'est pas exclue pour autant (e.g. *tous les enfants ont une voiture*), mais elle serait alors due à l'attribution d'un prédicat individuel (Kleiber 2011 : 150).

et donnent à penser que l'ensemble exhaustivement perçu comporte des éléments homogènes et dénombrables, par analogie à la notion mathématique d'ensemble. Le cas de *tout ça*, soulève à cet égard un « paradoxe » intéressant, car le marqueur se montre susceptible de renvoyer à des sous-graphes (ou réseaux référentiels) mal délimités en M. D'autres expressions avec *tout* posent également problème à cette approche :

(288) « Elle » avait aussitôt abordé l'Indienne qui rasait les murs et, après une brève lecture de son regard noir en amande, l'avait embrassée chaleureusement, en guise de bienvenue : « Tu peux me demander **tout ce que tu veux** ! » avait-elle souri sous les spots, comme si le monde était aussi grand. (Férey, C., *Mapuche*, 2012 < Frantext)

Cet exemple montre également la difficulté de dénombrer les éléments de manière exhaustive, comme ceux-ci sont en quantité manifestement indéfinie. Il en va encore de même ci-dessous, au vu de l'application de l'opérateur *tout* à une notion réputée infinie :

(289) **Tous les nombres premiers** sont impairs à l'exception de 2 (web, <http://education.toutcomment.com>, consulté le 16.09.2016)

Schnedecker (2008) signale un autre contexte où *tout* met à mal la dénombrabilité et l'exhaustivité du domaine. Il s'agit de son emploi pronominal, qui se rapproche selon l'auteur d'un terme massif, du fait qu'il se combine avec les marqueurs *ad hoc* : le partitif *de*, l'anaphore par *en*, le quantifieur *un peu de*, dont nous fournissons une illustration ci-dessous :

(290) Que nous conseillez-vous ? Un menu pour les bons clients ? Ah ! Préparez-nous l'assiette du chef, ce que vous voulez, je vous fais confiance. Avec **un peu de tout**, très bien. (Mréjen, *Eau sauvage*, 2004 < *ibid.* : 2165)

Selon Schnedecker, l'absence de N restreignant son domaine rend celui-ci « ouvert, hétérogène et massif » (*ibid.* : 2169). Elle relève également sa disposition à la récapitulation :

(291) À cause de haut-parleurs, c'est la musique qui vous attire, et puis, plus près, voilà les paroles, dialogues et bruitages, l'histoire. On entend **tout**. (Bon, *Mécanique*, 2001 < *ibid.* : 2162)

Dans ce cas, il va de soi que la référence de *tout* englobe les éléments énumérés. Mais elle s'ouvre à d'autres possibilités référentielles. Selon Schnedecker, le fait que *tout* pronominal puisse référer à des entités factuelles, contrefactuelles ou potentielles empêche la réalisation effective du parcours du domaine de quantification. La solution choisie est de considérer le parcours en question non pas comme exhaustivement réalisé, mais comme seulement virtuellement ébauché ou simulé.

Nous préférons considérer pour notre part que l'ensemble évoqué par *tout SN* est à interpréter non pas au sens mathématique, mais qu'il recouvre plus généralement une classe-objet (Grize 1990, cf. *supra* Ch.III §2.3) dont les éléments ne sont pas obligatoirement homogènes,

discrets ou dénombrables. C'est en tout cas ce que suggère l'emploi de *tout ça*, comme on va le voir.

Venons-en aux différents contextes d'apparition de la séquence *tout ça* : celui-ci peut se retrouver en relation de dépendance grammaticale avec son entourage ; on le trouve aussi en position détachée, redoublé par un pronom clitique ; il peut encore fonctionner comme apposition ; et pour finir, il apparaît en clôture d'énumération.

### 3.2 *tout ça régi*

Au niveau micro-syntaxique, il peut occuper toutes les places régies dévolues aux SN, comme ci-dessous en position de régime direct d'un verbe :

(292) Aussi on l'aimait et, les jours de réjouissance, nombreux étaient les compères qui se pressaient autour de lui, heureux d'entendre ses calembredaines, ses histoires, ses drôleries, les provoquant au besoin, et le ramenant, sans en avoir l'air, aux sujets de conversations qu'il préférait. Ou bien ils commentaient ses récits d'un petit clignement d'yeux à l'adresse de la société, comme pour en faire valoir la saveur, toute la verve rare et puissante. Sacré Poloche, on ne savait pas où il allait chercher **tout ça** ! (Moselly, E., *Terres lorraines*, 1907 < Frantext)

Dans cet exemple, *tout ça* fonctionne non seulement de manière résomptive mais suggère même plus que les référents préalablement introduits. En outre, il participe, à la faveur de son caractère token-réflexif, à l'expression d'un discours indirect libre rendant compte de l'expérience d'auditeur du narrateur.

### 3.3 *tout ça détaché*

Le segment *tout ça* peut apparaître comme élément détaché, redoublé par un pronom, ici le clitique *ça* en position sujet :

(293) parce qu'on avait en général huit vaches | \_ | le mulet | \_ | deux deux chèvres | \_ | et souvent | | un ou deux moutons | \_ | et **tout ça** ça mange | \_ | il faut emmagasiner du foin pour tout l'hiver

Dans ce cas, *tout ça* englobe notamment les objets préalablement introduits, mais envisage ouvertement le détail de cet ensemble avec une certaine latitude (*en général, souvent, un ou deux*).

A noter qu'on trouve sans problème des détachements dans l'écrit littéraire :

(294) [Yersin, un médecin suisse, est en fonction en Indochine] Yersin hésite, parce qu'il craint de devoir se mettre en règle un de ces jours avec les autorités françaises. Il a sollicité sa naturalisation et depuis n'a jamais effectué son temps sous les drapeaux. Mais **tout ça** pour lui c'est fini. C'est son ancienne vie. (Deville, P., *Peste et Choléra*, 2012)

Le syntagme *tout ça* récapitule les préoccupations du personnage à l'égard de sa situation irrégulière, mais semble suggérer en même temps l'existence d'autres tracas implicites, tout en participant, là aussi, à l'expression d'un discours indirect libre.

### 3.4 *tout ça* appositif régissant un circonstant

Le syntagme *tout ça* se retrouve également dans une position qu'on pourrait analyser comme appositive, y régissant un complément subséquent :

- (295) ben euh ben même si t'es juste financièrement tu peux t'inscrire et pis on t'envoie | et pis sinon t'as la possibilité d'imprimer chez toi les documents et puis euh | \_ | et pis au moins la secrétaire elle est pas obligée de passer deux semaines avant euh | \_ | plein d'heures à la photocopieuse euh | \_ | enfin voilà | \_ | **tout ça** pour dire | \_ | que euh je suis très contente de cette commission parce que on parle vraiment de | \_ | moi j'aime bien tout ce qui est euh | \_ | systèmes électoraux (ofrom)

Dans l'exemple ci-dessus, *tout ça* fonctionne en apposition d'une succession d'énoncés dont la délimitation syntaxique se montre par ailleurs indéfinie. La séquence *tout ça pour dire* illustre bien l'exploitation argumentative de *tout ça* en contexte récapitulatif (cf. *enfin voilà*) qui consiste à mettre fin au développement précédent dans le but d'introduire la conclusion jugée pertinente. On retrouve le même schéma ci-dessous, produisant des rendements discursifs analogues :

- (296) donc on travaille beaucoup en interdisciplinarité | avec des ergothérapeutes des physiothérapeutes des assistantes sociales | \_ | des diététiciennes | \_ | donc on a plusieurs corps de métiers | \_ | il faut déjà qu'on trouve une harmonie entre nous entre des cultures bien différentes | \_ | et **tout ça** pour converger vers un vers une personne dans le but de la maintenir à domicile (ofrom)

Dans l'exemple précédent, *tout ça* permet de récapituler le développement pour introduire la finalité du propos. Le complément introduit par le biais de *tout ça* n'exprime cependant pas toujours la finalité. Il illustre toutes sortes de fonctions circonstancielle, par exemple ci-dessous la caractérisation d'une manière de parler :

- (297) Alors ma mère se levait, s'approchait du fauteuil, désignait les arrivées de gaz, de chaque côté de la glace, et entamait un petit cours d'hygiène d'où il ressortait que j'étais une enfant, que bien souvent la clientèle adulte est porteuse de germes et que, de toute façon, les becs de gaz n'étaient pas là pour les chiens. **Tout ça** avec une politesse exquise. Puis elle allait se rasseoir. (Signoret, S., *La nostalgie n'est plus ce qu'elle était*, 1976< Frantext)

Dans cet exemple tiré de l'écrit littéraire, *tout ça* enchaîne comme auparavant sur un développement, ici à la suite d'un discours narrativisé.

Enfin, l'exemple suivant nous permet de faire la transition avec la dernière partie de la section, car il illustre un cas de *tout ça* enchaînant sur une liste, tout en servant de relais à l'expression d'une cause :

(298) alors, tu aurais étudié six ans, tu aurais été au collège, tu aurais été au gymnase, tu m'aurais coûté tant d'argent, **tout ça**... (il répéta le mot), **tout ça** pour des enfantillages ! (Ramuz, F., *Aimé Pache, peintre vaudois*, 1911 < Frantext)

La première occurrence de *tout ça* peut dans un premier temps être interprétée comme une particule d'extension de la liste (*cf.* ci-après). Mais le prolongement de la structure dément cette analyse et en fait un élément recteur à l'instar des exemples précédents, servant à introduire une cause, ici sur le ton du reproche.

### 3.5 *tout ça dans les listes*

On relève enfin un dernier contexte d'occurrence pour *tout ça*, la configuration de liste, principalement en dernière position de l'énumération<sup>222</sup> :

(299) Et ça se fait comment que l'internet est arrivé si tard en France ? Parce que en général la France est un pays qui adore le progrès, la technique, les avions **tout ça**. Alors, euh, pourquoi c'était pas comme ça pour l'internet ? (pfc)

Dans cet extrait, *tout ça* intervient à la suite d'éléments lexicaux occupant la position de complément d'objet direct du verbe *adore*. On est en droit d'invoquer, comme dans ses autres emplois, un fonctionnement résomptif : il comprend, dans sa référence, les objets mentionnés dans le cadre de l'énumération qu'il clôt. Mais comme dans les cas précédents, il fait bien plus que ça : tout en se présentant comme constituant postiche de la liste (*cf. supra* §5.1.3), il signale justement qu'il y a davantage à considérer que les ingrédients explicitement listés : il pousse l'interlocuteur, en simulant une certaine connivence, à envisager une classe-objet hétéroclite dont il révèle l'ampleur.

### 3.6 *Bilan sur les emplois de tout ça*

Dans tous ses emplois, le syntagme *tout ça* a pour vocation de renvoyer à une classe-objet floue, *i.e.* dont les membres hétérogènes sont inférables à des degrés de certitude divers, à laquelle appartiennent pour sûr certains membres explicitement introduits en amont. Il semble donc cumuler une fonction résomptive et « extensible ».

Néanmoins, nous avons décidé de restreindre la suite de notre investigation au dernier contexte d'occurrence de *tout ça* – la configuration de liste – pour plusieurs raisons : contrairement aux premiers contextes de *tout ça* que nous n'avons pas eu de peine à rencontrer à l'écrit « travaillé », la position de fin de liste y figure plus rarement. Nous l'y

---

<sup>222</sup> Cette position semble privilégiée, mais non exclusive. En effet, *tout ça* est également attesté en tête de liste (il est dans ce cas clairement régi) ou alors dans une position intermédiaire (voir aussi Bilger 1989 : 100) : *j'commençais à en avoir marre de faire le feu de de bois tout ça charbon les cendres et tout* (cfpp)

avons relevé ponctuellement, mais presque exclusivement dans des discours directs ou des écrits à la première personne :

(300) Je dis là des choses très négatives sur moi, mais je n'ai pas du tout l'esprit Mai 68. Je ne partage pas du tout cette utopie révolutionnaire, la poésie dans la rue, **tout ça**. J'ai horreur de la « fête », ça ne m'intéresse pas. Je suis un homme d'ordre ! (Grenier, C., *La vie possible de Christian Boltanski*, 2007 < Frantext).

Si nous nous tournons vers sa variante « prestigieuse » *tout cela* (cf. *supra* §5.2.3), on constate une semblable retenue en position de liste à l'écrit<sup>223</sup>, alors que les trois autres contextes y sont bien représentés<sup>224</sup>.

Tout porte ainsi à croire que l'occurrence de *tout ça* dans les listes est assez spécifique de l'oral spontané (Bilger 1989, Ferré 2011, Secova 2014), ou du moins, des genres de l'immédiat, où il commute, non pas tant avec des SN comme dans les autres environnements syntaxiques, mais plutôt avec des particules d'extension de liste, comme *et cetera*. Cette différence de paradigme nous pousse dès lors à explorer en détail son fonctionnement référentiel dans cette configuration et le développement des rendements pragmatiques et interactionnels qui l'accompagne.

---

<sup>223</sup> Quant à *tout cela* à l'oral, il est, sans surprise, plutôt rare : dans la base OFROM, il apparaît quatre fois (contre 282 *tout ça*), jamais dans CFPP (contre 333 *tout ça*) et six fois dans PFC (contre 1381 *tout ça*) (en date du 19.07.2016). Parmi ces six occurrences, quatre proviennent d'une même locutrice belge et les deux autres de deux locuteurs d'Afrique (Algérie et République centrafricaine). Et dans les trois corpus oraux confondus, seul un *tout cela* pourrait faire l'objet d'une analyse en contexte de liste, encore que son statut syntaxique demeure ambigu : *Les Pygmées, ils mangent seulement ce qu'ils, qu'ils trouvent dans la forêt, les gibiers, et puis même les inyams, tout cela, ce que nous mangeons, eux aussi, ils mangent aussi*. (pfc, RCA Bangui). On peut y voir le dernier constituant d'une liste ou alors un complément d'objet détaché.

<sup>224</sup> En voici trois exemples : 1) *La salle était sombre et parcourue en tous sens par des gens se dépêchant, malfaiteurs, policiers, hommes de loi, prêtres et journalistes je suppose. Tout cela faisait sombre, de sombres formes se pressant dans un espace sombre*. (Beckett, S., *Molloy*, 1951 < Frantext) ; 2) *C'est Mme Badiou qui m'a fait aimer Montaigne. Mes copines trouvaient cela austère. Montaigne, c'est la tolérance, l'ouverture aux autres, le goût des autres, c'est la raison ! Il essaie de se mettre à la place des autres pour penser, pour voir comment ils pensent. Tout cela c'est Montaigne*. (Dupuy, A., *Journal* [...], 2013 < Frantext) ; 3) *J'eus aussi le pénible devoir de remettre au Général deux photographies de sa mère sur son lit de mort. (Il n'eut pas le bonheur de la retrouver vivante à son retour en France.) Il les prit avec une sorte de brutalité où je vis de la pudeur et les mit de côté. Tout cela sans intérêt et sans que j'y prête une grande attention*. (Mauriac, C., *Aimer de Gaulle*, 1978 < Frantext).

#### 4. Etude de *tout ça* dans les listes à l'oral

Comme on vient de l'évoquer, le marqueur *tout ça* intègre un paradigme de marqueurs d'extension de liste, parmi lesquels on peut mentionner *et cetera*, *et ainsi de suite*, *et autres N*, *et compagnie*, *et consorts*, *et tout (et tout)*, *et tout le toutim / le tremblement / le reste / le tralala / le bastringue*, etc.), *(et) patati et patata*, *(et) machin*, *ou comme ça*, *SN comme ça*, *ou n'importe quoi*, etc. Des études sur d'autres langues attestent l'existence d'une catégorie similaire<sup>225</sup> qui a reçu plusieurs dénominations : *generalized list completers* (Jefferson 1990, Selting 2007), *set-marking tags* (Ward & Birner 1993), *general extenders* (Overstreet 1999, Secova 2014). En français, elle reçoit les noms de *particules d'extension* (Dubois 1993, Ferré 2011) ou encore de *clôtureurs* (Kahane & Pietrandrea 2012).

Malgré la diversité des formes possibles, ce sont les marqueurs *et tout*, *tout ça*, et *et cetera*<sup>226</sup> qui dans l'ensemble, sont les mieux attestés à l'oral dans les relevés des travaux antérieurs<sup>227</sup> (Ferré 2011, Secova 2014) et dans les bases que nous avons examinées. Figurent également en bonne place, quoique dans des proportions plus faibles, les variantes en *comme ça* du type *(ou) truc(s)/chose(s) comme ça* et le cas de *(et/ou) machin*<sup>228</sup> (Secova 2014). Quant aux autres marqueurs, ils semblent intervenir de manière beaucoup plus ponctuelle, voire pas du tout pour certains, dans les bases de données examinées<sup>229</sup>. Si notre objet d'étude principal est *tout ça*, il nous arrivera de le mettre en perspective avec *et tout*, au vu de leur parenté morphologique (quantificateur de totalité) et d'un fonctionnement pragmatique très proche :

(301) alors on s'arrêtait quand on voulait euh | \_ | quand ça nous plaisait ***et tout*** (ofrom)

La séquence *tout ça* est elle aussi régulièrement, mais non systématiquement, précédée du coordonnant *et* :

(302) bon ben ça pose des problèmes de mainten- enfin de maintenance euh de de mise à jour ***et tout ça*** euh voilà (pfc)

On peut signaler que l'on trouve, de façon éparse, d'autres coordonnants devant *tout ça* :

(303) Non, même au Canada, on peut aller, si on va à Vancouver ou en Californie ***ou tout ça***<sup>230</sup> ouais. (pfc)

---

<sup>225</sup> Pour l'anglais: *and all (that stuff)*, *and stuff*, *or something*, *and everything*, *or whatever*, etc.; pour l'allemand : *und so (weiter)*, *oder so (was)*, etc. (Overstreet 2005)

<sup>226</sup> La particule connaît de nombreuses variantes graphiques : *et cetera*, *etcetera*, *et caetera*, *et caetera*, *et cétéra*, *etcétéra*, etc.

<sup>227</sup> Voir Dubois (1992) et Secova (2014) pour une analyse distributionnelle, la première à partir de deux corpus de français québécois, la seconde à partir de trois corpus de français de l'Hexagone.

<sup>228</sup> Sur *machin* et *ou comme ça*, voir Béguelin & Corminboeuf (à par.).

<sup>229</sup> A noter qu'il ne faut pas négliger les facteurs diatopiques. A partir d'un corpus de français québécois, Dubois (1992) relève des particules semblant spécifiques à cette aire géographique, du genre les variantes en *affaire (des affaires de même, toutes ces affaires-là, une affaire comme ça, etc.)* ou en *patente (toutes ces patentes-là, des patentes comme ça, des patentes de même, etc.)*.

<sup>230</sup> Rare devant *tout ça* (cf. aucune attestation dans OFROM), au contraire de la présence fréquente de *ou* devant *comme ça* (Corminboeuf 2016, Béguelin & Corminboeuf à par.).

(304) Moi, ça serait plutôt le tricot, la couture, la broderie, tout ce qui est féminin dans le fond hein. Mais pas tellement des loisirs autrement, de, ni dans le sport, **ni tout ça**.  
(pfc)

Enfin, mentionnons la possibilité d'occurrence d'un *tout* « pronominal » non précédé d'un coordonnant :

(305) c'est, c'est, c'est par rapport à la mutuelle euh bon **tout** et comme ici en plus il a il a changé de de domicile (pfc)

L'analyse comme marqueur d'extension du morphème *tout* isolé reste néanmoins beaucoup plus difficile à identifier que les autres.

Cette section est envisagée de la manière suivante : nous proposons un aperçu des travaux existants sur les marqueurs d'extension et en particulier sur *tout ça* (§4.1). Après quelques précisions sur les données (§4.2), nous nous penchons sur la configuration de liste – étudiée sous les angles syntaxique, prosodique, sémantico-pragmatique et interactionnel – configuration qui offre un moule particulièrement propice à l'émergence des marqueurs d'extension (§4.3). Nous nous intéressons ensuite en détail au fonctionnement respectivement référentiel (§4.4) puis pragmatique de *tout ça* (§4.5).

#### **4.1 Travaux antérieurs**

##### **4.1.1 Les particules d'extension**

Les particules – ou marqueurs – d'extension ont pour caractéristique de figurer en clôture de syntagme ou de clause, marquant à cet égard la référence à l'ensemble auquel apparten(n)ent entre autres le ou les membre(s) ainsi introduit(s) en amont (Dubois 1992, Overstreet 2005). Ils sont souvent composés d'un coordonnant (facultatif), d'un quantifieur ou d'une marque de comparaison (facultatifs) et d'un N général. Leur usage est en outre reconnu comme un procédé fondamentalement interactionnel véhiculant une fonction interpersonnelle créatrice d'effets d'intersubjectivité et mettant en œuvre des stratégies de politesse ou d'atténuation (Overstreet 1999). Les particules d'extension sont vues comme des manifestations typiques de la langue parlée spontanée, bien qu'elles soient également attestées, dans des proportions bien plus faibles, à l'écrit formel (Secova 2014). A cet égard, elles sont perçues, sans surprise, comme des défauts de langage dans l'opinion collective, jugées imprécises et dépourvues de sens (Dines 1980).

Etant donné les fonctions typiquement pragmatiques qu'elles sont susceptibles de développer et leur sens plus procédural que référentiel, on les ramène généralement à la catégorie des « marqueurs discursifs » (Dubois 1993, Overstreet 1999, Secova 2014), le seul critère qui les en distingue étant leur position contrainte (position terminale d'un syntagme).

#### 4.1.2 *tout ça* particule d'extension

En ce qui concerne plus particulièrement le marqueur (*et*) *tout ça* en français, Bilger (1989) s'y intéresse dans le cadre de l'approche pronominale du GARS. Remarquant qu'il est susceptible de clore un paradigme de répétitions lexicales, elle souligne toutefois le caractère implicite de la liste entamée, dont la (ou les) réalisation(s) lexicale(s) donne(nt) un indice sur le champ sémantique à inférer. En comparant la présence avec l'absence du coordonnant *et*, elle distingue un effet de sens respectivement additif *vs* appositif et englobant pour le second (*ibid.* : 100-101).

Secova (2014) compare les types de marqueurs d'extension en français dans trois corpus distincts, afin d'y étudier les facteurs diachronique et générationnel sur la distribution des marqueurs. Dans son propre corpus de jeunes adultes parisiens enregistrés dans les années 2007-2009, c'est clairement la variante *et tout* qui est la plus fréquente (76%), suivie à égalité de *tout ça* (4%), *machin* (4%) et *truc(s) comme ça* (4%), le reste des occurrences se répartissant entre huit autres formes. Dans le corpus CFPP des années 2000, dont les locuteurs sont issus de toutes les catégories d'âge, *et cetera* est le mieux représenté (28%), suivi de *et tout* (22%), *tout ça* (16%) et *et tout ça* (9%), parmi 28 formes attestées. Enfin, le dernier corpus, composé d'enregistrements des années 1980 de locuteurs français d'âge variable, indique *tout ça* comme la variante préférée (33%), suivie d'*et cetera* (18%), d'*et tout* (10%) et *choses comme ça* (10%), parmi 16 formes utilisées.

L'analyse met ainsi en lumière une préférence indéniable d'*et tout* dans le corpus le plus récent et parmi les jeunes adultes (confirmée également par le sous-groupe de jeunes dans CFPP) au détriment de la forme *tout ça*, qui était prédominante dans le corpus le plus ancien. Secova s'intéresse donc de plus près à *et tout* et dégage les diverses fonctions pragmatiques qu'il assume (rôle de ponctuant, d'intensificateur ou à l'inverse de mitigeur, mise en œuvre de stratégies argumentatives, etc.). Au vu de ces indices, l'auteur soutient l'hypothèse d'un processus de grammaticalisation le concernant, se traduisant par un blanchiment sémantique et la diminution de sa valeur référentielle, au profit d'un fonctionnement pragmatique, dont le stade le plus abouti est celui de ponctuant du discours. Au contraire, elle considère la variante *tout ça* comme conservant un fonctionnement davantage référentiel, qui maintient l'indication de l'extension d'un ensemble.

#### 4.1.3 Bilan

Ce bref exposé des travaux antérieurs appelle plusieurs commentaires et nous conduit à préciser nos positions.

D'une part, des réserves peuvent être émises à l'égard de la notion de *marqueur discursif* (aussi appelé *marqueur pragmatique*), dont voici une définition couramment citée :

A pragmatic marker is defined as a phonologically short item that is not syntactically connected to the rest of the clause (*i.e.*, is parenthetical), and has

little or no referential meaning but serves pragmatic or procedural purposes (Brinton 2008 : 1)

Pour le français, on retient dans cette catégorie, selon leur contexte d'occurrence, des items aussi variés que *ben, enfin, donc, quoi, tu vois, tu sais, voilà, en fait, etc.*

Le problème que pose ce genre de définition est que les caractéristiques qui sont attribuées aux marqueurs<sup>231</sup> ne sont pas spécifiques aux éléments invoqués : elles s'appliquent en effet à toutes sortes d'autres items de la langue. Par exemple, le critère de réduction phonétique se retrouve dans *je sais > j'sais > chsais > chais*, sans pour autant qu'on traite celui-ci comme un marqueur discursif, contrairement à certains emplois de *tu sais > t'sais*. Il en va de même concernant le critère de mobilité syntaxique, qui caractérise toutes sortes de constituants divers (circonstants, adverbes d'énonciation, appositions, verbes recteurs faibles, etc.). Ainsi, malgré un vague faisceau de propriétés communes, celles-ci ne peuvent être rigoureusement tenues pour définitives. On a par ailleurs souvent reproché le caractère « fourre-tout » de la catégorie (Kerbrat-Orecchioni 2005 : 49), mettant dès lors sérieusement en cause sa pertinence. Aussi préférons-nous nous passer de la notion, sans toutefois nier le fonctionnement pragmatique des items qui nous intéressent. Précisons que nous approuvons à l'inverse la reconnaissance d'une classe de marqueurs d'extension, pour autant qu'on y fasse entrer des éléments partageant la même distribution.

D'autre part, des réserves peuvent être formulées à propos du cadre théorique de la grammaticalisation adopté par Secova (2014), qui comporte en effet plusieurs postulats discutables, parfois contradictoires (Béguelin, Corminboeuf & Johnsen 2014). Parmi ceux-ci, citons par exemple les principes de *gradualité* et d'*unidirectionnalité*, qui situent les formes concernées sur des parcours évolutifs graduels à sens unique. Bien que ces théories admettent la possibilité de coexistence simultanée de deux usages d'une même forme, qui représenterait un indice de changement en cours, il nous paraît plus prudent d'examiner les facteurs contextuels motivant les emplois respectifs, plutôt que de mettre cela au compte d'une hypothèse déterministe. Par ailleurs, ce point de vue sur le changement linguistique occulte complètement la dimension idiosyncratique des emplois, alors qu'il va de soi que le recours à certaines expressions peut être récurrent chez certains locuteurs, comme on va le voir, mais totalement inexistant chez d'autres. Bien que nous soyons généralement en accord avec l'identification des fonctions pragmatiques et les analyses interprétatives de Secova, nous tenons donc à prendre nos distances avec l'hypothèse sous-jacente de grammaticalisation d'*et tout*. Nous allons d'ailleurs voir que les différents rendements observés sont également imputables à *tout ça*.

#### **4.2 Remarques sur les données**

La méthodologie de récolte des données ayant déjà été présentée dans l'avant-propos de cette partie, nous nous limitons à quelques précisions spécifiques au marqueur étudié. Nous avons

---

<sup>231</sup> Voir à cet égard Fernandez (1994), Brinton (1994), Dostie (2004), Dostie & Push (2007), Denturck (2008).

donc recherché dans les bases indiquées les suites *tout ça* (ainsi que les suites *et tout*), parmi lesquelles nous retenons, après vérification manuelle, celles qui s'inscrivent dans la configuration souhaitée de liste. Avec toute la prudence que ce genre de relevé suppose, signalons à titre indicatif que nous avons recueilli, sur des versions antérieures et plus restreintes des bases de données<sup>232</sup> : 35 *tout ça* en configuration de liste sur les 67 résultats<sup>233</sup> de la requête de la séquence sur OFROM (52%), 166 sur 297 dans CFPP (56%)<sup>234</sup>.

### 4.3 Propriétés des listes

Nous envisageons dans ce travail les listes, ou énumérations – notions que nous considérons synonymes – comme des structures visant à l'élaboration référentielle d'une classe-objet et qui se caractérisent par des parallélismes à différents niveaux, à savoir les niveaux lexico-syntaxique, sémantico-pragmatique et prosodique. Toutefois, l'activation de tous ces parallélismes n'est pas systématiquement requise pour qu'on interprète une liste de la sorte : mais il semble que plus nombreux et visibles sont les parallélismes, mieux reconnue en tant que telle est la liste (Johnsen & Avanzi, à par.). On verra d'ailleurs qu'une liste, si tant est qu'on puisse la considérer comme telle, peut n'être ébauchée que par un seul item, suivi d'une particule d'extension. Avant d'observer le fonctionnement de *tout ça* dans ces configurations, il nous paraît important de caractériser plus précisément ces structures aux différents niveaux signalés pour comprendre pourquoi elles sont favorables à l'accueil du marqueur.

#### 4.3.1 Aspects lexico-syntaxiques

Au niveau micro-syntaxique, on peut considérer la liste comme une série additive d'items lexicaux occupant une même position syntaxique sur un axe paradigmatique :

(306) Et ça se fait comment que l'internet est arrivé si tard en France? Parce que en général la France est un pays qui adore le progrès, la technique, les avions **tout ça**. Alors, euh, pourquoi c'était pas comme ça pour l'internet ? (pfc)

La liste comprend ainsi plusieurs éléments instanciant la place d'argument objet du verbe *adore*. L'approche du GARS représente ce genre de phénomène sous la forme de « grilles » disposées sur l'axe paradigmatique vertical (Blanche-Benveniste 1990a, Blanche-Benveniste *et al.* 1990), que nous illustrons avec l'exemple précédent :

<sup>232</sup> En octobre 2012, env. 15h d'enregistrement sur OFROM et 39h sur CFPP. Pour rappel, nous avons renoncé à des comptages sur PFC en raison de problèmes de doublons dans les résultats. Mais nous ne nous privons pas pour autant du recours aux données de la base en question.

<sup>233</sup> Les suites non Pro-SN incluses, comme *pas du tout ça*.

<sup>234</sup> Quant à *et tout*, nous en avons relevé 71 occurrences clôturant une liste sur 88 dans OFROM (80%) et 154 sur 182 occurrences dans CFPP (85%). A titre indicatif et tous emplois confondus, on relève au 16.02.2017 : 307 (*et tout ça*) 392 *et tout* (sans *et tout ça*) dans OFROM ; 333 (*et tout ça*) et 194 *et tout* dans CFPP. On constate que le corpus OFROM, le plus récent (enregistrements à partir de 2008), ne confirme pas le déséquilibre relevé entre *et tout* et *tout ça* dans le corpus personnel de Secova (2014). Il est vrai cependant que l'âge des locuteurs d'OFROM est beaucoup moins homogène.

(307) la France c'est un pays qui adore le progrès  
la technique  
les avions  
*tout ça* (pfc)

Nous prenons cependant nos distances avec la conception aixoise de la liste qui intègre les phénomènes de « bribes du discours » (Blanche-Benveniste 1990a : 14) comme les « bredouillages, hésitations, maladroites, reprises » (Blanche-Benveniste *et al.* 1990 : 20) :

(308) on mettait oh même pas la moitié d'une cuillère à soupe  
d'une cuillère à café (*ibid.* : 21)

S'il est vrai qu'il n'est pas toujours évident de distinguer entre série additive et « disfluences », d'où l'émergence d'ambiguïtés d'analyse (Blanche-Benveniste 2011, Kahane & Piertrandrea 2012), nous considérons que les secondes ne relèvent pas du même phénomène car elles ne sont pas vouées, au niveau référentiel, à la création délibérée d'une classe-objet et elles ne se distinguent pas par des parallélismes récurrents aux autres niveaux (sémantico-pragmatiques, prosodiques). D'ailleurs, sur le plan syntaxique, un élément distinctif semble négligé dans l'approche aixoise : l'apparition régulière de coordonnants dans les séries additives (en ce qui nous concerne, la présence possible de *et* devant *tout ça*) vs leur absence notoire dans le cas des disfluences. Précisons que nous admettons toutefois dans la catégorie des listes le cas de formulations multiples (Blanche-Benveniste 2011) visant par exemple à l'adéquation de la dénomination, ou à la recherche de nuances, parce qu'elles contribuent, contrairement aux disfluences à proprement parler<sup>235</sup>, à créer une classe (de dénominations, de propriétés, etc.). En témoigne ci-dessous la caractérisation nuancée d'un même objectif :

(309) par exemple j'ai pas tout à fait fini ce que je dois faire ben voilà | \_ | j'arrête pour euh |  
\_ | pour être en forme le lendemain pour pouvoir euh **bien suivre les cours bien écouter tout ça**<sup>236</sup> (ofrom)

Les énumérations se rencontrent également au niveau macro-syntaxique, où les items mis en parallèle représentent cette fois des « énonciations » :

(310) d'accord ces jeunes | \_ | jouent le jeu ils sont | \_ | **ils sont enthousiastes ils ils ont de l'énergie ils ont des idées tout ça** ce qui est bien | \_ | mais franchement pas à pas à Berne | \_ | pas à cet âge-là et pas à Berne (ofrom) = (3)

Il nous semble légitime d'intégrer ces successions d'énonciations d'ordre macro-syntaxique aux phénomènes de liste, d'une part au vu de la répétition des patrons lexico-syntaxiques et

<sup>235</sup> Kahane & Piertrandrea (2012) les traitent tous deux dans la même catégorie, celle des entassements qu'ils appellent *de dicto*.

<sup>236</sup> Une écoute de l'extrait sonore met en évidence une prééminence prosodique sur la syllabe [te] de *écouter*, qui oriente vers une interprétation de *tout ça* comme particule plutôt que comme régime direct du verbe, cf. *infra* §4.3.3.



(315) on se retrouve comme ça des week-ends pour faire un peu de la musique **on compose on joue on a des concerts en Valais tout ça** (ofrom)

La notion d'isotopie demande néanmoins à être assouplie, la « catégorie créée » (Schiffrin 1994) par la liste ne manifestant pas toujours un sème générique préinscrit dans le contenu lexical ou dans des schémas d'actions types, mais étant susceptible d'émerger d'un point de vue « subjectif » sur le monde (Schiffrin 1994) ou de facteurs contextuels (Dubois 1992, Johnsen 2011). L'exemple suivant illustre ainsi une séquence d'arguments émanant d'une instance tierce et subjective :

(316) puis lui c'est là qu'il s'est lancé en disant que c'était terrible euh au Pakistan parce que le | \_ | le gouvernement laïc à l'époque euh | \_ | **c'était inacceptable que c'était beaucoup mieux en Iran tout ça tout ça** | \_ | donc un type euh un fanatique de l'islam (ofrom)

La liste intervient dans le cadre d'un discours rapporté et entame une série de critiques à l'égard du gouvernement laïc au Pakistan, autrement dit elle reflète le point de vue subjectif de l'énonciateur cité. Il est par conséquent difficile de reconnaître un schéma d'actions prédéfini qui engloberait typiquement les items formulés, étant donné qu'ils évoquent ici des jugements de valeur ('l'aberration d'un gouvernement laïc au Pakistan', 'les conditions plus favorables de l'Iran').

Au niveau pragmatique, on peut s'interroger sur la relation que le locuteur établit entre les différents items : énumère-t-il des référents distincts ou procède-t-il à des désignations multiples d'un même référent (Blanche-Benveniste 2011) ? Dans la liste suivante, il est évident que chaque membre de la liste renvoie à un individu distinct :

(317) t'étais dans l'couloir en train d'parler avec **Madame Chauvel Madame Beaumont Madame Cardémont tout ça** (cfpp)

Dans d'autres cas, le locuteur semble plutôt viser un même référent pour lequel il cherche la formulation appropriée :

(318) par exemple j'ai pas tout à fait fini ce que je dois faire ben voilà | \_ | j'arrête pour euh | \_ | pour être en forme le lendemain pour pouvoir euh **bien suivre les cours bien écouter tout ça** (ofrom) = (309)

Comme le souligne Blanche-Benveniste (2011), on ne peut pas toujours savoir si le locuteur tente de reformuler un même concept ou s'il cherche plutôt à dissocier différents objets par des nuances fines.

Certains auteurs distinguent le caractère fermé ou ouvert d'une liste (Paveau & Rosier 2009, Selting 2007). En effet, une liste peut se présenter comme fermée lorsque des éléments du contexte linguistique (annonce, récapitulation) indiquent le nombre d'items concernés ou lorsque les items sont eux-mêmes dénombrés dans la procédure de liste. Un coordonnant

précédant directement le dernier élément peut également constituer l'indice que l'auteur présente sa liste comme exhaustive :

(319) on était nomades | \_ | on avait | \_ | **une maison à** | # | \_ | **une maison à** | # | \_ | **et une maison à** | # | \_ | et partout on avait | \_ | des champs (ofrom)

On ne peut néanmoins tirer la même conclusion de la présence du coordonnant dans *et tout* (*ça*). Si celui-ci peut servir à indiquer, au niveau segmental, la fin de la structure de la liste, le syntagme dans son ensemble indique au niveau référentiel la non-exhaustivité des ingrédients de l'ensemble évoqué (Bilger 1988, Overstreet 1999, Ferré 2011, Johnsen 2011).

Enfin, au niveau interactionnel, la construction de listes a été considérée comme une activité volontiers collaborative (Jefferson 1990, Erickson 1992, Selting 2007) : un interlocuteur peut facilement prendre part à l'activité en ajoutant un ou des membres à la liste, démarche aboutissant à une « co-production of a list in progress » (Jefferson 1990 : 89). Nos données attestent ce genre de collaboration :

(320) L1: d't'façon c'est toujours pareil dans une école hein va y avoir un peu euh: **les racailles**  
L2: <ouais  
L1: **les**> **gothiques** euh  
L2: les les  
L1: **les shals**  
L2: voilà [rires] + **les bourges** voilà <**tout ça** ouais  
L1: voilà > (cfpp)

Nous aurons l'occasion de revenir sur cet exemple *infra* (§4.5.1), car le marqueur *tout ça* permet ici des manœuvres intersubjectives particulièrement intéressantes.

### 4.3.3 Aspects prosodiques

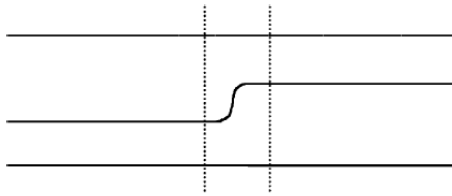
La prosodie des listes à l'oral spontané n'a pas fait l'objet d'études systématiques. Le travail le plus approfondi à notre connaissance est celui de Selting (2007). A partir de l'observation d'un corpus de conversations en allemand, l'auteur met en évidence les contours intonatifs des syllabes accentuées des membres de listes. Elle distingue plusieurs patrons prosodiques selon que la liste est fermée ou ouverte. Dans le cas des listes fermées, elle note que les éléments composent généralement une seule unité prosodique dotée d'un contour descendant se terminant dans le grave, sur un schéma de *downstep*<sup>238</sup> c'est-à-dire un abaissement tonal progressif sur chaque item. A l'inverse, dans le cas des listes ouvertes, les membres sont prononcés dans des unités prosodiques distinctes, avec un parallélisme remarquable entre les deux premiers membres de la liste : répétition des contours intonatifs, intensité et durée

---

<sup>238</sup> Selting (2007) donne la définition suivante : « 'Downstep' here, as in most of recent intonology, denotes that successive pitch peaks in an intonation unit are produced on successively lower pitch height, *i.e.* on a descending line ».

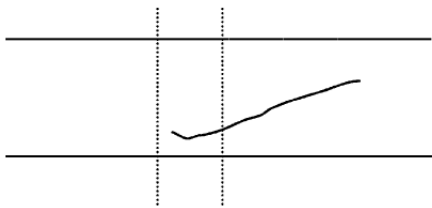
comparable, ceci sans *downstep*. Selting relève sept types de contours pour les syllabes accentuées finales des listes ouvertes. La plupart se terminent par une fréquence haute ou moyenne, souvent atteinte après un palier montant (*upward staircase*) se terminant par un plateau haut (ou légèrement descendant) (*ibid.* : 506-507) :

*welcher BAND, welches SCHEInungsjahr und so weiter* (p. 509)



ou alors via un contour montant continu :

*für die rEIfen, und so wEIter, einige reparaTURN* (p. 517):



Malgré la diversité des types répertoriés, c'est surtout la répétition du contour qui rend la liste remarquable. En ce qui concerne le dernier membre d'une liste ouverte, Selting remarque qu'il peut être prononcé soit sur le même schéma que les membres précédents, s'auto-désignant précisément comme élément non final, soit avec une prosodie distincte, souvent un contour descendant, signalant la complétude, non pas du contenu de la liste, mais de l'activité même de liste.

Pour le français parlé, une étude d'Auchlin & Simon (2004) est consacrée à quelques extraits de listes. L'hypothèse des auteurs est que la reconnaissance d'un schéma de liste se fonde essentiellement sur deux composantes : le rythme, caractérisé ainsi : « régularité du retour des temps accentués, espacement régulier des attaques, accents initiaux par élément mis en liste » ; et la modulation : « dynamisme accru dans la réalisation des tons, élargissement du registre tonal, allongement (ou réduction) de la durée des syllabes accentuées, etc. » (*ibid.* : 188). Les auteurs remarquent à juste titre que les listes sont très perceptibles pour un participant d'une interaction, mais paradoxalement beaucoup moins lorsqu'on ne se fie qu'à sa transcription prosodique. Les auteurs accordent une attention particulière à la prosodie du premier élément d'une liste. Celui-ci peut être prosodiquement marqué comme tel, doté ainsi d'une valeur de projection (annonçant l'initiation d'une énumération), ou alors n'acquiescer le

statut de premier élément d'une liste qu'*a posteriori*, de par la similitude syntaxique, prosodique et/ou sémantique qu'il entretient avec les items ultérieurs. Lorsqu'il se démarque, il peut le faire sur son attaque (pause, prise de souffle) ou sur sa syllabe finale (proéminence, exagération du contour) (*ibid.* : 190). En principe, la syntaxe et la prosodie fonctionnent de manière complémentaire ou redondante. Auchlin & Simon (2004) relèvent cependant des cas de « conflit entre les instructions données par la syntaxe et l'empaquetage prosodique » (*ibid.* : 190) et illustrent des cas où la prosodie prend le dessus sur la syntaxe pour imposer un schéma de liste sur des structures qui, sur le plan strictement syntaxique, n'en auraient pas le statut (voir aussi Selting 2007 : 519). Ils relèvent toutefois que toutes les listes ne sont pas marquées sur le plan intonatif : certaines présentent simplement des groupes intonatifs juxtaposés (contour continuatif bas rehaussé ou haut en position d'accent final) dominés par le dernier d'entre eux (ton plus haut ou alors infra-bas) marquant la clôture. Cela n'empêche pas le respect de contraintes d'eurythmie : alignement des attaques, ajustement syllabique ou de tempo visant à la régularité rythmique, scansion par des prises de souffles, etc. Les auteurs mettent en relation la prosodie plus ou moins marquée des listes avec l'engagement de ceux qui les produisent : l'énergie manifestée, tout comme les efforts de reproduction du pattern sont interprétés comme des signes iconiques de l'investissement des locuteurs dans leur discours.

Une étude que nous avons menée (Johnsen & Avanzi *à par.*) sur un corpus de français parlé de Suisse romande montre que les contours des syllabes terminales des éléments initiaux ou internes d'une liste manifestent un contour de type continuatif /H\*H%/<sup>239</sup> avec trois variantes possibles : un contour plat et haut (H\*H%) ; un contour haut montant (LH\*H%) et un contour plat et haut, avec montée effectuée sur la syllabe pénultième du membre de l'énumération (H+H\*H%). On en conclut donc, comme Selting (2007), qu'on a bien une répétition du type de contour sur les éléments de la liste, mais pas systématiquement de la même variante, bien qu'on observe une réalisation dans 61% des cas de la combinaison H\*H% + H\*H%. Un aspect plus caractéristique de la liste se situe au niveau temporel des syllabes terminales des items de la liste. Sans surprise, les durées des syllabes finales des éléments de la liste sont, comme toute syllabe accentuée, plus longues que la durée syllabique moyenne de la liste. Mais chose plus intéressante, la syllabe du premier membre de la liste est près de deux fois plus longue que la durée syllabique moyenne, sachant qu'une syllabe accentuée « ordinaire » est en principe une fois et demie plus longue que les autres (Delattre 1965). Nous avons montré que cette longueur spectaculaire pouvait être due soit au sur-marquage de l'activité de liste, par projection à partir du premier membre de la liste à venir, soit à l'inverse à l'hésitation du locuteur du fait de la recherche lexicale en cours des éléments à suivre.

Quant à la prosodie de l'item *tout ça* ou *et tout*, elle est diversement réalisée, selon qu'elle mime le contour des éléments qui précèdent, qu'elle se fonde de manière atténuée dans une

---

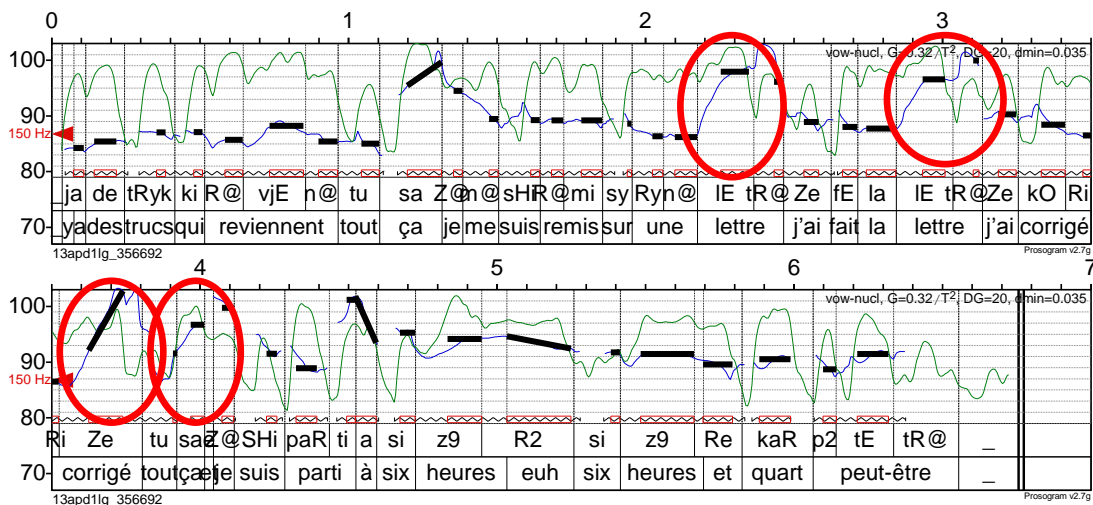
<sup>239</sup> Conventions de transcription selon le modèle F-ToBI (Delais-Roussarie *et al.* 2015) : L = ton bas ; H = ton haut ; \* = accent lexical ; % = frontière prosodique majeure.

période tout en signalant la continuation, ou à l'inverse qu'elle corresponde à une fin de période. Quoiqu'avec des interprétations différentes des nôtres, Ferré (2011) observe une même diversité de réalisations des contours prosodiques pour les « particules d'extension » *et cetera*, (*et*) *tout ça* et *et tout*.

Signalons pour finir que la prosodie peut parfois désambiguïser une analyse syntaxique. Ainsi en va-t-il de l'exemple suivant accompagné de son prosogramme<sup>240</sup> :

(321) y a des trucs qui reviennent tout ça je me suis remis sur une lettre j'ai fait la lettre j'ai corrigé **tout ça** et je suis parti à six heures six heures et quart peut-être (pfc)

Figure 8 : Prosogramme 1



Dans cet exemple, la scansion rythmique par le retour de syllabes exagérément accentuées (dont les contours sont entourés) est très manifeste. La syllabe finale de *tout ça* reproduit le contour intonatif des proéminences des items précédents de la liste (cf. la fréquence fondamentale en trait fin à laquelle se superpose une version stylisée en trait épais). Le syntagme constitue à cet égard un groupe intonatif à lui seul, de la même façon que les membres précédents. Si l'on ne se fiait qu'à la transcription orthographique, on pourrait interpréter *tout ça* comme le complément du verbe *j'ai corrigé*. Si tel avait été le cas, on aurait vraisemblablement noté une absence de proéminence sur la dernière syllabe de *j'ai corrigé*, permettant ainsi de regrouper le syntagme *tout ça* avec ce qui précède, le tout formant ensemble un groupe intonatif. Or, la séparation entre les deux groupes est au contraire bien nette à l'écoute (avec des proéminences sur [Ze] de *corrigé* et sur [sa] de *tout ça*) : *tout ça* est de ce fait clairement exhibé comme le dernier constituant d'une liste, ici composée d'une succession d'énonciations, qu'on pourrait schématiser ainsi (► indiquant une relation de continuation entre les énonciations) :

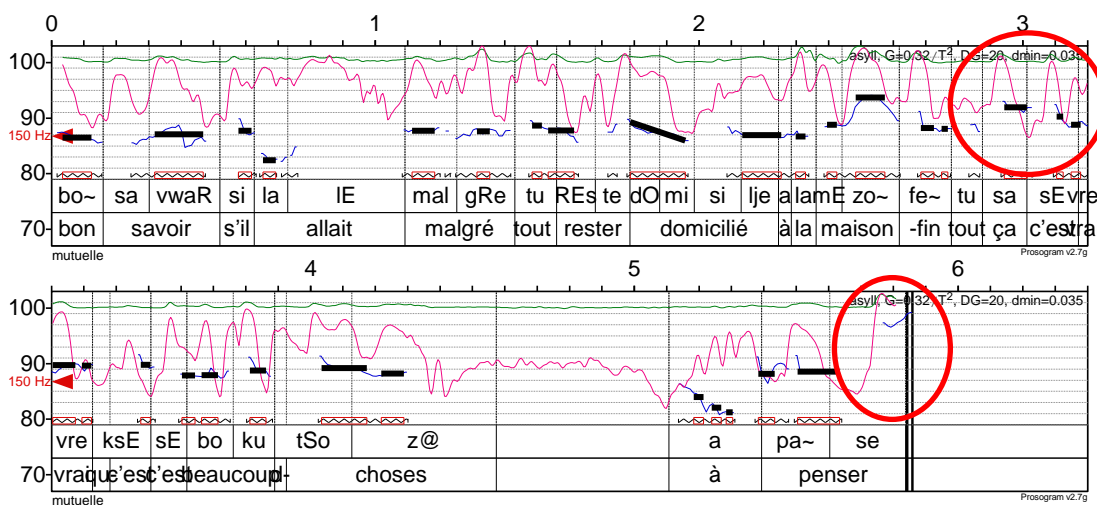
<sup>240</sup> Outil développé par Mertens (2004).

(322) y a des trucs qui reviennent ► tout ça ► je me suis remis sur une lettre ► j'ai fait la lettre ► j'ai corrigé ► *tout ça* ► et je suis parti à six heures six heures et quart peut-être

Il n'en va pas de même pour l'exemple suivant, où la réalisation de *tout ça* n'est pas accompagnée des traits caractéristiques d'une liste, comme en témoigne la figure correspondante :

(323) comme ici en plus il a il a changé de de domicile bon savoir s'il allait malgré tout rester domicilié à la maison enfin *tout ça* c'est vrai que c/ c'est beaucoup de choses à penser (pfc)

Figure 9 : Prosogramme 2



En l'occurrence, *tout ça* ne paraît pas entrer dans un schéma rythmique particulièrement régulier ni présenter les caractéristiques accentuelles d'une liste. En effet, le syntagme *c'est vrai* qui le suit prolonge naturellement le groupe intonatif auquel il appartient (pas de pause, pas de marque particulière d'initiation d'un nouveau groupe : la fréquence fondamentale à l'initiale de la syllabe [sE] se situe au même niveau que la fréquence de la syllabe [sa] qui précède). Par conséquent, *tout ça* est intégré dans un groupe intonatif de plus grande envergure, dont le signal de regroupement est une syllabe accentuée subséquente ([se] de *penser*, en trait fin), qui domine les syllabes précédentes depuis [fe~]. On en conclut que *tout ça* n'est pas aligné sur le paradigme du SV *savoir s'il allait malgré tout rester domicilié à la maison*, mais constitue très vraisemblablement le sujet redoublé de *c'est beaucoup de choses à penser*.

#### 4.4 Fonctionnement référentiel

Nous pouvons à présent nous concentrer sur le fonctionnement de *tout ça* dans ces contextes énumératifs, en nous interrogeant en particulier sur deux propriétés apparemment contradictoires, à savoir son fonctionnement résumptif et sa capacité à évoquer une classe-objet de plus grande envergure.

#### 4.4.1 Un pointage résumptif ?

Au vu de la position terminale de *tout ça* dans les listes et du fonctionnement volontiers résumptif de ses emplois en général (cf. *supra* §3), il y a de bonnes raisons de considérer qu'il englobe dans sa référence les divers objets préalablement introduits. D'ailleurs, d'autres indices de récapitulation se combinent régulièrement à *tout ça*, ainsi que l'illustrent les exemples ci-dessous via les particules *bon voilà quoi* et *enfin* :

(324) L1 : ouais + et les filles c'est le le sac aussi c'est important

L2 : Vanessa Bruno ?

L1 : Vanessa Bruno euh Gérard Darel euh *tout ça bon voilà quoi* (cfpp)

(325) on s'occupait au début de tout c' qui était l'matériel donc extrêmement multiple puisque y avait aussi bien des des couleurs des pastels des clous des *enfin tout ça* (cfpp)

Dans l'exemple (324), la locutrice L1 entreprend l'énumération de marques de sacs à main, à savoir *Vanessa Bruno*, puis *Gérard Darel* et clôt sa liste via le syntagme *tout ça* accompagné des ponctuants *bon voilà quoi* qui permettent de récapituler et de finaliser le propos. L'exemple (325) illustre l'élaboration d'une liste de matériel (*des couleurs, des pastels, des clous*) terminée par *tout ça*, lui-même précédé de la particule *enfin* jouant également un rôle de synthèse (Bertrand & Chanet 2005).

Mais on voit bien que l'analyse uniquement résumptive du marqueur est insuffisante. Trois arguments viennent appuyer ce point de vue. Il y a d'abord le fait que la liste se voit parfois prolongée à la suite du marqueur *tout ça*. C'est le cas de l'exemple (324) ci-dessus, que nous avons délibérément abrégé et dont voici le contexte postérieur :

(324') L1 : Vanessa Bruno euh Gérard Darel euh *tout ça bon voilà quoi*

L2 : que d'la marque qu'on paye cher

L3 : *Longchamp Séquoia et compagnie quoi*

L1 : ouais voilà quoi + (cfpp)

Ce prolongement, par ailleurs effectué par une autre locutrice et approuvé (*ouais voilà quoi*) par la première, inclut explicitement d'autres membres à l'ensemble désigné par *tout ça*. D'ailleurs, le syntagme *et compagnie* renchérit sur cette extension. Cet extrait illustre bien le fait que le réseau évoqué ne se limite pas aux membres cités en amont.

D'autre part, certains exemples contiennent une reformulation lexicale à la suite de *tout ça*, indiquant explicitement un ensemble plus vaste que celui constitué des seuls items préalablement introduits :

(326) donc j'ai eu la chance, là, de euh, de voir ça, sinon, bon après, c'est les petites vacances, qu'on passe, quand on était jeunes, avec les parents en été, bon ben, en l'occurrence c'était Nice, Cannes, *enfin tout ça, toute la Côte d'Azur* (pfc)

Ainsi, *tout ça* désigne allusivement un ensemble sous-déterminé comprenant Nice et Cannes, que le locuteur juge bon de spécifier via la reformulation *toute la Côte d'Azur*. On remarque qu'*enfin* n'est ici pas à interpréter comme une marque de synthèse finale mais plutôt comme un indice de changement au niveau de la planification discursive qu'on pourrait gloser par « j'interromps le détail de ma liste pour aller droit au but ».

Par ailleurs, on a déjà signalé *supra* (ex. (311)) que les listes pouvaient être structurellement constituées d'un seul élément lexical précédant *tout ça*. Il paraît dès lors peu vraisemblable qu'ils soient voués à récapituler, via un marqueur de totalité, un ensemble composé en tout et pour tout d'un seul objet-de-discours.

Au final, il nous semble qu'on peut admettre une analyse résumptive de *tout ça* dans la mesure où il comprend dans sa référence celle des items (de l'item) préalablement mentionné(s). Mais il indique en même temps que la liste entreprise ne reflète pas exhaustivement sa référence, qui va bien au-delà. C'est d'ailleurs cet aspect que mettent en évidence les appellations *general extenders* (Overstreet 1999, Secova 2014) ou *particles d'extension* (Dubois 1993, Ferré 2011).

#### 4.4.2 L'inférence d'un ensemble ?

Dans le cadre de la théorie de la pertinence (Sperber & Wilson 1986), on pourrait considérer que *tout ça* convie le destinataire à enrichir par inférence le contenu d'un message à partir des informations explicites et contextuelles à disposition, dans le but d'atteindre le sens que le locuteur cherche à communiquer. D'une part, le syntagme *tout ça* clôt une liste au niveau structurel, d'autre part, il désigne un ensemble dont la teneur est partiellement implicite au niveau référentiel. Ce faisant, il incite à en inférer d'autres membres potentiels :

(327) c'est vrai que + à Paris c'est cher quand même *tous les cafés tout ça* ça revient cher on a pas trop les moyens (cfpp)

(328) c'est vrai que + à Paris c'est cher quand même *tous les cafés les boutiques les restaurants* ça revient cher on a pas trop les moyens (exemple modifié)

En effet, en fonction du contexte et de l'orientation argumentative du discours, on peut récupérer l'isotopie de départ et en inférer d'autres éléments congruents (Bilger 1989, Ward & Birner 1993). Ainsi, en ne nommant qu'un élément de la liste, le locuteur réalise une économie de moyens au regard des effets escomptés<sup>241</sup>.

Mais à l'inverse, la théorie de la pertinence permet également de rendre compte du fait qu'un traitement inférentiel peut s'avérer trop coûteux, conduisant l'interprète à se contenter d'un référent sous-déterminé pour les besoins de l'échange (Berrendonner 1990a, Overstreet 2005, Jucker *et al.* 2007). L'exemple suivant témoigne de cette situation :

---

<sup>241</sup> On peut aussi invoquer la maxime de quantité de Grice (1975) qui enjoint de ne pas fournir plus d'information que nécessaire.

- (329) SC : Et puis autrement euh, bon bah je sais pas, il m'a demandé un truc sur le euh. Il il redemande le euh, le cholestérol et puis alors les euh. Et puis un ionogramme.  
 E : Un quoi ?  
 SC : Ionogramme, c'est les. <E : C'est quoi ?> Bah l'ionogramme, c'est les euh. Euh, c'est les chlorures, les euh. Les chlorures, euh, **tout ça**, oui les euh.  
 E: Et ça se,(X) je connais pas. Je sais pas à quoi ça sert. (pfc)

L'interlocutrice SC évoque ici visiblement une classe-objet du référent 'ionogramme' dont elle est capable de mentionner l'un des ingrédients, à savoir 'les chlorures'. Etant donné la complexité de ce que la notion de *ionogramme* recouvre (elle désigne un examen médical analysant la teneur d'un liquide organique en électrolytes), la locutrice utilise *tout ça* de manière allusive pour s'éviter d'en formuler le détail (visiblement inconnu) : *tout ça* désigne ainsi l'objet « indiscret » dont 'les chlorures' représente l'un des composants hétéroclites, au demeurant indénumérables. On peut remarquer, à cet égard, que l'interlocutrice E ne se satisfait pas de cette allusion, s'attendant à une explication plus précise : sa réaction nuance l'observation de Dines (1980) selon laquelle l'emploi des *general extenders* génère d'habitude des marques d'approbation de la part de l'interlocuteur. L'interlocutrice E, en l'occurrence, ne « joue pas le jeu ». Cette situation illustre un cas de dissension sur l'état de M : en traçant minimalement un référent, un locuteur n'est pas à l'abri d'éventuelles demandes de clarification.

On reconnaît ici le comportement *nonchalant* de la locutrice (Berrendonner 1990a) qui, pour des raisons d'économie, mais également pour épargner à son interlocutrice des détails jugés non pertinents, se contente d'une évocation vague :

une représentation apparaît d'autant plus économique à manipuler qu'elle est moins complexe et moins analytique. C'est pourquoi il peut être avantageux de ne « tracer » discursivement la pensée que par un minimum de repères cognitifs vagues [...] (*ibid.* : 150).

Par conséquent, le destinataire a tout avantage à se contenter de l'indication minimale d'une classe d'objets entretenant des rapports entre eux, sans chercher à la compléter (voir aussi Jucker *et al.* (2003).

En somme, on constate que selon les circonstances et le calcul de la pertinence, *tout ça* invite ou non l'interprète à inférer d'autres objets « de la même sorte » que celui ou ceux introduit(s). En tout cas, il n'oblige pas systématiquement à une restitution, a fortiori dans les cas où le référent désigné s'apparente à une classe-objet au contenu inanalysable.

#### **4.5 Fonctionnement pragmatique**

Dans cette section, nous observons quelles sont plus précisément les fonctions assumées par *tout ça*, quelles stratégies du locuteur son emploi reflète et à quels effets discursifs il donne lieu.

#### 4.5.1 Marquage de l'intersubjectivité

Le marquage de l'intersubjectivité est régulièrement invoqué pour l'usage des *particules d'extension* (Dubois 1992, Overstreet 1999, 2005, Secova 2014), en ce qu'elles sollicitent de l'interlocuteur une participation active à la construction de représentations communes. Overstreet (2005 : 1851) définit l'intersubjectivité comme « the process by which individuals manage to achieve interpersonal understanding despite subjective differences ». L'auteur insiste sur la présomption de cette communauté de savoir et d'expérience plutôt que sur sa réalité. On constate fréquemment un effet d'empathie, voire de connivence qu'on pourrait gloser par « tu vois ce que je veux dire ». Ci-dessous, l'exemple (330) illustre l'ébauche d'une liste de lieux interrompue par l'évocation de représentations via *tout ça* accompagné de l'adverbe *là*, ce dernier fonctionnant non seulement comme un marqueur proprement spatial, mais aussi comme un désignateur de l'espace mental partagé (*cf. supra* Ch.III §5.2.4) :

(330) ils ont fait des dans j' sais plus quelle année mais dans quelles années c'était + tout plein de plein de HLM en bas du cimetière + dans la rue Gaston Leriaux *tout ça là* (cfpp)

La combinaison *tout ça là*, de par sa nature token-réflexive s'appuyant sur la situation d'énonciation, invite les interlocuteurs à partager leurs représentations cognitives, leurs espaces mentaux. Il en va de même ci-dessous :

(331) c'était un un ensemble d'usines + et d'ateliers + qui était très fréquentes à Montreuil + dans les années cinquante + même + derrière là<sup>242</sup> où ils ont construit où y a une partie de de de de de l'administration qui s'est installée *là* au truc des immigrés *là tout ça* (cfpp)

De nouveau, l'adverbe *là* est utilisé à plusieurs reprises pour mobiliser les représentations communes et se figurer mentalement la scène et fonctionne à ce titre comme particule de clôture ou ponctuant (*cf. supra* Ch.III §5.2.4 et *infra* §4.5.2). Avec *tout ça*, ils contribuent ensemble à solliciter l'activité de co-construction discursive.

Comme on l'a déjà mentionné *supra* (§4.3.2), *tout ça* est susceptible d'intégrer des listes co-construites par plusieurs locuteurs. Pour mémoire :

---

<sup>242</sup> Contrairement aux suivantes, cette occurrence est principalement spatiale et accompagne d'autres expressions contribuant à situer un lieu particulier (*derrière là où...*).

- (332) L1:d't'façon c'est toujours pareil dans une école hein va y avoir un peu euh: les racailles  
 L2: <ouais  
 L1: les> gothiques euh  
 L2: les les  
 L1: les shals  
 L2: voilà [rires] + les bourges voilà <**tout ça** ouais  
 L1: voilà > (cfpp) = (320)

Le syntagme *tout ça* évoque ainsi l'ensemble amorcé collaborativement par les participants et continuant dans un espace commun implicite. A noter qu'on rencontre des co-énonciations de listes terminées par *et tout* également :

- (333) L1 : moi j'habite dans des banlieues j'ai toujours habité dans des banlieues où y a d'plus en plus de de Noirs de gens métissés  
 L2 : <mmh mmh de mélange oui>  
 L1 : **et tout** + et moi au contraire ça m'plaît hein (cfpp)

Tous ces exemples illustrent avantageusement l'invitation ou la participation des interlocuteurs à la construction de représentations communes.

#### 4.5.2 Rôle de ponctuant

A sa fonction d'extension implicite, on peut ajouter le rôle de ponctuant qu'est susceptible d'endosser *tout ça*. Vincent & Demers (1994) définissent un ponctuant comme un élément prosodiquement et pragmatiquement rattaché à ce qui précède<sup>243</sup> qui se dote d'une fonction de structuration du discours : il agit en cela au niveau de la segmentation des unités syntaxiques et prosodiques, mais aussi aux niveaux interactionnels et argumentatifs (Vincent 1993). Secova (2014) insiste sur la vocation de ponctuant du marqueur *et tout*, contrairement à *tout ça*. Ce rôle de ponctuant est bien illustré ci-dessous :

- (334) et pis y a une euh donc y a tout le monde qui part dans | \_ | tout le FBI et la CIA qui partent dans tout dans toutes les hy les hypothèses possibles **et tout** [...] ils sont de de sortie dans un restaurant **et tout** et pis euh | elle monte dans sa voiture **et tout** | \_ | pis elle se prend un petit peu la tête avec son son amant (ofrom)

Tout en signalant qu'elle passe sous silence une partie des détails du scénario policier en question, la locutrice emploie de manière récurrente le marqueur *et tout* à la fin de clauses syntaxiquement complètes. En tant que ponctuant, le syntagme *et tout* sert non seulement de signe de démarcation syntaxique, mais il contribue également à maintenir le rythme et la fluidité du récit, de même que sa progression informationnelle : en se permettant de « réduire » de la sorte les péripéties décrites, la locutrice manifeste sa volonté de passer à autre chose (« je te passe les détails ») et présente celles-ci comme dépendantes d'un

<sup>243</sup> Cf. aussi le ponctuant *quoi* situé, selon Teston-Bonnard, Lefevre & Morel (2011) en fin de rhème (ou noyau) et portant sur celui-ci, contrairement à d'autres particules comme *enfin* ou *bon* qui focalisent ce qui suit.

dénouement à venir (Secova 2014 : 291). Dans l'exemple (334), la séquence narrative, dont nous n'avons pas reproduit la suite, se conclut comme attendu par une chute (l'échec d'une tentative de meurtre) ainsi qu'un commentaire global sur l'intrigue et son auteur.

Le marqueur *et tout* se rapproche d'autres particules discursives servant à scander le discours, comme le ponctuant *quoi*<sup>244</sup>. Il serait intéressant de mettre en perspective les emplois de plusieurs types de ponctuants. On remarque que ces particules peuvent s'avérer totalement absentes chez certains usagers, mais au contraire particulièrement productives chez d'autres<sup>245</sup>, comme chez la locutrice ci-dessus ou encore chez celle-ci :

(335) on avait gardé contact ***et tout*** et puis euh | \_ | comme j'avais eu l'occasion | \_ | enfin il m'avait demandé si je voulais | \_ | si je voulais y aller ***et tout*** et puis comme j'ai eu l'occasion une fois alors je me suis dit | pourquoi pas pourquoi pas profiter ***et tout*** et puis là justement c'était vraiment | \_ | c'était très intéressant parce que il est | \_ | enfin il a fait presque deux semaines euh | \_ | avec moi ***et tout*** il m'a fait visiter euh | \_ | euh quelques villes [...] si j'étais allée seule en touriste vraiment européenne ***et tout*** euh | \_ | on m'aurait voilà on m'aurait fait vraiment découvrir les trucs touristiques ***et tout*** ce que j'ai eu aussi (ofrom)

Dans ce genre d'occurrences, la position syntaxique du paradigme de la liste ébauchée nous échappe, *et tout* enchaînant sur un seul item, souvent au rang d'une clause, si bien que la structure en liste n'est même plus perceptible (Secova 2014).

Si la fonction de ponctuant est bien illustrée par *et tout*, on remarque que *tout ça* n'est pas en reste. Il intervient ci-dessous de manière analogue en tant que « tic stylistique » chez une locutrice<sup>246</sup> :

(336) c'est vrai que + à Paris c'est cher quand même tous les cafés ***tout ça*** ça revient cher on a pas trop les moyens [...] la photographie ***tout ça*** qui nous intéresse plus enfin à notre âge + + ou voilà ou des trucs sur les groupes sur des artistes ***et tout ça là*** on ira voir [...] + oui voilà les tableaux d' Picasso ***tout ça*** voilà pas peut-être pas enfin [...] y a la M.J.C. pas loin y a les conservatoires ***tout ça*** [...] ouais anniversaires alors les évènements ***tout ça*** toutes les soirées on est au courant [...] ben + c'est plus enfin ça c'est enfin si jamais j' vais XX c'est plus quand même sur la réforme pour les lycées ***et cetera tout ça*** les suppressions de postes ***tout ça*** et c'est aussi en général parce que y a quand même enfin plein d'inégalités ***tout ça*** enfin même pour les immigrés ***tout ça*** enfin c'est pour plein d' choses (cfpp)

<sup>244</sup> Voir à cet égard Chanet (2001), Beeching (2007) et Teston-Bonnard, Lefevre & Morel (2011).

<sup>245</sup> Cf. aussi la récurrence de *quoi* chez certains usagers : *je dois rendre un truc mec mais | \_ | totalement inintéressant je dois faire un film avec ça quoi | \_ | sur la base de rien mon gars mais comme c'est chaud quoi ils m'ont dit ouais tu te démerdes quoi ils m'ont | il y a genre ma cheffe elle m'a donné tout le projet elle m'a dit tiens quoi | \_ | demain je dois faire le storyboard tu sais des petites des petites tu dessines une petite BD quoi | puis genre ce que ça va être le film quoi image par image (ofrom).*

<sup>246</sup> Afin de faire ressortir la récurrence de *tout ça* chez une même locutrice, nous n'avons reproduit que ses interventions, mais il faut imaginer qu'elles s'insèrent dans une conversation à plusieurs dont les tours interviennent dans les passages non retranscrits que symbolise le signe [...].

Si l'on compare cet exemple avec les deux précédents, il faut reconnaître avec Secova (2014) que *tout ça* se positionne à la suite d'un « operand » mieux identifiable (souvent un, parfois deux SN) plutôt qu'entre deux clauses qui se succèdent dans une routine narrative. Selon Secova, l'identification de la structure de liste rend le fonctionnement de *tout ça* plus référentiel que celui d'*et tout*. Il nous semble que la présence du désignateur token-réflexif *ça* joue également un rôle dans cette interprétation référentielle du marqueur. Mais à notre avis, cela n'empêche pas *tout ça* d'associer à sa fonction référentielle des rendements pragmatiques apparentés à celui d'un ponctuant.

### 4.5.3 Approximation de paroles rapportées

Overstreet (2005 : 1855) note l'affinité des particules d'extension avec le discours rapporté et Secova (2014 : 291) relève en particulier l'adéquation d'*et tout* dans ce genre de contexte, où il manifeste également un rôle de ponctuant :

(337) c'est vrai que ça été vraiment un accompagnement parce que au début il me disait | \_ | ouais moi au début j'aimais écouter que les voix de femmes mais maintenant j'aime bien les voix d'hommes *et tout* pis | \_ | il a commencé à s'intéresser lui-même | \_ | à ça (ofrom)

(338) elle m' a dit elle voulait venir me chercher elle me dit "ouais j'passe te prendre *et tout*" "oh non non c'est bon j'prends j'prends les transports" et à chaque fois c'était comme ça "ouais mais attends mais tu tu fais trente minutes de train *et tout* tu prends l' RER euh trois changements [...]" (cfpp)

Cet emploi permet de hiérarchiser les types de discours (respectivement les discours hôte et rapporté) tout en signalant l'incomplétude et l'approximation des paroles rapportées. Contrairement à une idée reçue<sup>247</sup>, le discours direct ne garantit nullement la fidélité des propos effectivement tenus. Le marqueur *et tout* est donc particulièrement révélateur de cette contrefaçon des paroles par le locuteur citant qui marque de la sorte ses distances à leur égard et se prémunit par ce biais d'un éventuel reproche d'infidélité. Il indique ainsi la nécessité de relativiser le caractère littéral, exact ou catégorique des propos. On reconnaît ici la fonction de mitigation, dont Sadock (1977) dote les marqueurs d'approximation (*environ, à peu près, approximativement*) :

the role of an approximator in the pragmatic theory is to trivialize the semantics of a sentence, to make it almost unfalsifiable, to hedge in a genuine sense. (p. 437)

Secova (2014 : 291) justifie l'avantage d'*et tout* dans ces contextes par sa brièveté, jugeant plus fastidieux le recours à des marqueurs plus longs comme *et cetera* ou *et tout ça*. Mais c'est oublier que *tout ça* n'est pas plus long à prononcer qu'*et tout*, et qu'il se retrouve également aux confins de discours directs :

<sup>247</sup> Par exemple Grevisse & Goosse (2011 :§414a) : « Le rapporteur les [les paroles] reproduit censément telles quelles, sans les modifier. C'est le **discours** (ou **style**) **direct** ».

(339) donc elle s'est dit elle s'était dit "oui je travaillerai pour avoir de l'argent pour faire ça à mes enfants" **tout ça** et voilà donc euh (cfpp)

(340) ben ils essayaient ils essayaient de s' plaindre en disant aux policiers enfin "vous avez pas le droit **tout ça**" ils disaient "mais mais tu veux que j't'embarque donc bon" (cfpp)

Il est par ailleurs intéressant de remarquer qu'il demeure une incertitude sur la prise en charge énonciative des marqueurs<sup>248</sup>. En effet, on peut se demander si l'on doit mettre l'occurrence d'*et tout* et de *tout ça* au compte de l'énonciateur citant ou de l'énonciateur cité. En écoutant les extraits, on remarque que la contrefaçon mélodique du discours direct est volontiers maintenue sur la particule *et tout* mais non sur *tout ça*. Cela suggère qu'*et tout* fait partie intégrante des paroles mimées, consciemment déformées et réduites : il se fond dans les propos rapportés. A l'inverse, *tout ça* semble traduire un comportement davantage métadiscursif de la part du locuteur citant, qui signale par ce biais qu'il y a d'autres propos analogues à considérer.

#### 4.5.4 Euphémisme

Le rôle de mitigation ou d'atténuation déjà évoqué ci-dessus (Overstreet 2005 : 1855, Secova 2014 : 289) peut s'accompagner de rendements euphémiques, contribuant à adoucir la crudité de certaines réalités ou propos. Ci-dessous, l'emploi de *tout ça*, précédé d'un *enfin* à fonction rectificative, atténue la vulgarité exprimée via l'expression *m'entuber*.

(341) moi la coopé je la paie donc les gens | qui me tapent sur la gueule je leur donne du fric pour le faire les g- enfin | je les paie pour aller faire des séances à Berne je les paie pour réfléchir | à comment m'entuber enfin **tout ça** je les paie et puis je les paie bien (ofrom)

Dans l'exemple suivant, *tout ça* évoque allusivement toute la classe-objet du sujet tabou de la maladie :

(342) Ben, c'était bien, parce que bon, c'était des handicapés, il y avait beaucoup de malades **tout ça** [...] des fois, on connaît qu'il y a des gens qui sont malades **tout ça**. [...] Et puis il y a plein de malades de cancer **et tout** qui souffrent énormément (pfc<sup>249</sup>)

Enfin, signalons un dernier exemple à rendement euphémique avec *et tout* :

(343) [L2 vient de confier à L1 qu'elle a entretenu une relation passagère avec une connaissance mutuelle]

L1 : (ah ouais)<sup>A</sup> (et vous vous êtes embrassés)<sup>S</sup> (**et tout et tout**)<sup>Q</sup>

L2 : (on s'est embrassés)<sup>S</sup> (mais pas **et tout et tout**)<sup>A250</sup> (oral, au vol)

<sup>248</sup> Dans CFPP, les guillemets sont le fait des transcripateurs.

<sup>249</sup> Locutrice de la Réunion dont la langue maternelle est le créole, et le français la langue seconde.

<sup>250</sup> Les exposants reproduisent sommairement la fonction de l'intonation des énoncés telle qu'elle a été perçue : selon l'approche fribourgeoise, S = intonation continuative ; Q = intonation interrogative ; A = intonation exclamative ; F = intonation conclusive.

L'intention de L1 est de suggérer implicitement via *et tout et tout* une succession d'actes typiques (entamée explicitement par la phase du baiser) appartenant au script d'une relation amoureuse passagère, manifestement orientée vers un aboutissement plus intime. Autrement dit, L1 cherche à se renseigner sur la tournure des événements, à propos duquel un questionnement explicite passerait pour indiscret. Ici, L1 invite clairement L2 à une résolution de la sous-détermination, par ailleurs aussitôt opérée : pour preuve, la réplique immédiate et symétrique de L2, qui concède une partie du script, tout en niant subtilement l'épilogue présumé... Dans ce cas, il faut reconnaître le fonctionnement fondamentalement référentiel d'*et tout et tout*, dont la vocation est bien de désigner un référent tabou.

#### 4.5.5 Emphase

Comme on vient de le voir, le marqueur *et tout* est susceptible de se voir rédupliquer, pour marquer une forme d'emphase de l'incomplétude (« et bien plus encore ») par « iconicité » (Overstreet 2005 : 1853<sup>251</sup>). En effet, nous avons observé qu'*et tout* est particulièrement propice au redoublement<sup>252</sup>, comme l'illustre encore l'exemple suivant :

(344) et euh j'suis revenu et puis y a le chef de service qui m' regarde il fait "ouais non euh c'est pas d'sa faute il est comme ça **et tout et tout**" j'fais "ouais mais + XXX non c'est pas des des [...] c'est pas des excuses" j'lui dis "c'est pas + c'est pas une façon d'apprendre + on crie pas sur les gens on est pas des animaux" [...] (cfpp)

Outre l'incomplétude des propos, la réduplication souligne ici l'intensité de l'excuse citée dans la tentative de sauver *in extremis* « la peau » d'un collègue. Nous avons relevé le phénomène avec *tout ça* également :

(345) puis lui c'est là qu'il s'est lancé en disant que c'était terrible euh au Pakistan parce que le | \_ | le gouvernement laïc à l'époque euh | \_ | c'était inacceptable que c'était beaucoup mieux en Iran **tout ça tout ça** | \_ | donc un type euh un fanatique de l'islam (ofrom) = (316)

Dans cet exemple, la locutrice rapporte une partie des paroles d'un individu (cette fois au discours indirect) et glose le reste par la réduplication de *tout ça*, qu'on peut interpréter iconiquement comme l'abondance des arguments délibérément passés sous silence, à l'égard desquels on infère contextuellement la distanciation de la locutrice citante. Bien que cet exemple soit la seule occurrence trouvée avec *tout ça* dans les bases examinées, le phénomène ne semble pas pour autant isolé. Nous avons en effet retrouvé ce genre de réduplication dans des dialogues d'ouvrages de fiction :

<sup>251</sup> Elle relève ce phénomène avec *und so* en allemand et *etcetera* en anglais.

<sup>252</sup> Voire à une répétition plus importante, comme en témoigne cet exemple d'un locuteur algérien : *j'ai abandonné carrément, à l'époque je sortais avec une fille, son père c'était un opticien, bon c'était une belle fille quand même donc euh. Euh, elle était d'Alger, ils étaient d'Alger. Ce qui fait j'ai, j'ai, j'ai eu quand même une petite idée sur le, sur l'optique et tout, mais j'ai, j'ai jamais pensé que je pouvais être un opticien et tout et tout et tout et tout* (pfc, Chlef, Algérie)

- (346) - Andromède ? répéta Raquel en plissant les yeux.  
 Elle portait le pull noir, déjà usé jusqu'à la corde, de l'année précédente.  
 - Ça vient de la mythologie grecque, non ? Le nom me dit vaguement quelque chose  
 mais je ne me souviens pas de l'histoire.  
 - Victime sacrificielle, Persée, la Méduse, **tout ça, tout ça**... répondit Vic à ma place  
 (Gray, C. *Evernight II*)

#### 4.5.6 Association à d'autres éléments

##### 4.5.6.1 *machin tout ça*

Il arrive que *tout ça* (et *et tout*) s'associe à l'emploi de *machin*, généralement lié au code oral (Mihatsch 1996) du fait des conditions de productions favorables (spontanéité, manque de temps, ignorance, etc.) au recours à un terme postiche et passe-partout (cf. *supra* Ch.III §5.1.3) :

- (347) j'ai dit écoute euh | \_ | il faut il faut que tu trouves c'est comme | pour la manière  
 d'apprentissage | mh mh | \_ | \_ | tu sais le le fait de devoir réviser **et tout ça** | \_ | elle me  
 dit maintenant ça va très bien **machin tout ça** euh | c'est vrai | \_ | vrai | c'était | \_ | trop  
 cool (ofrom)

L'association des deux éléments contribue à véhiculer comme précédemment une valeur iconique d'emphase. Mais *machin* semble ajouter à l'incomplétude des propos le marquage de l'indifférence au niveau lexical (Béguelin & Corminboeuf à *par*). A noter à nouveau l'affinité des marqueurs avec le discours rapporté. Il en va de même ci-dessous avec *machin et tout* :

- (348) le coût d'la vie + quand même + quand j'entends + et c'est c'est + dire qu'il faut manger  
 autant de légumes et **machin et tout** mais attendez + quand quand y a quel- y a  
 quelqu'un qui touche trois cent cinquante à quatre cent + (mm) cinquante euros par  
 mois (mm) (cfpp)

##### 4.5.6.2 *tout ça mais*

Les syntagmes *tout ça* et *et tout* entrent régulièrement dans des mouvements concessifs où ils contribuent à appuyer une argumentation dont la conclusion attendue se voit aussitôt contredite, souvent via un énoncé introduit par *mais* (Secova 2014 : 292)<sup>253</sup>. Les marqueurs renforcent ainsi implicitement l'argument concédé en en suggérant d'autres, avant l'intervention d'un dénouement anti-orienté :

- (349) j'aurais pu le placer lui dans il y avait des des accueils pour les enfants pour euh | \_ |  
 dont les mamans travaillaient **et tout ça mais** je trouvais ça vraiment trop (ofrom)

La locutrice évoque l'existence de structures de garde d'enfants pour les mères actives, argument qui favorise une conclusion type du genre 'donc il y avait des raisons d'en profiter'. Le marqueur *et tout ça* contribue ici à élaborer l'argument avancé, en renchérissant sur la

<sup>253</sup> Cf. *and all (that) but* en anglais et *und so aber* en allemand (Overstreet 2005 : 1859).

catégorie des mères actives pour mettre en évidence l'existence de ces accueils. Or, la conclusion implicite, dont l'inefficacité est préfigurée dans l'usage du conditionnel *j'aurais pu le placer*, se voit explicitement contrecarrée par le jugement asserté ultérieurement, exprimant une forme d'excès auquel mènerait l'alternative. Il en va de même dans l'exemple ci-dessous avec *et tout*, qui amplifie les qualités prédiquées (*hyper bien écrit, le personnage vachement intéressant*) sur les livres évoqués par la locutrice qui travaille comme lectrice pour une maison d'édition :

(350) Alors ça si par exemple j'ai eu arrivé de lire des bouquins où euh, euh, mais noirs, sordides, enfin tu vois, un truc euh. et euh, par contre hyper bien écrit euh, le personnage vachement intéressant **et tout mais** euh, j'étais obligé de dire euh, ben non. C'est pas, c'est pas pour vous quoi (pfc, Dijon)

La conclusion attendue par ce renchérissement ('donc je les ai retenus') se voit elle aussi niée au moyen de l'énoncé introduit par *mais*. On peut donc voir dans ces exemples une variante du marquage d'emphase visant à renforcer un argument pour ensuite mieux nier, via *mais*, sa conséquence en mettant en avant un dénouement argumentatif en apparence inattendu.

## 5. Conclusion

Au terme de cet examen des occurrences de *tout ça*, nous espérons avoir montré comment ce syntagme est susceptible de développer toutes sortes de rendements contextuels au-delà de sa fonction référentielle sous-déterminée.

Nous avons tout d'abord mis en évidence les spécificités de son contexte d'occurrence, à savoir les listes. Tout en occupant structurellement sa position finale, il désigne au niveau référentiel un ensemble hétéroclite implicite minimalement ébauché par les premiers items de la liste. Le caractère sous-déterminé et implicite du référent favorise l'apparition de toute une gamme de fonctions pragmatiques. Le marqueur *tout ça* est ainsi susceptible de participer à diverses sortes de stratégies pragmatiques et interactionnelles, telles que le marquage de l'intersubjectivité, la progression et le rythme d'un récit, la synthèse et l'approximation d'un discours rapporté, l'emphase, l'atténuation, etc.

Nous avons relevé çà et là quelques différences d'emplois entre *tout ça* et *et tout*, ce dernier apparaissant plus affranchi d'un rôle référentiel, comme le relève Secova (2014). *Tout ça* s'illustre au contraire davantage dans la référence à un ensemble vague, dont on peut identifier certains ingrédients. La présence du démonstratif *ça*, comme nous l'avons déjà noté, y est sans doute pour quelque chose. Bien que nous reconnaissons cette différence de comportement référentiel entre les deux expressions, il nous importe de nous distancier de l'hypothèse de grammaticalisation postulée par Secova à propos d'*et tout*, illustrant un stade avancé de changement sémantico-pragmatique. Les arguments invoqués sont sa évolution diachronique croissante, sa prédominance chez les jeunes adultes observée dans les corpus et le développement de fonctions pragmatiques au détriment d'un rôle référentiel.

A propos des deux premiers aspects d'ordre quantitatif, il nous semblerait utile de tenir compte du critère de préférence individuelle en vérifiant si certains locuteurs exploitent massivement tel ou tel marqueur ; ce critère est susceptible de biaiser le bilan chiffré de la représentativité du marqueur<sup>254</sup>. A cet égard, il faut noter que le corpus personnel de Secova, celui qui fait ressortir la prépondérance d'*et tout* chez les jeunes adultes parisiens, est constitué de seulement 14 locuteurs (contre 57 dans CFPP et 95 dans le corpus des années 1980). Enfin, concernant le développement des rendements pragmatiques d'*et tout*, on constate qu'il n'est pas propre à ce marqueur, *tout ça* étant susceptible, comme on l'a vu, de contribuer aux mêmes effets. Plutôt que de s'appuyer sur une hypothèse déterministe, il nous semble donc plus prudent et pertinent d'expliquer les motivations contextuelles des différents emplois et marqueurs possibles, en vue d'une vision globale du fonctionnement des particules d'extension. Ceci n'étant pas l'enjeu de notre thèse, nous remettons ce travail à plus tard...

---

<sup>254</sup> Après une recherche grossière sur CFPP, on constate par exemple que 5 entretiens sur 24 comprennent 121 occurrences d'*et cetera* sur les 212 résultats (57%) ; qu'un entretien contient 43 occurrences d'*et tout* sur les 272 résultats de 37 entretiens (16%) ; que 5 entretiens sur 37 contiennent 112 occurrences de *tout ça* sur 333 (34%).



## Chapitre V

### Fonctionnement référentiel et pragmatique de *ILS* à valeur sous-déterminée

*C'est ça qu'ils avaient anéanti, avec leur rapport lapidaire : l'humanité de ma mère. Qui ça ils ? Je ne le savais pas. (Yves Aubrymore, Qui a tué Frajdla Cinnamone ?, p. 51)*



## 1. Introduction

Le pronom conjoint de 6<sup>e</sup> personne connaît un fonctionnement particulier qui se distingue par le caractère sous-déterminé de sa référence. On lui réserve à ce titre des traits caractéristiques<sup>255</sup> : genre masculin, nombre pluriel, position sujet et interprétation humaine du référent, ainsi que l'illustre cet exemple :

(351) quand on fréquentait avec mon mari | \_ | euh on allait se promener jusqu'au bout de la |  
% | de de la Vallombreuse là | \_ | alors y avait c'était l/ | c'était là qu'ils avaient  
construit tous les bâtiments de | \_ | de Mont-Goulin (oral, ofrom)

Le fonctionnement référentiel d'une telle occurrence s'écarte sur bon nombre de points des propriétés typiquement associées au pronom de 3<sup>e</sup> personne dans les ouvrages de grammaire ou de linguistique (*cf. supra* §3 et §3.3) :

- il n'est pas lié à un antécédent textuel
- il n'est pas voué à exprimer la continuité référentielle
- il ne renvoie pas à un référent saillant ni bien déterminé

Un certain nombre de travaux d'origines diverses se sont penchés sur cet emploi de *ils* (et ses interprétations), tantôt décrit comme *indéfini*, *indésigné*, *collectif*, *institutionnel*, *arbitraire*, *impersonnel*, etc. ou sur ses « homologues » à travers les langues. Nous présentons ci-dessous une synthèse critique de ces travaux, en commençant par le traitement que lui réservent les grammaires ou autres usuels du français (§2.1). Nous rendons compte ensuite des études en linguistique française sur la question (§2.2), à la suite de quoi nous rapportons quelques résultats en psycholinguistique (§2.3), avant d'accorder notre attention aux travaux en typologie des langues (§2.4). Nous proposons en seconde partie notre propre approche empirique de la question (§3) portant sur une collection de données attestées en français (§3.1). Notre étude, qui vise à mettre au jour les circonstances d'emploi de *ils*, consiste en un classement commenté des occurrences de *ils* fondé sur les indices contextuels favorisant ou non sa résolution référentielle (§3.2). Pour terminer, nous comparons l'usage de *ils* avec d'autres ressources linguistiques manifestant des rendements pragmatiques similaires (§3.4).

---

<sup>255</sup> Pour une vue d'ensemble des caractéristiques morphologiques et sémantiques du pronom clitique de 3<sup>e</sup> personne en général, voir *supra* (Ch.II §2).



## 2. Travaux antérieurs

### 2.1 *Le traitement de ils dans les grammaires*

Selon Haase (1898), l'ancien français exprimait le « sujet indéterminé » par la 6<sup>e</sup> personne, d'abord dans la seule désinence verbale à la manière du latin, puis à travers la présence du pronom *ils*, utilisé avec un sens analogue au *on* « contemporain ». A son époque, il en relève le caractère populaire : « [c]et emploi du verbe avec ou sans pronom, assez fréquent chez les écrivains du XVI<sup>e</sup> siècle, apparaît quelquefois au XVII<sup>e</sup>, et aujourd'hui même, il est familier à la langue populaire » (*ibid.* : 4). Il cite Vaugelas comme un adepte de cet usage, en particulier de la relative « qu'ils appellent » considérée comme « un latinisme qu'on ne rencontre pas ailleurs » (*ibid.*) :

(352) Quand on cite un livre ou un chapitre, ou que l'on nomme un pape ou un roi, ou quelqu'autre chose semblable, il faut se servir du nombre adjectif ou ordonnant et non pas du substantif ou primitif, **qu'ils appellent**, comme on fait d'ordinaire dans les chaires et dans le barreau. (*Remarques*, I, 215)

Parmi les auteurs du XVII<sup>e</sup> siècle, Haase cite Racine :

(353) Agrippine : J'avouerai les rumeurs les plus injurieuses ;  
Je confesserai tout, exils, assassinats,  
Poison même...  
Burrhus : Madame, **ils** ne vous croiront pas (*Britannicus*, III, III)

Cet exemple se distingue toutefois du précédent, à nos yeux, par le fait que le pronom renvoie à un référent déjà préconstruit par le discours en amont, à travers les structures actanciennes des verbes de parole *avouer* et *confesser*, impliquant sémantiquement un/des destinataire(s) aux propos (à *qqn*), même sans réalisation explicite.

Dans leur *Précis de grammaire historique de la langue française*, Brunot & Bruneau (1949 : 378) nomment « indéfinie » cette valeur de la 6<sup>e</sup> personne, comparable à l'usage latin, qui y aurait recouru à défaut de disposer d'un pronom indéfini. Les auteurs situent cet emploi latinisant comme antérieur au XVIII<sup>e</sup> siècle et l'illustrent notamment par un extrait de Montaigne :

(354) Et chose de trop grande obligation à qui ne peut beaucoup tenir : l'apprêt donne plus à espérer qu'il ne porte. On se met souvent sottement en pourpoint pour ne sauter pas mieux qu'en saye. *Nihil est his qui placere volunt tam adversarium quam expectatio*. **Ils** ont laissé par écrit de l'orateur Curio que, quand il proposait la distribution des pièces de son oraison en trois ou en quatre, ou le nombre de ses arguments et raisons, il lui advenait volontiers, ou d'en oublier quelqu'un, ou d'y en ajouter un ou deux de plus. (*Essais* III, IX, nous élargissons le contexte)

A partir du XVIII<sup>e</sup> siècle, l'emploi est traité différemment : il « est exclusivement populaire [et] renvoie à des personnes mal déterminées, mais connues (et que l'on hait ou que l'on n'apprécie guère) » (*ibid.*).

Il est paradoxal que le phénomène puisse être traité tantôt de latinisme, tantôt d'emploi populaire, selon la période considérée. Un examen diachronique des conditions d'apparition des emplois de *ils* pourrait apporter un regard intéressant sur la question pour savoir s'il s'agit d'usages semblables ou différents. Quoi qu'il en soit, *ils* se justifie aux yeux des philologues jusqu'en moyen français, comme un effet de grammaire seconde conscient (Blanche-Benveniste 1990b), en tant que survivance du latin et en l'absence d'un pronom atone spécifiquement indéfini (*on* étant employé à ce moment-là aussi bien en tant que substantif qu'en tant que pronom indéfini). A partir du XVIII<sup>e</sup> siècle, du fait de l'emploi désormais bien établi de *on* indéfini, mais sans doute aussi pour d'autres raisons à rechercher, *ils* ne bénéficie plus de la même légitimité et est qualifié de populaire.

Ce traitement réservé à l'usage de *ils* se confirme dans les grammaires du français moderne, où il représente le pendant familier et populaire de *on*. Ainsi en va-t-il chez Sandfeld (1970) :

Dans le langage populaire et aussi en style familier, *ils* s'emploie pour *on* : *Ils ne veulent de moi nulle part* D.T. Alp. 129. Le peuple désigne par *ils* tous ceux qui gouvernent, qui dirigent, qui ont de l'influence tout en étant invisibles et inconnus pour la foule : *Ils ne voudront peut-être pas me la donner, la médaille* Duh Mart. 159. *Pourquoi qu'on nous traite comme des veaux ? J'les en préviens ; j'vas tout plaquer ! « Les », c'était l'Etat-major, les généraux, la France* Benj. Gasp. 69. (p. 33)

Le philologue étend toutefois son emploi à la « langue courante ». Dans l'extrait, on notera dans le dernier exemple cité l'emploi régime *les*. Sandfeld ajoute l'attestation de la forme au féminin : « *elles* sert parfois à désigner les femmes en général » : « *Elles sont toutes pareilles* Z. Féc. 496 » (p. 34). Sandfeld suggère quelques hypothèses sur les motivations de ce type d'emplois, telles qu'un intérêt particulier pour des individus qui occupent l'esprit du locuteur, ou « un besoin euphémistique d'éviter de prononcer le vrai nom » (p. 32-33). On peut rapprocher ces usages de ceux qu'il considère comme « sans antécédent » (ou dont l'antécédent est « tout simplement suggéré par le contexte »), comme celui-ci :

(355) (on entend la voix d'une personne endormie par le chloroforme :) Qu'est-ce que c'est que cela ? – C'est l'effet du chloroforme. Tu ne savais pas ça ? Sous l'influence du chloroforme, **ils** se mettent à délirer. Birabeau. Chifferton I. 18. (*ibid.* : 41)

Chez Von Wartburg & Zumthor (1973), l'emploi de *ils* est mentionné comme un « cas extrême » de pronom personnel, relevant de la langue parlée et utilisé « pour désigner l'ensemble très vague des personnes, des êtres auxquels on pense en formulant la phrase » (p. 311). Les grammairiens y associent en outre l'expression d'« une nuance affective de crainte et d'hostilité », le rapprochant par là du rendement euphémique dégagé par Sandfeld. Wagner & Pinchon (1962 : 169) le restreignent également à un registre familier et le dotent de traits

subjectifs tels que l'ironie et le mépris à l'égard de personnes « qu'on ne peut ou qu'on ne veut pas désigner d'une façon explicite ». Parfois, le pronom ne ferait qu'évoquer de manière floue des individus partageant une fonction spécifique délibérément tue :

(356) Vous pourrez demander tout à l'heure au wagon-restaurant s'**ils** en ont (Butor < Wagner & Pinchon : *ibid.*)

Dans leur guide pratique des difficultés, Grevisse & Lenoble-Pinson (2009) accordent un paragraphe particulier à *ils* et y résument les caractéristiques dégagées par leurs prédécesseurs :

ILS, dans la langue parlée familière, s'emploie sans antécédent comme indéfini, souvent dans un sens plutôt méprisant, pour désigner un groupe plus ou moins déterminé d'individus, particulièrement ceux qui détiennent l'autorité : ILS ont fini par l'arrêter (Ac.). – ILS sont en grève (Id.) – ILS ont encore augmenté les impôts ! – « ILS font tout ce qu'ILS peuvent pour nous embêter. » « Qu'est-ce qu'ILS ont encore inventé ? » ILS, ce sont, suivant les cas, ou simultanément, l'Etat, le gouvernement ou le Parlement, la majorité et l'opposition, mais surtout les bureaux. ILS, ce sont ceux qui décident (Peyrefitte). – De même, on dit parfois *elles* pour désigner les femmes en général : Sans ELLES, rien ne serait possible. (Grevisse & Lenoble-Pinson 2009 : 243)

C'est à peu de choses près la caractérisation qu'en font Grevisse & Goosse (2011 : §632), y compris au niveau formel concernant l'absence d'antécédent. A cet égard, les auteurs du *Bon usage* commentent l'exemple ci-dessous en suggérant paradoxalement de « chercher l'antécédent dans la situation qui vient d'être décrite » :

(357) En sortant spontanément au moment d'un accident on m'aurait peut-être considéré seulement comme voisin et mon secours médical aurait passé pour gratuit. S'*ils* me voulaient *ils* n'avaient qu'à m'appeler dans les règles et alors ça serait vingt francs (Céline, *Voyage au bout de la nuit*, p. 403 < *ibid.*)

Les auteurs se gardent bien de fournir des indications sur cette quête de l'antécédent, en orientant par exemple sur le rôle du *on* en amont ou en mettant à disposition un contexte plus large en amont<sup>256</sup>.

Riegel *et al.* (2009 : 361) décrivent dans un premier temps le procédé comme relevant de la référence par défaut (par opposition aux références anaphorique et déictique), à l'instar du pronom indéfini *on* ou d'autres indéfinis. Contrairement à ceux-ci, *ils* ne serait pas susceptible d'une interprétation générique et renverrait « par défaut au référent spécifique le plus immédiatement accessible à partir des informations fournies par le reste de la phrase : *Ils* [= ceux qui ont le pouvoir d'augmenter les impôts = les gouvernants] *ont encore augmenté les impôts* » (*ibid.*). Néanmoins, quelques pages plus loin, les auteurs intègrent ce genre d'emploi au fonctionnement déictique du pronom de 3<sup>e</sup> personne, notamment illustré par cet exemple :

---

<sup>256</sup> Après vérification, il se trouve que le narrateur relate précédemment un grand fracas survenu dans la rue, suivi de gémissements.

(358) **Ils** ont encore augmenté le prix de la vignette (*ibid.* : 366).

L'explication suivante en est fournie : « on interprète contextuellement *ils* comme l'instance collective qui a pris la mesure décrite par le prédicat ». Figure encore à ce paragraphe l'exemple ci-dessous :

(359) J'aime bien cet orchestre : **ils** jouent tous remarquablement (*ibid.*)

Dans ce cas, « l'antécédent collectif se trouve distribué par *ils* sur l'ensemble de ses éléments » (*ibid.*). On peut mesurer la confusion qui règne autour de cet emploi de *ils*, que Riegel *et al.* (2009) traitent tantôt comme une référence par défaut, tantôt comme une référence déictique, ou encore anaphorique (implicitement, *cf.* la notion d'antécédent), à l'appui d'exemples très proches les uns des autres...

En somme, recourant à une conception substitutive ou textuelle des pronoms, les grammairiens se montrent empruntés face à cet emploi de *ils* pour lequel les notions usuelles d'antécédent, de représentant ou d'indéfini, etc. s'avèrent peu ou pas du tout opératoires. Dans un état de langue plus ancien et dans les travaux d'« érudits », l'usage de *ils* s'explique comme une survivance du latin et se voit légitimé de la sorte. Dès lors que l'emploi s'observe chez les usagers ordinaires, il écope d'un certain discrédit de la part des grammairiens<sup>257</sup>. On le ramène systématiquement à des registres de langue peu prestigieux (courant, familier, populaire, etc.) ou au *médium* oral ainsi qu'à la subjectivité du locuteur (hostilité, crainte, mépris, ironie, réserve, etc.), comme en témoigne encore cet extrait :

Le fameux *ils* désigne n'importe quelle catégorie sur laquelle on veut décharger sa colère. (« Le manuel de la grammaire française », <http://www.gabrielwyler.com>)

Dans le même esprit, l'écrivain Pierre Daninos, en simple observateur, élabore une définition sur le ton de l'humour :

*Ils* : troisième personne du pluriel, souvent réduite à une seule lettre (« Y vont encore nous embêter longtemps avec tous leurs trucs ? »), adoptée par les Français pour désigner l'origine de tous leurs maux : députés, percepteurs, communistes, fascistes, piétons, automobilistes, fonctionnaires, gouvernement, Américains, Russes, etc. Tout est la faute de cette troisième personne. Jamais de la première. (Daninos, *Le Jacassin*)

A l'égard des rendements pragmatiques de *ils*, on remarque plus généralement que les grammairiens se concentrent sur le cas particulier d'une identité référentielle délibérément dissimulée. Or, nous avons montré *supra* (§4.1) que la sous-détermination pouvait être due à d'autres facteurs. Ces aspects nous paraissent ainsi trop simplement traités par les grammaires d'aujourd'hui qui se limitent souvent à reprendre leurs prédécesseurs (en plus du contenu, les mêmes exemples sont souvent cités), alors que nous disposons à l'heure actuelle de données

---

<sup>257</sup> Riegel *et al.* (2009) se gardent néanmoins de remarques sur le registre ou sur les intentions du locuteur. A noter encore que nous n'avons trouvé aucune référence à cet emploi chez Wilmet (2010) ni chez Arrivé *et al.* (1986).

beaucoup plus diversifiées qui permettent d'élargir le champ d'observation et de nuancer ces commentaires.

## 2.2 *Le traitement de ils en linguistique française*

Dans la littérature en linguistique française, plusieurs auteurs ont traité de la question, avec des conceptions toutefois assez différentes les unes des autres. À part l'étude de Kleiber (1992b)<sup>258</sup> portant sur l'interprétation de ce cas spécifique, les autres travaux abordent le sujet dans le cadre de problématiques plus générales : le système des marques de 3<sup>e</sup> personne chez Goudet (1983), la polysémie du pronom *il* chez Lebas (1997), le statut des faits jugés déviants chez [Reichler-]Béguelin (1993b). On ne peut donc pas dire que le phénomène ait suscité énormément d'attention dans le domaine français. Nous passons en revue ci-dessous les différentes propriétés dégagées par ces auteurs sur l'emploi de *ils*, au niveau référentiel (§2.2.1) et au niveau morpho-sémantique (§2.2.2) pour en tirer un bilan provisoire.

### 2.2.1 Un référent sous-déterminé

Les auteurs s'accordent généralement sur le fait que l'usage de *ils* concerné implique un référent sous-déterminé, mal identifié, dont on ignore les caractéristiques distinctives, comme l'illustrent ces exemples<sup>259</sup> :

(360) *ils* vont construire un hôtel (Goudet 1983 : 13)

(361) (en remarquant le brouillard) Tu as vu ? On dirait qu'*ils* ont démonté la montagne !  
(Lebas 1997 : 46)

Les observations des linguistes confirment la présence de ce flou référentiel :

La communication orale s'accommode couramment de *ils* collectifs désignant des référents absents de la situation d'énonciation et non mentionnés préalablement, qu'ils soient inférés du contexte ou utilisés, si je puis dire, ex nihilo, en « deixis mémorielle ». ([Reichler-]Béguelin 1993b : 103)

[...] il est très courant et utile d'établir d'autorité, pour certaines actions, l'existence d'acteurs dont on ne connaît rien. (Lebas 1997 : 46)

[...] elle [= la langue] attache du prix à l'indésignation totale : [...] la langue suscite un indésigné pur, d'ailleurs pluriel puisqu'aussi bien le singulier n'apparaît plus essentiel, le *ils* indésigné. (Goudet 1983 : 13)

Néanmoins, le degré ou la manière d'envisager cette sous-détermination varie d'un auteur à l'autre, allant de « l'indésignation totale » (*ibid.*), à la nécessité d'un « assignement référentiel » du groupe en question (Kleiber 1992b). Goudet (1983) n'étudie pas vraiment le processus interprétatif en jeu, se concentrant sur une analyse structurale du système des marques de 3<sup>e</sup> personne. Kleiber (1992b) propose à l'inverse une analyse sémantique détaillée du procédé à l'œuvre, dans une approche restrictive : il requiert pour l'interprétation de *ils* un

---

<sup>258</sup> Cette étude constitue également le chapitre 7 de son ouvrage de (1994b).

<sup>259</sup> Les auteurs ne fournissent malheureusement aucune indication sur le caractère attesté, forgé ou repris des exemples.

domaine de référence circonscrit menant à l'identification du groupe visé (à défaut de l'identification de ses membres) à travers des éléments textuels ou contextuels :

(362) A Paris, *ils* roulent comme des fous (*ibid.*)

(363) *Ils* ont encore augmenté les impôts (*ibid.*)

(364) [A demande à B qui sort du concert]  
- Qu'est-ce qu'*ils* jouaient ? (*ibid.*)

Selon Kleiber, l'ensemble visé correspond à la notion de *groupe* émanant de la typologie de Cruse (1986), qui exige une organisation en *système* : « plus que des attributs en commun, les membres d'un groupe ont un but ou une fonction en commun, ce qui assure à l'ensemble un facteur cohésif plus grand que celui que possèdent les membres d'une classe » (Kleiber 1992b : 337). Sont reconnues de la sorte des entités comme une famille, un jury, un comité, une ville, une équipe, etc. Au contraire, sont exclus de la notion un troupeau, une foule, un groupe (*sic*), la bourgeoisie, etc.

[Reichler-]Béguelin (1993b) critique cette approche en questionnant le bien-fondé des jugements grammaticaux de Kleiber (1992b) à l'égard d'exemples fabriqués dont le « caractère déviant [...] lui permet de « montrer » le fonctionnement de *ils*, et de manifester l'existence supposée d'une contrainte linguistique » (p. 101). Ainsi, en adaptant les données jugées déviantes par Kleiber aux contraintes sémantiques qu'il invoque lui-même, [Reichler-]Béguelin montre l'arbitraire du traitement différentiel auquel devraient aboutir les énoncés ci-dessous (fautif vs grammatical) :

(365) \*J'étais pris dans la foule. *Ils* ont failli m'étouffer (Kleiber 1992b)

(366) J'étais retenu dans la ville. *Ils* ont failli m'étouffer (exemple modifié < [Reichler-]Béguelin (1993b)

Du reste, elle montre qu'il n'est pas rare de relever même dans l'écrit standard des enchaînements pronominaux sur de simples noms d'agrégat ou de classe, réputés impossibles par Kleiber car non conformes à la notion de *groupe* :

(367) Mon fils n'aura pas le chagrin de commander la noblesse de la vicomté de Rennes et de la baronnie de Vitré : *ils* l'ont élu malgré lui pour être à leur tête. (Mme de Sévigné < *ibid.*)

(368) Suite aux élections, je suis atterré de voir nos autorités dans leur léthargie et le peu de courage qu'*ils* ont à traiter ces problèmes (presse < *ibid.*)

Pour Goudet (1983), le cas où *ils* renvoie à un groupe déterminé n'est en fait qu'une éventualité du fonctionnement général de *ils* « indésigné » :

Evidemment, *ils ont encore augmenté les impôts* existe aussi et a pu suggérer que *ils*, représente "les autorités officielles". Mais [...] cette interprétation : ils

= "les autorités" n'est qu'un cas particulier de l'emploi général de *ils* indiquant un groupe imprécis dont *je* est exclu. (p. 13)

Face à ces différences interprétatives, Lebas (1997) propose de distinguer deux sous-types de *ils* : i) les occurrences où l'on pourrait déterminer l'existence d'un référent à l'intérieur d'un cadre circonstanciel et aptes à supporter un prédicat particulier (les exemples de Kleiber *supra* (362), (363), (364)) : dans ce cas, *ils* représente un thème doté de « suffisamment de structure [...] pour postuler son existence extérieure au prédicat » (p. 44) ; ii) les emplois non « thématiques », où le référent n'a pas de « définition objective » (p. 46) et ne s'interprète pas indépendamment du prédicat. L'existence d'un actant indéterminé est en quelque sorte établie *de facto* par l'énoncé même :

(369) *Ils* ont envoyé tous les présidents et les rois à la conférence-anniversaire de l'ONU.  
(mais qui donc ? !) (*ibid.*)

L'idée que *ils* puisse évoquer un référent non déterminé fait écho aux observations de Yule (1982) sur des données en anglais, qui souligne que l'identification référentielle stricte n'est pas toujours nécessaire à l'interprétation du pronom de 6<sup>e</sup> personne et de l'anaphore en général :

(370) Well I saw a demolition order there actually – a few months ago – *they* said *they* were going to demolish some of the flats – which is a pity – I don't know what *they*'re doing with Edinburgh though – as long as *they* don't do what *they* did with Glasgow (p. 319)

Cet exemple illustre à ses yeux une situation de référence indéterminée où le pronom sert de simple indication de l'existence d'un groupe-agent anonyme dont l'identité n'apparaît pas pertinente. L'enjeu, dans l'interprétation de cet extrait, n'est pas d'identifier l'agent, mais plutôt de comprendre les agissements dénotés. Ainsi, plutôt que de porter tous ses efforts sur l'interprétation référentielle d'un pronom atone, il est plus pertinent de diriger son attention vers le prédicat du message communiqué, autrement dit, l'information nouvelle. Plus généralement, l'étude vise à remettre en question la conception traditionnelle de l'anaphore, qui, d'après Yule, n'est pas toujours guidée par un objectif de résolution référentielle.

Au vu de ces différents points de vue sur la question, il semble judicieux d'envisager soit différents types de *ils*, soit qu'il supporte différents degrés de sous-détermination référentielle. Les deux exemples attestés suivants illustrent ces différences :

(371) je lui ai répété plusieurs fois que ça restait une opération | \_ | hein qu'y avait pas de risque zéro mais que c'était une opération entre guillemets banale | \_ | sans risque et couramment courue | \_ | et pis ben il s'est avéré que l'opération a eu des complications que le cheval a été | couché pendant deux jours pis finalement *ils ont dû l'euthanasier* deux jours après (ofrom)

(372) On a vu dans la Liberté<sup>260</sup> qu'*ils ont retrouvé un cadavre* dans l'Aar hier dans l'après-midi dans la zone où on nageait... (courriel, 03.08.2012)

Dans le premier exemple, le contexte global et le prédicat permettent d'inférer une classe dont c'est le rôle d'« euthanasier », tandis que dans le second exemple, ni la situation ni le prédicat ne contribuent à arrêter un type de référent déterminé, le prédicat 'retrouver un cadavre' n'étant pas dévolu à un seul type de référent.

## 2.2.2 Indices morpho-sémantiques

Plusieurs études avancent certaines caractéristiques morpho-sémantiques propres à l'emploi de *ils* sous-déterminé : le nombre pluriel, le genre masculin et les traits [+humain] et [+délocuté] (Goudet 1983, Kleiber 1992b, Lebas 1997). Outre le trait [+délocuté] qui fait consensus parmi les auteurs car inscrit de fait dans le sens de la 3<sup>e</sup> personne, nous évoquons ci-dessous brièvement la manière dont les trois autres traits sont présentés dans la littérature en linguistique française.

### 2.2.2.1 Le nombre pluriel

Le nombre pluriel est constamment mis en avant comme une caractéristique de l'emploi à valeur sous-déterminée du pronom de 3<sup>e</sup> personne. Son interprétation, toutefois, n'est pas traitée de la même manière selon les auteurs.

Chez Goudet, *ils* « indésigné » est par définition pluriel et ne connaît pas d'alternance en nombre, par opposition à son pendant « désigné » qu'il considère comme son homonyme. Cependant, ce pluriel morphologique, comme le *on*, ne serait pas marqué du point de vue sémantique :

(373) *ils* ont ouvert une pharmacie (Goudet 1983)

Selon Goudet, *ils* pourrait se paraphraser ici par *des gens* ou *quelqu'un*, « ce pluriel pouvant éventuellement recouvrir un actant singulier, mais dont on ne sait rien » (p. 13) : l'énoncé « ne signifie pas que deux pharmaciens soient associés ; il ne postule même pas que le pharmacien, d'ailleurs inconnu, soit marié ! » (p. 14).

Dans son étude, Goudet (1983) place *ils* en concurrence avec *on*, les deux formes couvrant « l'autre indésigné » (*ibid.*), indifférent au nombre sémantique. Mais, aux dires de l'auteur, l'emploi croissant de *on* comme 4<sup>e</sup> personne incluant le locuteur (*cf. nous*) favoriserait l'emploi de *ils* comme « indésigné pur » (p. 13). Arnavielle (1984) conteste cependant ce scénario diachronique : se référant à Haase (1898) et Brunot & Bruneau (1949) (*cf. supra* §2.1), il soutient que les emplois de *ils* à valeur indéterminée et de *on* à valeur de *nous*<sup>261</sup> sont attestés depuis longtemps, sans que ces coexistences posent problème : « le français

<sup>260</sup> Il s'agit du nom d'un quotidien régional suisse.

<sup>261</sup> Par exemple dans ce vers de Corneille, « On n'a tous deux qu'un cœur qui sent mêmes traverses » (Polyeucte) où *on* = *nous*, *i.e.* le mari et la femme.

contemporain ne fait que réactualiser, avec des valeurs expressives sans doute différentes, ce que le XVI<sup>e</sup> et le XVII<sup>e</sup> siècles connaissaient déjà » (p. 48).

Pour Kleiber (1992b), *ils* collectif serait marqué en nombre du point de vue sémantique. Il prend d'ailleurs ses distances à l'égard du rapprochement régulièrement fait entre *on* et *ils*. Selon lui, les deux formes ne sont pas substituables dans toutes les situations :

(374) On sonne (on = quelqu'un / des gens) (< Kleiber 1992b)

(375) \*Ils sonnent (= des gens) (< Kleiber 1992b)

L'interprétation « existentielle » de *on* ne peut être rendue pas l'emploi d'un *ils* collectif. Cela montre selon Kleiber que l'emploi de *ils* est à placer du côté de la définitude plutôt que d'une quelconque partition propre aux indéfinis : le pluriel exprimerait une véritable pluralité sémantique correspondant, comme dans les cas du défini pluriel, à une forme de totalité, obligatoirement restreinte par des éléments du contexte (*cf.* la notion restrictive de groupe *supra* §2.2.1).

#### 2.2.2.2 *Le genre masculin*

La marque de genre masculine est également exhibée comme un signe distinctif de l'emploi de *ils* à référence sous-déterminée (Goudet 1983, Kleiber 1992b, Lebas 1997). Selon Goudet (1983), *ils* « indésigné » n'alterne pas plus en genre qu'en nombre, contrairement à son homonyme « désigné ». En effet, le marquage au féminin conduit d'après lui inévitablement à l'interprétation d'un actant déterminé :

(376) *elles* viennent encore de barrer la rue (< *ibid.* : 13)

Dans cet exemple, *elles* « vise incontestablement un actant désigné, des femmes grévistes ou des femmes policiers, des vaches ou des fourmis, des trombes d'eau ou des branches élaguées » dont la valeur est récupérable en contexte. Il constitue le féminin de l'énoncé avec *ils* « désigné ». Ainsi, l'énoncé suivant présente une homonymie :

(377) *ils* viennent encore de barrer la rue (*ibid.*)

Dans son sens « désigné », *ils* peut renvoyer à toutes sortes d'individus déterminés « par exemple des moutons échappés ». Dans son sens « indésigné », *ils* représentent forcément des humains (*cf.* ci-dessous). Si l'approche en termes d'homonymie nous apparaît contre-intuitive et trop coûteuse, on peut toutefois reconnaître que la simple marque du féminin apporte une distinction sémantique supplémentaire à *ils*, pour sa part non marqué.

Dans une approche plus générale du pronom de 3<sup>e</sup> (et 6<sup>e</sup>) personne, [Reichler-]Béguelin n'exclut pas *a priori* le féminin de l'emploi collectif, citant la fameuse formule :

(378) *Elles* sont bien toutes les mêmes<sup>262</sup>

Si la plupart des auteurs excluent ce type d'emplois au féminin des occurrences « collectives », c'est sans doute parce qu'en reflétant une dénomination sous-jacente féminine, la marque de genre contribue à une forme de détermination.

### 2.2.2.3 *Le trait [+humain]*

Enfin, la plupart des auteurs attribuent à l'emploi de *ils* à valeur sous-déterminée un trait exclusivement [+humain], comme déjà anticipé dans la remarque de Goudet (1983) ci-dessus. Le trait [+humain] permettrait notamment de se passer d'une récupération nominale du référent : selon Kleiber (1992b), cet emploi de *ils* se distingue d'autres emplois indirects du pronom par le fait qu'il ne conserve pas le N tête d'un syntagme antécédent, par opposition à l'emploi générique indirect comme ci-dessous :

(379) J'ai acheté une Toyota, car *elles* sont bon marché (<*ibid.*>).

Conformément à la contrainte du trait [+humain], il juge l'exemple ci-dessous mal formé :

(380) Le troupeau avançait doucement. *Ils* broutaient tranquillement de l'herbe.

ou du moins, comme ne relevant pas du même type d'emploi. Kleiber se réclame de l'analyse de Goudet (1983) ci-dessus montrant qu'un changement de genre conduit à un assignement référentiel particulier : au masculin, le pronom pourra renvoyer à des moutons par exemple, et au féminin à des brebis.

Une fois de plus, la restriction au trait [+humain] semble se justifier par le fait qu'une dénomination sous-jacente du référent impliquerait, pour ces auteurs, un référent déterminé, tandis que la référence humaine se passe de l'intermédiaire d'une étiquette lexicale.

A nouveau, [Reichler-]Béguelin (1993b) relativise cette contrainte en considérant l'interprétation humaine comme le résultat d'inférences contextuelles et la prédominance des référents humains comme conséquence de « l'anthropocentrisme des discours ambiants » (p. 106). Elle soutient que l'interprétation en général des pronoms dépend des connaissances partagées des interlocuteurs et qu'un exemple forgé comme ci-dessous apparaît plausible sans qu'il véhicule de trait [+humain] :

(381) A Berlin, *ils sont fermés* le samedi après-midi. (<*ibid.*>)

### 2.2.3 **Bilan**

Au terme de cette synthèse, certains points gagneraient à être éclaircis. D'abord, la diversité des données exploitées et les descriptions proposées mettent en avant l'existence de différents degrés de sous-détermination, voire, de types d'emploi distincts de *ils*. En effet, il est tantôt

---

<sup>262</sup> Cf. aussi la citation *supra* (§2.1) de Sandfeld : *Elles sont toutes pareilles*.

possible d'inférer, via le contexte, des attributs typants du référent dont les membres restent certes anonymes (*cf.* Kleiber 1992b) :

(382) J'ai été hier soir au concert. *Ils* jouaient la 9<sup>e</sup> symphonie. Le directeur m'a appris qu'*ils* la jouaient tous les vingt ans (< *ibid* : 337)

Tantôt, la variable introduite demeure largement sous-déterminée, les seuls attributs dont on dispose étant fournis par le cadre référentiel et le prédicat éphémère :

(383) [En réaction à un article de presse sur le prénommé Tanabe, doyen de l'humanité] Le dernier plaisir de ce fossile vivant consistait à décortiquer une ou deux gambas. Il en mangeait d'ailleurs de moins en moins, car la cuisine à l'huile ne lui réussissait pas... Pauvre Tanabe! Bientôt tu seras accueilli au nirvana : *ils ont installé* à l'entrée un stand de gambas frites où tu te goinfreras à l'œil et là, pas trop d'huile... (Eric Faye, *Nagasaki*, p.17) = (236)

L'une des difficultés majeure de ces emplois est de savoir si les interlocuteurs eux-mêmes ont réellement en tête un référent déterminé ou non :

(384) dimanche on se baignait et aujourd'hui c'est l'hiver non mais *ils* sont fous (oral, au vol dans un secrétariat, 21.05.2015)

En effet, comment décider si la locutrice incrimine avec humour des responsables particuliers, aussi fictifs soient-ils, ou si elle ne fait qu'exprimer son étonnement face aux caprices du temps ?

Comme on l'a vu, certains auteurs abordent la question des *ils* à référence sous-déterminée sous l'angle de l'anaphore, avec des conceptions plus (Kleiber 1992b) ou moins restrictives (Yule 1992, [Reichler-]Béguelin 1993b). Mais peut-on encore parler d'anaphore en (384), lorsqu'aucune valeur stable ne peut être attribuée à la variable introduite ? Il nous paraît dès lors nécessaire, pour étudier la question, de situer le phénomène par rapport à la notion d'anaphore, qui, on l'a vu, représente une problématique bien plus complexe qu'elle n'en a l'air. A cet égard, les travaux présentés dans les deux sections qui suivent évacuent à notre sens un peu trop rapidement la question de l'anaphore.

### 2.3 *Traitement de they en psycholinguistique*

En psycholinguistique, le traitement du pronom de 6<sup>e</sup> personne a fait l'objet de deux études complémentaires sur l'anglais (Sanford *et al.* 2008, Filik *et al.* 2008). Les auteurs appellent le pronom *Institutional They* (Sanford *et al.* 2008) parce qu'il sert d'agent par défaut pour une situation stéréotypée, comme dans *On the train, they served really bad coffee*. Cet emploi est considéré comme un sous-type de pronoms caractérisé comme « sans antécédent explicite ». Les auteurs partent du constat qu'intuitivement, de tels emplois ne paraissent pas poser de problèmes d'interprétation. Ils font l'hypothèse que, par opposition aux anaphores pronominales au singulier, ces pronoms au pluriel n'exigent pas d'identification référentielle

au moyen d'un antécédent. En effet, dans la littérature antérieure en psycholinguistique, on a tenté de montrer que l'absence d'antécédent explicite entraînait pour les pronoms anaphoriques des coûts de traitement cognitif<sup>263</sup> (Sanford *et al.* 1994, Greene *et al.* 1994). Cependant, les expérimentations ne portaient que sur des pronoms au singulier. L'enjeu des expérimentations de Sanford *et al.* (2008) et de Filik *et al.* (2008) est précisément de s'intéresser au cas du pluriel « sans antécédent ». Les deux expérimentations sont dotées de designs distincts, mais elles recourent à un matériel semblable : la première s'appuie sur des mesures de temps de lecture par *eye-tracking* (Sanford *et al.* 2008), la seconde sur des mesures de potentiels évoqués au moyen de l'électroencéphalographie (Filik *et al.* 2008) et toutes deux font varier la présence et l'absence d'antécédent, telles qu'illustrées par ces items :

(385) *a. antécédent explicite, pronom au pluriel*

The in-flight meal I got from the staff was more impressive than usual. In fact, *they* courteously presented the food as well. I made a note to travel with this company again.

*b. antécédent explicite, pronom au singulier*

The in-flight meal I got from the stewardess was more impressive than usual. In fact, *she* courteously presented the food as well. I made a note to travel with this company again.

*c. pas d'antécédent, pronom au pluriel*

The in-flight meal I got was more impressive than usual. In fact, *they* courteously presented the food as well. I made a note to travel with this company again.

*d. pas d'antécédent, pronom au singulier*

The in-flight meal I got was more impressive than usual. In fact, *she* courteously presented the food as well. I made a note to travel with this company again. (nous soulignons)

Les résultats des expériences confirment les intuitions des auteurs : ceux-ci n'observent pas de différence de coûts de traitement entre les conditions avec ou sans antécédent pour le pronom au pluriel, contrairement à la condition au singulier, où l'absence d'antécédent provoque des difficultés de traitement. Pour expliquer cela, les auteurs invoquent deux facteurs en jeu : ils ont recours à la notion de *sous-spécification* (Sanford & Sturt 2002) pour l'interprétation du pronom, celui-ci faisant office d'agent factice (*dummy agent*) induisant une représentation « suffisamment bonne » (*good enough*, Ferreira, Ferraro, & Bailey 2002) pour la compréhension<sup>264</sup>. D'autre part, les auteurs avancent une différence liée au nombre grammatical du pronom : alors qu'un pronom au singulier exigerait une résolution immédiate, à défaut de quoi émergeraient des perturbations de traitement qui mènent à l'inférence d'un référent nouveau, un pronom au pluriel ne requerrait pas d'assignement référentiel immédiat. En effet, les psycholinguistes soutiennent que l'absence de perturbations révèle qu'il n'y a pas

<sup>263</sup> Voir cependant les résultats divergents de Cornish *et al.* (2005) *supra* (Ch.II §4.2.1).

<sup>264</sup> Cf. *supra* (Ch.III §4.3) la notion de *shallow processing*.

de processus inférentiel d'introduction d'un référent nouveau pour l'agent, qui de fait, demeure sous-spécifié.

Il faut ici noter que la plupart des travaux en psycholinguistique, dont ceux-ci, se fondent sur une conception essentiellement textuelle de l'anaphore en surévaluant le rôle joué par l'antécédent (parfois le seul critère déterminant). Or, nous avons déjà montré que la notion d'anaphore ne repose pas sur des critères segmentaux. A cet égard, les stimuli proposés appellent quelques remarques : la paire d'énoncés c) et d) de (385) installe un contexte préalable qui contient déjà l'actant (implicite) en question, à savoir, l'agent du verbe *get*, qui dans son schéma actantiel, implique l'existence de ce dernier. Le traitement des pronoms respectifs ne met donc pas réellement en jeu l'inférence d'un nouvel objet, mais fournit une valeur référentielle à la variable sémantiquement impliquée. En outre, on peut constater en général que pour les besoins des analyses expérimentales, impliquant la maîtrise de tous les paramètres afin de mener à bien les mesures et les statistiques qui en découlent, les items apparaissent dépouillés de tout contexte et dans un environnement aseptisé. D'où l'impression de stimuli souvent caricaturaux qui occultent la réalité et la variété des facteurs en jeu dans les procédés référentiels.

Tout en saluant l'effort entrepris par quelques psycholinguistes pour prendre en considération des phénomènes de sous-spécification et de traitement superficiel (*shallow processing*) – là où la plupart des confrères présupposent un traitement analytique et toujours approfondi du langage – il nous semble donc que les résultats obtenus gagneraient à être complétés par une analyse de terrain, qui tiennent compte des paramètres contextuels à l'œuvre dans les productions réelles des usagers.

## **2.4 Traitement de la 6<sup>e</sup> personne en grammaire générative et en typologie des langues**

### **2.4.1 Interprétation arbitraire**

Dans le cadre théorique générativiste, ce type d'emploi du pronom de 6<sup>e</sup> personne est décrit par Suñer (1983) en termes de référence arbitraire, par analogie avec l'interprétation des PRO (*i.e.* lorsqu'ils ne sont pas contrôlés par un antécédent comme dans *Just PRO to sit here should be forbidden*). Malgré les prédictions contraires de la théorie du *Gouvernement et du Liage*, Suñer étend cette analyse aux pronoms nuls non anaphoriques, ainsi nommés *arbitrary pro*, qui se distinguent des PRO par leur accord en personne et en nombre. L'auteur prend le cas de l'espagnol :

(386) *pro llaman a la puerta*

‘‘they’ are knocking at the door’ (*ibid.*)

Cette construction dite « impersonnelle » présente à ses yeux une interprétation « indéfinie » ou « non spécifiée » (p. 189). Suñer note l'ambiguïté que manifeste cet énoncé entre une lecture arbitraire et une lecture spécifique. Elle considère cependant que dans le cas de la

lecture arbitraire, le nombre de référents demeure indéterminé, malgré l'usage de la marque de pluriel<sup>265</sup>. Plutôt qu'une indication sémantique de multiplicité, elle considère que la marque formelle de nombre sert à indiquer la nature arbitraire de la référence.

Jaeggli (1986) approfondit l'étude de ce type de construction et y ajoute certaines caractéristiques. Il assimile ses conditions de vérité avec celles d'une construction avec un SN quantificateur existentiel en position de sujet comme dans *Someone is calling at the door* : « There is an x, x a human being, such that x is calling at the door » (p. 46). Cette propriété est selon lui importante pour distinguer entre l'interprétation arbitraire et l'interprétation homophone définie impliquant une réelle pluralité sémantique. L'auteur remarque que l'ajout du pronom *ellos* au début de l'énoncé espagnol (386) élimine la lecture arbitraire. Cependant, l'analyse arbitraire ne se limiterait pas aux langues « pro-drop » comme le suggère Suñer (1983), étant donné que l'anglais connaît par exemple ce type d'interprétation avec le pronom *they*. Jaeggli estime en effet que l'énoncé suivant représente l'homologue de la construction espagnole :

(387) They sell cigarettes at all gas stations. (< Jaeggli 1986 : 60)

Dès lors, il baptise la catégorie *arbitrary plural constructions*, incluant de la sorte aussi bien les pronoms nuls que les pronoms visibles véhiculant une interprétation arbitraire.

A la suite de ces travaux, Rizzi (1986b : 509) envisage un indice sémantique nommé *arb index* qui recouvre l'ensemble des propriétés suivantes : [+humain, +générique, ±pluriel], responsable des interprétations *arbitraires* des pronoms. Néanmoins, le trait [+générique], au vu de l'attestation d'interprétations arbitraires dans des contextes épisodiques tels que (386), ne fait pas l'unanimité (Cabredo Hofherr 2003, 2014 : 6-7). Pour Cinque (1988 : 545), l'indice *arb* est susceptible de fournir diverses variantes contextuelles liées au caractère générique ou épisodique du prédicat, respectivement l'interprétation quasi-universelle vs l'interprétation quasi-existentielle. Néanmoins, les observations de Condoravdi (1989) et d'Alonso Ovalle (2000) montrent que ces aspects ne sont pas dépendants, les lectures quasi-existentielles étant compatibles avec des prédicats génériques<sup>266</sup> et inversement, les lectures quasi-universelles avec des prédicats épisodiques<sup>267</sup>. Ces constats amènent Cabredo Hofherr

<sup>265</sup> Jaeggli (1986 : 46) renvoie à la grammaire du castillan de Bello (1847=1977 : 265), qui signale que dans un énoncé comme *cantan en la casa vecina*, il ne faut pas supposer un agent pluriel, puisque cette construction pourrait être produite dans le cas où une seule personne est à l'origine du procès. A nos yeux, ces propos reflètent une conception arithmétique du nombre associant le pluriel à l'addition d'éléments et le singulier à l'unité (cf. *supra* Ch. II §2.3). Or, les faits suggèrent d'envisager une approche moins binaire du nombre.

<sup>266</sup> Lecture quasi-existentielle en contexte générique :

(grec)	To apogevma sinithos	pro poulane	pagoto s afti	ti gonia.
	Le après-midi d'habitude	vendent.3pl glace	dans	ceDET coin
	L'après-midi, ils vendent des glaces à cet endroit. (Condoravdi 1989)			

<sup>267</sup> Lecture quasi-universelle en contexte épisodique :

(esp)	Ayer en España celebraron	el día del	trabajo
	Hier en Espagne pro célébrer.PST.3PL	DET jour de + DET travail	
	Hier, en Espagne, ils ont célébré le jour du travail (Cabredo Hofherr 2014 : 7, adapté de Alonso Ovalle 2000)		

(2003, 2014) à conclure que les interprétations arbitraires des pronoms de 6<sup>e</sup> personne ne constituent pas de simples variantes pragmatiques de l'indice *arb* tel que proposé par Rizzi (1986b) ou Cinque (1988).

#### 2.4.2 Typologie des pronoms R-impersonnels

La comparaison des possibilités d'emploi de la marque de 6<sup>e</sup> personne dans diverses langues pour exprimer le caractère arbitraire du référent incite Cabredo Hofherr (2003, 2014) à proposer une classification nuancée des interprétations en présence, qui ne représentent pas des manifestations d'un phénomène homogène mais requièrent des analyses distinctes. Cabredo Hofherr (2014) recourt à une appellation et une définition de Siewierska (2011) pour caractériser les pronoms concernés :

R-impersonals have the appearance of regular, personal constructions but feature a subject which is human and non-referential. (Siewierska 2011: 57-58)

En se fondant sur une comparaison entre les emplois de ce type de 6<sup>e</sup> personne notamment en anglais, français, russe, espagnol et arabe marocain, Cabredo Hofherr (2014 : 15-16) relève plusieurs cas de figure et élabore une grille typologique que nous adaptons ci-dessous :

**Figure 10** : Lectures des 3pl sans antécédent d'après Cabredo Hofherr (2014)<sup>268</sup>

- i) **Lecture existentielle ancrée (avec ancrage temporel)**  
Tocan a la puerta (esp.)  
(Ils) frappent à la porte (=quelqu'un est en train de frapper à la porte)
- ii) **Lecture existentielle vague (sans ancrage temporel)**  
Ils ont trouvé une moto dans la cour
- iii) **Lecture existentielle inférée (inférence à partir d'un résultat)<sup>269</sup>**  
Aquí han comido mariscos (esp.)  
Ici, (ils) ont.3pl mangé des fruits de mer (= quelqu'un)
- iv) **Lecture corporative (prédicats avec un sujet pragmatiquement contraint aux membres d'une certaine profession)**  
Volvieron a aumentar el IVA (esp.)  
(Ils) ont encore augmenté la TVA
- v) **Lecture universelle locative (permise par la présence d'un locatif)**  
En España hablan español (esp.)  
En Espagne, (ils) parlent espagnol
- vi) **Lecture évidentielle (avec verbe de parole indiquant une information de 2<sup>nde</sup> main)**  
Dicen que en esa casa vivió Darwin (esp.)  
On dit que Darwin a habité dans cette maison

<sup>268</sup> Afin de recouvrir les interprétations d'autres pronoms R-impersonnels, Cabredo Hofherr ajoute une septième catégorie, à savoir la lecture quasi-universelle sans restriction, comme dans *on ne mange pas avec les doigts*, où *on* représente 'les gens en général', sans « restriction de domaine » contrairement aux catégories iv) et v) (p. 15). Quoi qu'il en soit, cette lecture ne concernant pas la 6<sup>e</sup> personne, nous la laissons de côté.

<sup>269</sup> L'énoncé résulte d'une inférence opérée à partir de la perception de certains indices.

L'auteur (2003) justifie ces distinctions d'une part par des critères de comparaison interlinguistiques : par rapport à l'espagnol qui admet toutes les lectures, le français exclut avec *ils* les emplois i) et iii)<sup>270</sup> (auxquels on peut ajouter les emplois vi) au vu des paraphrases données). L'arabe semble quant à lui tolérer l'emploi i) mais pas le iii). L'existence d'une lecture évidentielle (ajoutée dans 2014), résulte des observations du pronom avec des verbes de parole faites par Siewierska & Papasthati (2011) et se justifie par le fait que certaines langues (par exemple le finnois et l'estonien) ne connaissent que cet emploi (Siewierska 2011 : 72). D'autre part, Cabredo Hofherr fait intervenir des critères « distributionnels », en l'occurrence, le type de prédicat, la présence d'un locatif et l'ancrage temporel. Enfin, un aspect important de la typologie consiste en l'opposition effectuée entre interprétations universelle et existentielle.

Quelques aspects de ce classement méritent d'être relevés et discutés. Tout d'abord, on peut se demander si ces interprétations ne recouvrent pas l'emploi de désignateurs sous-spécifiés de manière plus large. Cabredo Hofherr ouvre d'ailleurs la voie en intégrant les marqueurs du type *on-man* (ou similaires dans d'autres langues). Mais plus généralement, on observe que d'autres formes sous-spécifiées se prêtent à plusieurs de ces lectures, comme les arguments implicites des verbes en constructions absolues :

(388) Je la trouve en train de lire Ø, tranquillement assise à la table en bois verni du salon sous le lustre en cristal. (lecture existentielle ancrée) (Radulescu, D., *Un train pour Trieste*, 2010)

(389) Pendant sa longue patience de combattant clandestin, Jacques a beaucoup lu Ø, mais c'est la première fois depuis de nombreuses années qu'il [...] (lecture existentielle vague) (Sonnay, J.-F., *Le tigre en papier*, 1990)

(390) La femme qui sait lire Ø peut ainsi directement entendre le message des Saintes Ecritures. (lecture universelle) (Berriot-Salvadore, E., *Les femmes dans la société française de la Renaissance*, 1990)

On pourrait même envisager la compatibilité de *un N* avec la plupart des interprétations de la grille ci-avant<sup>271</sup>. Ces pistes devraient toutefois être explorées sur des données attestées pour pouvoir être véritablement prises en compte. Mais cela pourrait suggérer que la grille n'est pas propre à un type de marqueur mais plutôt aux marqueurs sous-spécifiés en général, ouverts à plus de possibilités interprétatives que des désignateurs plus spécifiés.

Revenons ensuite aux catégories proposées. Cabredo Hofherr décrit les lectures corporative (iv) et universelle locative (v) comme partageant l'expression d'un groupe maximal via la 6<sup>e</sup>

<sup>270</sup> Autrement dit, Cabredo-Hofherr refuse les emplois du genre *Ils frappent à la porte*, dans le sens *quelqu'un/on frappe à la porte* (lecture spécifique i) ou *Ici, ils ont mangé des fruits de mer* dans le sens de *on a mangé des fruits de mer* (lecture inférée iii).

<sup>271</sup> Cf. *Une personne* frappe à la porte (lecture existentielle ancrée) ; *Une personne* a mangé des fruits de mer (lecture inférée) ; En Espagne, *une personne* parle espagnol (lecture universelle locative) ; *Une personne* est responsable de ses propres agissements (lecture « universelle » ?).

personne, nécessitant toutefois l'intervention d'un élément restricteur du domaine, étant donné qu'*ils* ne peut s'interpréter comme 'les gens en général' (cf. aussi Kleiber 1992b). En cela, ces emplois se rapprochent de l'interprétation des SN définis (Cabredo Hofherr 2014 : 40). Dans le premier cas, c'est la nature sémantique du prédicat qui permet de restreindre le domaine, en identifiant forcément un agent prototypique, en l'occurrence exerçant une même profession. Dans le second cas, le groupe s'interprète par rapport à un lieu évoqué dans le contexte via un locatif. Néanmoins, nous allons voir (§3.2.1) que ces contextes ne sont pas toujours aussi prototypiques et que des indices plus variés interviennent dans le processus d'interprétation, rapprochant ces occurrences de l'anaphore.

A défaut d'un élément restricteur (il nous semble toutefois difficile d'envisager un énoncé sans contexte de production), Cabredo Hofherr range les emplois sous une interprétation existentielle (catégorie i-ii-iii<sup>272</sup>), sauf l'emploi évidentiel (vi). L'auteur rapproche ces emplois de celui des indéfinis dans (2003) mais révisé son analyse en les apparentant dans (2014 : 41) aux arguments implicites<sup>273</sup>. Néanmoins, l'hypothèse d'une analyse existentielle ne nous apparaît pas suffisamment argumentée : elle n'est justifiée que par la représentation logique ou par les gloses à l'aide de quantificateurs comme *quelqu'un vs les gens*, dont on doit tenir l'équivalence pour acquise. Or, pour le français du moins, la lecture existentielle et l'équivalence des paraphrases ne nous semblent pas aller de soi (pour l'emploi (ii), le seul étant présenté comme possible en français). En outre, à notre connaissance, aucun linguiste n'a proposé d'interprétation existentielle du pronom de 3<sup>e</sup> personne en français<sup>274</sup>. Si elle ne nous paraît pas exclue, l'attestation d'une telle lecture nécessiterait à notre sens plus d'indices. Il nous paraît en effet prématuré d'imputer ce fonctionnement au pronom français sur la base d'une analogie avec un quantificateur existentiel ou avec le fonctionnement du pronom dans d'autres langues, sachant que chaque langue possède son propre système de la personne. D'ailleurs, Cabredo Hofherr (2014 : 15) présente l'emploi dit évidentiel (vi) comme échappant à une analyse en termes d'existentialité ou d'universalité, celui-ci n'étant « ni clairement existentiel ni clairement maximalisant », au vu vraisemblablement de sa compatibilité avec des paraphrases des deux types (*Il y a des gens / les gens*). A cet égard, il nous semble que la lecture vague (ii) pourrait mériter la même analyse pour le français.

Certes, la grille typologique de Cabredo Hofherr apporte des éléments tout à fait inédits dans l'interprétation du phénomène en en proposant une vue d'ensemble nuancée, et cela sur la base d'une documentation fouillée de la situation à travers les langues, dont nous n'avons pas

---

<sup>272</sup> Les lectures temporellement ancrée (i) ou vague (ii) se distinguent en fonction de la précision (i) ou non (ii) du point d'ancrage temporel : tandis que dans le premier cas, le point est constant, dans le second, la variable temporelle est existentiellement quantifiée, permettant une restriction d'intervalle (2014 : 13). La lecture inférée (iii), impossible en français selon Cabredo Hofherr, exprime un constat inféré d'une situation sur la base d'indices, souvent de nature perceptive. Contrairement à (i), un ancrage temporel fait défaut, la variable temporelle étant soumise à la quantification existentielle.

<sup>273</sup> Cependant nous montrons ci-dessus (388), (389), (390) que les arguments implicites peuvent conduire suivant les contextes à des interprétations diverses, y compris celles que Cabredo Hofherr considère comme universelles.

<sup>274</sup> Si Goudet (1983) rapproche *ils* de *on* pour le marquage de l'indésignation, il n'invoque jamais une analyse existentielle pour leur interprétation.

idée avant la lecture de ces travaux. Néanmoins, les résultats obtenus à partir de données dépourvues de tout contexte et sur la base de jugements d'acceptabilité demandent à notre sens à être complétés par une approche descriptive des usages réels des locuteurs dans ce domaine.

### 2.4.3 L'hypothèse de la grammaticalisation

C'est l'un des buts que poursuivent Siewierska & Papasthati (2011) à travers l'application de la typologie de Cabredo-Hofherr (version 2003) à un corpus de traductions parallèles du début du roman *Harry Potter and the Philosopher's Stone* (en néerlandais, anglais, français, allemand, grec, italien, polonais, russe et espagnol). Les auteurs recueillent parallèlement des jugements d'acceptabilité sur des données fabriquées par le biais de questionnaires soumis à des locuteurs natifs des langues respectives. Les occurrences du corpus, sélectionnées sur les critères d'absence d'antécédent et de la nature humaine aréférentielle de la 6<sup>e</sup> personne, sont appelées *third person plural impersonal constructions* (3pl IMPs). Elles sont en effet censées refléter un degré d'impersonnalité<sup>275</sup> : les emplois (iv) *corporatif* et (v) *universel locatif* sont ainsi considérés comme semi-impersonnels, car ils fournissent certaines informations sur l'identité de la collectivité impliquée, tandis que les interprétations (i) *ancrée*, (ii) *vague* (iii), *inférée* et *évidentielle* (vi) sont perçues comme complètement impersonnelles, car elles ne donnent, selon les auteurs, aucun indice sur l'actant du procès.

Les résultats du corpus montrent que l'usage des 3pl IMPs est très variable d'une langue à l'autre : les 3pl IMPs sont beaucoup plus fréquents dans la version russe (25% des occurrences de 6<sup>e</sup> personne) que dans la version originale anglaise (8%) ou que dans la traduction française (6,5%, la plus petite proportion !). D'après les questionnaires et les données, les locuteurs préfèrent globalement les emplois semi-impersonnels (*universels* et *corporatifs*) aux emplois dits complètement impersonnels. Les langues qui apparaissent les plus restrictives à l'égard de l'usage des 3pl IMPs se révèlent être l'allemand et le français, à propos desquelles les auteurs soulignent l'existence d'une construction concurrente de type *on/man*<sup>276</sup> pour l'expression d'agents impersonnels ou semi-impersonnels.

Dans une autre étude, Siewierska (2011 : 79) dégage une corrélation entre la nature pro-drop ou non des langues observées et la manière d'exprimer un agent humain dit « non référentiel ». Dans les langues européennes possédant les deux types de constructions MAN-IMPS<sup>277</sup> et 3pl-IMPS (e.g. *on vs ils* pour le français, *men vs ze* pour le néerlandais), les langues non pro-drop – donc à sujet réalisé comme, selon la *doxa*, le français – manifestent selon ses observations une gamme d'emplois plus diversifiée du premier (MAN-IMPS, i.e. *on*

---

<sup>275</sup> Dans une conception sémantique du terme, i.e. dans le cas où l'agent humain d'un procès demeure non spécifié (Siewierska 2008).

<sup>276</sup> A noter que les emplois de *on* (fr.) *man* (all.) et autres « équivalents » dans les autres langues ne se recouvrent pas (François 1984, Siewierska 2011).

<sup>277</sup> Siewierska (2011 : 60) donne la définition suivante : « The term MAN-IMP is used here to refer to an impersonal construction which denotes an unidentified human subject expressed by a word etymologically related to 'human' or 'man', such as *on* in French, *hom* in Catalan and *man* in German ».

pour le français) que du second (3pl-IMPS, *i.e.* *ils*) : cela se voit à travers le fait que dans ces langues, le pronom de 6<sup>e</sup> personne (*ils*) n'exploite pas toutes les valeurs possibles de la typologie, certaines étant prises en charge via les constructions en *on*, *men*, *man*, etc. À l'inverse, les langues pro-drop, ne possédant généralement pas de forme de type MAN ou similaire<sup>278</sup> (*e.g.* le russe, l'italien, l'espagnol, le grec, etc.), disposent d'un éventail d'emplois impersonnels ou semi-impersonnels plus vaste de la marque de 3<sup>e</sup> personne du pluriel<sup>279</sup>.

Ces différences d'emploi sont liées, selon l'auteur, à la nature pro-drop et reflètent un degré relatif de grammaticalisation des formes en question dans les langues respectives. En effet, Siewierska situe les emplois de la marque de 6<sup>e</sup> personne sur une échelle de blanchiment sémantique (*semantic bleaching*), manifestant le plus grand degré de référentialité à gauche et le moindre à droite. Les emplois à droite impliquent l'existence, dans une langue donnée, de ceux à gauche (p. 80) :

a) plural definite > b) plural semi-definite > c) plural non-specific > d) singular specific

L'emploi a) correspond à l'usage anaphorique référentiel de l'indice de 3<sup>e</sup> personne, l'emploi b) *grosso modo* aux emplois universels de Cabredo Hofherr, les emplois c) et d) aux emplois existentiels avec interprétation respectivement plurielle *vs* individuelle. Selon les observations de Siewierska, l'emploi d), c'est-à-dire où le référent est potentiellement singulier<sup>280</sup>, n'est disponible que dans les langues pro-drop. Elle nuance toutefois ce point en admettant une interprétation analogue dans les langues non pro-drop, pour autant seulement que le ou les individu(s) représente(nt) un *organizational grouping* (Myhill 1997), autrement dit, une institution, une organisation ou autre entité collective reconnue. Ses résultats dans (2008) montrent que cette lecture est admise en anglais et en néerlandais, mais, aux dires de ses informateurs, exclue en français, en suédois, en norvégien, en danois, en islandais ou en allemand. Quoi qu'il en soit, Siewierska (2011) soutient que les langues pro-drop ont tendance à exploiter l'échelle dans sa globalité alors que les langues non pro-drop sont plus restrictives à l'égard des emplois de l'extrémité droite :

[...] all the languages that exhibit the right-most stage on the cline [...] are pro-drop. And significantly none of the non-pro-drop languages in which the 3PL is realized by a weak form rather than an affixal one display the singular specific reading. (p. 81)

Se réclamant de Traugott & Hopper (1991), elle estime que la possibilité de manifester des valeurs référentielles aussi réduites qu'à l'extrémité droite de l'échelle est tout à fait caractéristique d'un degré élevé de grammaticalisation des pronoms sujets (Siewierska 2011 : 81). Dans cette perspective, un pronom atone (*i.e.* dans les langues non pro-drop) est

---

<sup>278</sup> Le catalan constitue une exception (forme *hom*).

<sup>279</sup> Mais qui peuvent avoir d'autres moyens de marquage de l'impersonnel comme la construction réflexive impersonnelle (Siewierska 2011 : 85).

<sup>280</sup> Cf. aussi Suñer (1983), Jaeggli (1986), Myhill (1997).

sémantiquement plus contraint par les marques de personne et de nombre qu'une simple marque affixale du verbe, plus susceptible de développer des interprétations alternatives.

L'hypothèse invoquée appelle plusieurs remarques. D'abord, nous avons déjà suggéré *supra* (Ch. II, §2) qu'en considérant les pronoms « clitiques » du français comme des affixes, il était nécessaire de revoir le statut du français par rapport au paramètre *pro-drop* (voire de réexaminer les fondements du concept *pro-drop*). Ensuite, le recours à des échelles interprétatives orientées et par là même, au cadre théorique de la grammaticalisation, nous apparaît discutable. En effet, l'ordre des interprétations est présenté comme un « parcours obligé » (*cf.* la relation d'implication invoquée) ne tolérant pas d'orientation inverse ni de sauts d'étape, comme le déplore Béguelin (2014 : 15). D'autre part, les échelles de ce type se dotent d'une visée prédictive, voire sont promues au rang de lois, formulées comme des contraintes externes au fonctionnement linguistique (*ibid.*). Ainsi en témoigne l'extrait ci-dessous :

Non-*pro-drop* languages may be expected to manifest a more restricted range of 3PL-IMP than *pro-drop* ones. (Siewierska 2011: 81)

Or, sur quoi repose ce genre d'affirmation ? Pour le français par exemple, l'exclusion de l'interprétation d) (*singular specific*) est décidée sur la base d'intuitions d'informateurs (Siewierska 2008). La question semble rapidement évacuée. En effet, on a vu qu'il existait un débat sur cette question en français, Goudet (1983) invoquant le caractère sémantiquement non marqué du pluriel et Kleiber (1992b) requérant au contraire la dimension plurielle du groupe visé. Au vu des positions contradictoires des chercheurs sur la question, il paraît nécessaire d'observer les faits sur le terrain.

#### **2.4.4 Bilan**

Les travaux générativistes ou en typologie des langues offrent une perspective élargie de la question, suscitant des réflexions intéressantes sur la question de l'universalité et de la productivité relative du phénomène<sup>281</sup> et suggérant des pistes d'analyse inédites. Toutefois, la portée des résultats est difficile à évaluer pour plusieurs raisons : premièrement, nous n'avons souvent pas la possibilité de remettre en cause les jugements avancés à l'égard des données typologiques, nos compétences linguistiques étant tout bonnement limitées dans la plupart des langues étudiées ; deuxièmement, une part des linguistes ne s'appuient que sur un très petit nombre d'exemples fabriqués, dépourvus de contexte et soumis à des jugements d'acceptabilité intuitifs (les leurs, ceux de leurs confrères ou d'informateurs des langues étudiées). Dès lors, des conclusions sont parfois hâtivement tirées sur des problématiques qui demanderaient une analyse plus fine, en particulier au sein d'une même langue ; quant à l'examen sur corpus de Siewierska & Papastathi (2011), la démarche empirique et contrastive se révèle pertinente et instructive sur la productivité interlinguistique du phénomène, mais

---

<sup>281</sup> Siewierska (2011) atteste l'usage des 3PL-IMPS dans un grand nombre de familles de langues : outre les langues occidentales, elle mentionne les langues eurasiennes, africaines, amérindiennes, les langues indigènes d'Océanie, d'Australie et de Nouvelle Guinée.

limitée à un corpus écrit d'un genre bien particulier, la fiction littéraire ; en outre, la distinction entre les emplois dits (semi-)impersonnels et les procédés anaphoriques, à nos yeux loin d'être évidente<sup>282</sup>, repose sur le critère segmental de présence ou absence d'antécédent, dont on a relevé les limites à maintes reprises.

## 2.5 Synthèse générale sur les travaux antérieurs

Les travaux se montrent pour le moins hétérogènes sur le plan théorique (la question de la référence en particulier) et méthodologique (jugements d'acceptabilité, données attestées ou non, choix des critères linguistiques, etc.) Par conséquent, l'extension des faits considérés apparaît très variable d'un auteur à l'autre.

On peut dégager *grosso modo* trois positionnements à l'égard de la relation entre *ils* à valeur sous-déterminée et la notion de référence :

- i) L'approche des grammaires du français qui, tout en abordant le problème avec une conception segmentale ou substitutive de l'anaphore, tente d'en distinguer *ils* tant bien que mal.
- ii) L'approche qui situe le problème au cœur des procédés de référence (Kleiber 1992, Lebas 1997, [Reichler-]Béguelin 1993b, Yule 1982). Pour [Reichler-]Béguelin (1993b) et Yule (1982), le phénomène remet en cause la conception même de l'anaphore courante et la question demande à être traitée avec généralité. Chez Kleiber (1992) et Lebas (1997), on observe à l'inverse une tentative de cibler des sous-types particuliers parmi les procédés de référence indirecte, dans une visée plus ou moins restrictive.
- iii) Enfin, l'approche qui évacue la question de la référence et de l'anaphore, soit en ignorant la problématique (Goudet 1983)<sup>283</sup>, soit en récusant purement et simplement le statut anaphorique du phénomène, en raison de l'absence d'antécédent (Suñer 1983, Jaeggli 1986, Cabredo Hofherr 2003, 2014, Siewierska & Papasthati 2011, Siewierska 2008, 2011, Sanford *et al.* 2008, Filik *et al.* 2008).

Quant à la méthodologie adoptée, elle dépend bien entendu des cadres théoriques exploités, mais aussi du choix des données à traiter. Ainsi, une grande part des données en jeu sont des exemples élaborés par le soin des auteurs et donc soumis à leur propre intuition ou à celle d'informateurs (Kleiber 1992b, Lebas 1997, Goudet 1983, Suñer 1983, Jaeggli 1986, Cabredo Hofherr 2003, 2014, Sanford *et al.* 2008, Filik *et al.* 2008, en partie Siewierska 2008, 2011 et Siewierska & Papasthati 2011) ; d'autres données représentent des collections d'exemples recueillis de sources authentiques diverses, mais en nombre limité (Yule 1982 en anglais,

---

<sup>282</sup> *Contra* l'avis de Siewierska (2011 : 81) : « If these [3PL-IMPS in non pro-drop languages] do arise, their grammaticalization is restricted to *semi-impersonal contexts in which they are least likely to be confused with referential 3PL forms* » (nous soulignons).

<sup>283</sup> Dans une optique structuraliste, l'auteur se limite à l'analyse componentielle du système linguistique des indices de 3<sup>e</sup> personne, sans s'intéresser à la procédure interprétative référentielle proprement dite.

[Reichler-]Béguelin 1993b, en partie Siewierska 2008) ; enfin, certaines données sont extraites d'un corpus bien spécifique (traductions parallèles d'un roman chez Siewierska & Papasthati 2011).

Malgré la diversité des approches, des méthodologies, des langues observées et de l'extension des faits, quelques caractéristiques récurrentes se dégagent de ces travaux, à quelques exceptions près<sup>284</sup> :

- a) La sous-détermination du référent de l'indice de 6<sup>e</sup> personne (reflétée par la diversité des adjectifs caractérisant l'emploi, tels que *indésigné, indéterminé, vague, collectif, non-spécifique, (semi-)impersonnel, arbitraire*, etc.) : elle peut être plus ou moins marquée.
- b) Le nombre pluriel, le genre masculin (pour les langues dont le pronom de 6<sup>e</sup> personne alterne en genre)<sup>285</sup>.
- c) L'interprétation humaine du référent.

Comme nous l'avons déjà signalé, pour le français, aucune étude empirique – c'est-à-dire fondée sur des données attestées – n'est entièrement consacrée à la question. Il nous semble donc important de documenter davantage le sujet sous cet angle, afin de compléter, corroborer ou remettre en cause les résultats et partis pris de nos prédécesseurs. Nous problématiserons en particulier la relation entre les emplois concernés et l'anaphore ou les procédés référentiels. Une attention scrupuleuse sera également accordée aux données orales spontanées, dont les conditions de production sont présentées comme particulièrement favorables par de nombreux linguistes (*e.g.* Yule 1982, [Reichler-]Béguelin 1993b, Siewierska 2008) ou grammairiens (Von Wartburg & Zumthor 1973, Grevisse & Goosse 2011).

---

<sup>284</sup> Rappelons que [Reichler-]Béguelin (1993b) se distancie des restrictions sur le genre et sur le trait humain.

<sup>285</sup> On pourrait ajouter la position sujet implicitement admise, au vu des données étudiées. Suñer (1983 : 189) mentionne toutefois cette exigence à propos des seules langues pro-drop. Quant à [Reichler-]Béguelin (1993b), elle rend compte d'emplois en position de régime et de déterminants possessifs.

### 3. Etude de *ILS* dans une collection de données en français

Cette section vise à étudier le comportement de *ils* et ses circonstances d'emploi dans des contextes authentiques où il manifeste un degré de sous-détermination référentielle. A ce titre, il s'écarte de la conception courante de l'anaphore pronominale, ainsi que le montre l'exemple ci-dessous :

(391) j'pense que c'est dans les années alors euh + euh dans les années euh soixante euh soixante-trois quatre cinq ça a commencé + en je je c'était l'arrivée des vélomoteurs du plastic dans en province enfin en province j'me souviens de à la campagne c'était frappant euh ***ils ont pris les anciens meubles ils ont proposé des cuisines en formica*** y a eu tout ce (cfpp)

Malgré la mise en place en M d'un domaine d'interprétation assez circonstancié (années soixante, à la campagne), *ils* ne renvoie pas à un référent déjà établi ni identifié dans M et les prédications qui le concernent ne permettent manifestement pas d'inférer les traits distinctifs d'un référent stabilisé dans les connaissances partagées. A l'inverse, elles invitent plutôt à un traitement sommaire de *ils*.

D'autres occurrences se laissent néanmoins plus volontiers décrire en termes d'anaphore indirecte (voir aussi (371) ou les « *ils* collectifs » de Kleiber comme (382)), à la faveur d'un processus inférentiel (respectivement collectif-classe pour (392) et scénario-actant pour (393)) :

(392) euh enfin la gérance a mis des | \_ | \_ | % | un appareil pour euh mesurer l'hygrométrie pis ***ils ont dit non non tout va bien*** | pis y avait les papiers peints qui se décollaient | \_ | euh ***ils les ont accusés*** de je sais pas | \_ | et pis | l/ le | de t/ | \_ | de d'ouvrir trop leur fenêtre (ofrom)

(393) ne tinquiètes ps!La on fait l'embarquement ms ***ils font un exercice***:enregistrer tout manuellement!Donc on n'est ps prêts d'embarquer!A toute.gros becs (Swiss sms corpus)

Cependant, la limite entre les emplois de *ils* donnant l'instruction d'unifier leur variable avec un objet à inférer et ceux qui demeurent pratiquement inanalysés du point de vue interprétatif nous paraît loin d'être évidente et mérite dès lors qu'on s'y attarde.

Par ailleurs, ce genre d'emplois sous-déterminés n'est pas pris en compte dans les modèles dominants des expressions référentielles. A cet égard, nous avons entrepris dans une étude antérieure (Johnsen 2014) un classement des occurrences de *ils* issues d'un corpus restreint, à savoir 26 enregistrements de la base PFC provenant de 13 points d'enquête de la francophonie. Au total, nous avons classé 218 occurrences de *ils*. Parmi celles-ci, 68% correspondaient à une situation « canonique », c'est-à-dire où la variable est unifiable avec un référent saillant et identifié en M, l'unique cas généralement prédit par les modèles existants. Le reste, tout de même 32%, se répartissait de la manière suivante : 8% renvoyait à un

référent disponible en M mais en retrait (peu saillant), 17% à un référent à inférer, et 7% à un référent largement sous-déterminé. Etant donné la taille du corpus, les chiffres ne peuvent pas être tenus pour représentatifs. Mais mêmes indicatifs, les résultats soutiennent globalement qu'une part non négligeable des faits est laissée pour compte dans les théories de la référence.

Afin de compléter cette étude antérieure, nous nous proposons d'élargir notre champ d'observation à des données provenant de sources plus variées, dans une approche qualitative. Avant d'entrer dans l'analyse proprement dite, il convient de rappeler qu'un procédé de résolution anaphorique typique recouvre la situation où l'objet présupposé par le pointeur se voit unifié avec un objet déjà présent en M. Dans les chapitres précédents, nous avons cependant dégagé plusieurs cas de figure qui s'en écartaient : au chapitre II, la notion d'anaphore indirecte permet de recouvrir la situation où, à défaut d'un objet valide en M, l'objet visé doit être inféré sur la base d'une relation avec d'autres objets présents ; au chapitre III (*cf. supra* §2.3), nous avons en outre énuméré plusieurs situations de sous-détermination référentielle, pour mémoire : i) les cas de cataphore, où l'instanciation de la variable se voit retardée, la valeur à attribuer étant livrée a posteriori ; ii) les cas où la valeur à affecter est absente et doit être conjecturée, mais l'ensemble des valeurs possibles étant vague, la variable demeure sous-déterminée ; iii) les cas où l'unification se fait avec un objet effectivement valide mais dont la composition et la délimitation demeurent incertaines (*e.g.* les anaphores résomptives).

Dans cette étude dédiée à l'indice de 6<sup>e</sup> personne *ils*, nous ne nous intéressons pas au cas de retardement de l'unification i), qui ne met pas en jeu une sous-détermination durable, ni au cas iii), car il concerne la référence à des objets non individués. C'est avant tout le rapport entre l'anaphore indirecte et la situation ii) qui occupera désormais notre attention. Comme constaté ci-dessus, il n'est pas toujours évident de déterminer si *ils* fait effectivement l'objet d'un traitement inférentiel (anaphore indirecte) ou si, la procédure étant trop coûteuse, sa variable demeure non instanciée. Nous tenterons donc de cerner de plus près les situations où le processus inférentiel aboutit à une hypothèse interprétative (§3.2.1) et celles où il n'y aboutit pas (§3.2.3), de même que les cas où l'analyse reste indécidable (§3.2.2). Etant donné les similarités déjà relevées par les auteurs entre l'usage de *ils*, du passif et de *on*, nous proposerons également, une analyse comparative de ces différents moyens de sous-détermination de l'agent (§3.4). Avant d'entamer l'analyse proprement dite, quelques précisions sur les données sont opportunes (§3.1).

### 3.1 Remarques sur les données

Nous avons effectué des recherches automatiques de la chaîne de caractères<sup>286</sup> (PFC et CFPP) ou du mot *ils* (OFROM) via les concordanciers respectifs. Comme indiqué dans l'*Avant-*

---

<sup>286</sup> La recherche par chaîne de caractères, seule possibilité dans PFC et CFPP, ne permet pas de distinguer l'occurrence pronominale *ils* des mêmes suites de lettres dans un mot (comme *files*, *détails*, etc.). En ajoutant un espace à la suite de la chaîne, on réduit le bruit. A l'inverse, l'ajout d'un espace avant se révèle une requête silencieuse, c'est-à-dire qu'elle exclut des occurrences, comme celles précédées d'une apostrophe (*e.g. qu'ils*).

*propos* de cette 2<sup>e</sup> partie, au vu du nombre de résultats<sup>287</sup>, nous nous sommes penchée sur les extraits d'enregistrements les plus féconds en *ils*, pour nous concentrer dans un second temps sur les cas les plus problématiques à analyser du point de vue de la référence.

Nous avons parfois rencontré des problèmes de transcription dans les bases respectives, en raison de situations d'homophonie entre les 3<sup>e</sup> et 6<sup>e</sup> personnes (*e.g. il mange vs ils mangent*) ayant induit une transcription à notre sens erronée. D'autres fois, la situation reste indécidable du point de vue de l'analyse, comme ci-dessous où le transcripateur a privilégié le pluriel, mais où l'on pourrait envisager un singulier étant donné le genre d'émission TV en question<sup>288</sup> :

(394) on regardait là sur ARTE par exemple y a eu cet été euh ++ vers huit heures du soir un ++ euh + des visites de pays où ***il(s) rencontrai(en)t les gens il(s) les regardai(en)t lire et il(s) les regardai(en)t manger*** aussi ++ cuisiner manger (cfpp)

Dans ce cas, le contexte approprié fait défaut pour reconstituer les intentions référentielles du locuteur. Des problèmes d'audibilité dus à la qualité sonore des enregistrements ou à des situations de chevauchement ont également surgi. En cas de doute sur la réalisation effective d'un *ils*, nous avons renoncé à prendre l'exemple en compte. On peut encore relever des cas d'indistinction perceptive entre le masculin et le féminin (*ils vs elles*), où c'est à nouveau le recours au contexte qui permet éventuellement de trancher. Enfin, la réalisation phonétique est parfois si ténue que l'on est en droit de s'interroger sur la présence effective ou non du morphème. Là également, le doute conduit à écarter une telle donnée des observations.

### 3.2 Classement des données

#### 3.2.1 Anaphore indirecte

Comme nous l'avons annoncé, une manière d'analyser certaines occurrences en *ils* est de les traiter comme des anaphores indirectes, mettant en jeu des types d'inférences par ailleurs routinisées (*cf. supra* Ch.I §3.3 et Ch.II §4.2.2). Le pronom conjoint *ils* requiert d'unifier sa variable avec une classe qu'il est possible d'inférer de l'existence d'un individu collectif (§3.2.1.1), d'un lieu (§3.2.1.2) ou d'un scénario impliqué (§3.2.1.3).

##### 3.2.1.1 Inférence collectif-classe

L'« anaphore indirecte collective » (Lammert & Lecolle 2014) est un procédé bien connu : *ils* permet dans ce cas de renvoyer aux membres d'une collection déjà établie en M (*cf. supra* ch. II §2.3). On peut définir une collection comme un ensemble qui se compose d'éléments<sup>289</sup> : le rapport entre les éléments (ou membres) et l'ensemble relève ainsi de l'*ingrédience* (ou *méréonymie*). Ci-dessous, on relève ce genre de rappel à la suite de l'introduction d'une

<sup>287</sup> Au 16 août 2016, PFC compte 7234 occurrences, OFROM 4590 et CFPP 3569.

<sup>288</sup> Par exemple une émission où un animateur part seul à la rencontre d'indigènes.

<sup>289</sup> Certaines définitions sont plus restrictives, ajoutant par exemple l'appartenance des éléments à une même catégorie. Voir Lammert et Lecolle (2014) pour une vue générale sur les N collectifs en français.

collection comme un groupe (395), une famille (396), une clientèle (397), une équipe (398) ou encore une génération (399) :

- (395) y avait une soir/ un soir euh | \_ | jeu de nuit donc c'est **un groupe** qui a dû organiser des un jeu de nuit dans la forêt | \_ | qui faisait assez peur d'ailleurs mais conduit | \_ | hein | \_ | ben conduit organisé alors *ils ont dû / \_ / organiser donc des jeux de nuit* (ofrom) = (132)
- (396) j'ai j'ai encore eu y a pas très m/ très longtemps | une fille de | % | qui a | \_ | elle a juste cinquante ans | \_ | et c'est **une famille où ils parlaient toujours patois** | \_ | elle s'est mariée en Singine avec un Suisse allemand (ofrom)
- (397) [sur une roulotte « take-away »] En ce jeudi de juillet, **la clientèle** défile – *ils sont parfois une douzaine à faire la queue* – et les mets proposés jouent l'éclectisme, entre l'Europe et l'Asie. Ce jour-là, pour moins de 10francs, il est ainsi possible de craquer pour un ragoût aux légumes avec sauce-crème au curry et pâtes, pour du poulet ou des crevettes «sweet & sour» avec riz au beurre, voire même pour une saucisse de veau avec pommes de terre sautées. (presse, *La Liberté*, 24.07.12)
- (398) C'est la rentrée ! Alors si vous avez un petit coup de blues ou besoin d'un moment de bien-être gustatif... venez savourez les bons plats préparés par **l'équipe du food-truck l'Epicurieux ! Ils sont sur le campus** les mardis et les mercredis de 11h30 à 14h00 durant le mois de septembre et viendront tous les jours à partir d'octobre. (Notification Facebook, 25.08.2015)
- (399) L1 : vous ne n'avez pas non plus dans les dans les décalages justement de générations comment vous vous situez par rapport à la + à **la nouvelle génération** est-ce que + vous êtes vous faites partie des gens qui n'supportent plus les incivilités je mets tous les guillemets  
L2 : [...] j' trouve qu'elles [mes filles] ont énormément évolué euh + dans un monde du venant d'une génération quand même euh n'ayant pas grand chose à proposer nous et euh + voilà un monde assez + + avec de moins en moins d' liberté si vous voulez mais j'*les trouve pas incivils* euh + non j' trouve qu'*ils pourraient plus encore même se se* révolter + + j' trouve qu'*ils vivent énormément en groupe* moi c'est ce qui m'aura marquée dans cette génération là c'est qu'ils vivent entre eux quoi et j'trouve que voilà au collègue aussi beaucoup les uns chez les autres (cfpp)

On relève ainsi des types de référents allant bien au-delà de la notion de *groupe* invoquée *supra* (§2.2.2.1) par Kleiber (1992b). On note également *contra* Kleiber que la position régime du clitique est possible (399), de même que la référence à des membres non humains (*cf.* [Reichler-]Béguelin (1993b)), comme l'illustre encore l'exemple ci-dessous, où en l'absence d'indices, on ne connaît d'ailleurs pas le genre du pronom *les* (les 'bêtes', les 'vaches' ?) :

- (400) la vie à la ferme ben en ce moment je suis en plein dedans vu que | \_ | je suis ben justement dans mon année sabbatique alors j'aide mon papa | \_ | euh | \_ | ben en gros

euh | \_ | le matin on se réveille assez tôt pour s'occuper **du bétail les** nourrir **les** traire |  
\_ | on **les** sort on fait des parcs (ofrom)

Il est intéressant de constater que la référence à la classe, via l'indice de 6<sup>e</sup> personne, semble parfois favorisée par l'expression d'un prédicat distributif : en (396), on peut supposer que le prédicat 'parler patois' s'applique aux membres dans un rapport de réciprocité. En (397), c'est bien l'addition des différents membres qui correspond à la collection quantifiée 'douzaine'. De même en (399), on attribuera la propriété de 'vivre en groupe' aux individus plutôt qu'à la collection. C'est encore le cas en (400), où le prédicat 'traire' concerne de la même manière les bêtes prises séparément. Mais la lecture collective est également bien représentée, comme le démontre (395), le même prédicat étant reformulé à propos de la collection, respectivement de la classe ('devoir organiser des/un jeu(x) de nuit'). Il y a donc de bonnes raisons de considérer qu'il s'agit d'une même action collective. De la même manière, on inférera vraisemblablement une interprétation collective du prédicat 'être sur le campus' en (398). Il se peut également que le prédicat demeure non-spécifié de ce point de vue, comme 'être incivil' et 'pouvoir se révolter' en (399), où il n'est pas pertinent à distinguer entre une interprétation collective ou distributive.

### 3.2.1.2 Inférence lieu-classe

Dans cette catégorie, on peut distinguer deux types de classe dont on infère l'existence à partir d'un lieu : la classe des simples occupants ou alors une classe de participants du scénario impliqué par ledit lieu.

#### i) Lieu-occupants

Dans l'extrait suivant, un lieu est introduit via un complément circonstanciel (*dans les pays africains*) servant de cadre pour les prédications qui suivent :

(401) Comme quand on va maintenant dans les pays africains, hein, où c'est qu'elle est la viande, dans les étalages en suspens, hein. Il y a pas de, il y a pas de contrôle sanitaire là, hein. Il y a les mouches qui se collent dessus, mais, **ils ont faim**, donc **ils mangent**, hein. Nous on la toucherait pas peut-être. (pfc)

Les occurrences de *ils* et leur prédication conduisent à inférer une classe correspondante à partir du contexte établi. La routine inférentielle lieu-occupants (*cf. supra* 3.3) offre ici un moule particulièrement propice à l'instanciation de la variable. On remarque pour cet exemple que le SN *les gens* mènerait à une interprétation assez semblable<sup>290</sup>.

Ci-dessous, le lieu est introduit via un circonstant, puis se voit reformulé par un SN, avant que ses occupants soient visés dans une parenthèse :

---

<sup>290</sup> Sur l'emploi des SN *les gens/des gens*, manifestant des indices de réanalyse vers un statut de pronom indéfini, voir Cappeau & Schnedecker (2014).

(402) Grâce à la générosité de mon employeur et au talent de Fabian Cancellara, je passe dix jours en Belgique. Le pays des moules, des frites, du chocolat – si, si, *ils y croient* ! – et des bières évidemment. (presse, *La Liberté*, rubrique *Plage de vie*, 13.04.11)

Outre la routinisation de l'inférence, l'emploi de *ils* au détriment d'un SN de gentilé (*les Belges*) est peut-être motivé par une stratégie d'évitement de redondance lexicale (O3, cf. Ch. I §6.3.6). Quant à *les gens*, on remarque qu'il induirait une différence d'interprétation : le cadre référentiel à considérer se révélerait ambigu (les gens en général ou les gens de la Belgique ?). Selon Cabredo Hofherr (2014 : 11), cela s'explique par le fait qu'*ils* implique nécessairement une « restriction de domaine », ici exprimé par le locatif, facultative pour *les gens* (qui peut signifier *les gens en général*).

Dans l'exemple suivant, le lieu dont les occupants sont évoqués n'est manifestement pas saillant au moment où apparaît *ils*. La locutrice introduit en effet d'autres objets qu'elle place au centre de l'attention et qu'elle est en train d'élaborer lorsque son interlocuteur opère une sorte de « backtracking » référentiel :

(403) E: Donc qu'est-ce que vous voyez cet été, là qu'est-ce qui vous tente ?

CB: Alors en fait, je voudrais partir euh, oui parce qu'en fait euh, quand j'avais dix-huit ans, j'ai gagné, bon l'année dernière hein, j'ai gagné à un hum, un concours, un hum, un prix aventure. Donc j'ai eu une bourse de treize mille francs, et je suis partie euh, je suis allée à Cuba. Il fallait choisir une destination, et euh, enquêter et rendre après un dossier à euh, au déc/ au rectorat. Bon je suis partie, toute seule, en sac à dos, et ça m'a vraiment plu, d'ailleurs je voudrais recommencer cette expérience, avec d'ailleurs Sylvain. Parce que cette année-là il avait pas pu venir parce qu'il faisait un B.T.S. Donc en fait je voudrais partir dans un pays où le niveau de vie est, pas très élevé, hélas, hein, parce que ça revient cher, et euh visiter en même temps, c'est-à-dire se planifier euh des, des visites et, et partir comme ça à l'aventure en sac à dos.

E: Et *ils aiment Fidel Castro euh* ?

CB: Ah non *ils ont très très peur de Fidel Castro*. (pcf)

Afin de répondre à la question de E à propos de ses projets de vacances, la locutrice CB revient sur un voyage fait à Cuba qu'elle a apprécié (*ça m'a vraiment plu*) pour motiver son désir de revivre une expérience similaire mais dans un lieu encore indéterminé (*je voudrais partir dans un pays où le niveau de vie est pas très élevé*). Alors qu'elle est en train de justifier ses critères pour ce prochain voyage (sur lequel était centrée la première question de E), son interlocuteur l'interrompt pour revenir sur l'expérience de Cuba et pour en savoir plus sur les habitants. Au vu de l'argumentation entamée, ce sont selon toute vraisemblance les éléments liés au projet futur qui apparaissent les plus saillants dans M, du moins au plus haut de la « pile » à cet instant-là du discours. On observe à cet égard le rôle déterminant du prédicat 'aimer Fidel Castro' pour l'interprétation du pronom. Cet extrait permet de montrer comment un interlocuteur oriente le discours en ramenant, selon ses propres intérêts, un objet au centre de l'attention. Par ailleurs, il s'attache à maintenir saillant aussi bien les occupants

que le lieu dans la suite de l'échange (*Vous avez vu ce, ce problème qu'ils ont actuellement ? [...] Mais, c'est beau Cuba non ?*).

Il arrive régulièrement que le lieu soit évoqué par le biais d'un simple SP adverbial cadratif, comme dans l'exemple de Kleiber (1992b), pour rappel :

(404) A Paris, *ils* roulent comme des fous. = (362)

Le circonstant, de même que le prédicat 'rouler comme des fous' servent ainsi de « restricteurs de domaine » pour l'inférence du référent. Nous traitons ce genre de procédé comme une anaphore indirecte (*cf.* aussi la qualification d'anaphore « divergente » de Kleiber 1990a), car il met en jeu un processus analogue d'unification de la variable avec une classe inférée du cadre référentiel délivré.

Dans ce genre de configuration, où la portée du cadratif est clairement établie, la commutation avec *les gens* semble bien fonctionner :

(405) PM: Mon père aimait bien discuter en patois avec tous ses copains, hein , oh oui, oui, oui, oui, oh oui, oui, oui. Oh oui, oh oui. Oh oui.

E : Je sais que *dans le Sud ils ont souffert de l'interdiction* des années vingt de parler patois. (pfc)

(406) Je sais que *dans le Sud les gens* ont souffert de l'interdiction (exemple modifié)

Dans l'exemple suivant, le circonstant cadratif consiste simplement en l'adverbe anaphorique *là*, pointant sur le lieu préalablement introduit :

(407) Par contre du côté de la branche euh, maternelle euh, ma mère c'était une Le Glouannec de l'escouade Gouarec elle était du pays Fisel, et *là ils chantaient breton* quoi (pfc)

(408) et *là les gens chantaient breton* quoi (exemple modifié)

Dans cette sous-catégorie, les prédicats formulés ne sont pas non plus univoquement collectifs ou distributifs et reposent sur des inférences contextuelles ; s'il nous semble plus plausible<sup>291</sup> d'envisager une lecture collective des sentiments exprimés – faim (401), admiration et peur (403), souffrance (405) – ou d'une croyance (402), les prédicats 'rouler comme des fous' (404) ou 'chanter breton' (407) nous paraissent à l'inverse compatibles avec l'une ou l'autre interprétation, la propriété pouvant être envisagée comme celle du groupe ou des individus qui le composent.

---

<sup>291</sup> Selon l'étude psycholinguistique de Frazer, Pacht & Rayner (1999) portant sur l'interprétation collective ou distributive d'énoncés dont le sujet se compose de deux SN coordonnés, les lecteurs manifesteront clairement, en cas d'ambiguïté théorique, une préférence pour l'interprétation collective, appliquée par défaut. L'occurrence d'un terme distributif désambiguïsant (*each*) situé après le prédicat entraînerait ainsi une réinterprétation distributive coûteuse de l'énoncé depuis le début. Chez les enfants paradoxalement, l'interprétation par défaut est au contraire distributive selon les auteurs: « For children, the distributive reading is strongly preferred, though perhaps for nonlinguistic reasons involving a conceptual preference for a one-one pairing » (*ibid.* : 101).

ii) Lieu-actant

La situation est un peu différente lorsque les membres ne sont pas de simples occupants du lieu, mais représentent un actant d'un schéma qu'on peut inférer d'un lieu particulier. Dans ce cas, une commutation avec *les gens* serait plus problématique, car elle ne rendrait pas compte du rôle actantiel en question, là où *ils* permet de l'inférer. L'exemple ci-dessous présente deux lieux différents, respectivement un 'hôtel', puis un 'café', auxquels il est possible d'associer des schémas respectifs impliquant une relation entre divers actants typiques :

(409) Et puis voilà, donc le premier soir on a couché à l'hôtel, parce que l'appartement qu'on avait réservé n'était pas libre, donc on a couché à l'hôtel. Alors l'hôtel, les bagages, ***il(s) voulais(en)t pas nos bagages***, c'était pas l'heure. Alors que normalement, c'était à onze heures du matin, ***ils nous ont dit*** : 'Non, à cinq heures de l'après-midi.' On a dit: 'Mais non, vous plaisantez.' Alors, quand même ***ils ont accepté de les mettre dans une petite pièce***, là-bas au fond [...] Eh bien, en arrivant là-bas, eh ben (rires), on est allé chercher un café pour euh déjeuner. Mais en Italie, c'est qu'on s'assoit pas pour déjeuner, c'est tout au, au bar. Oh ben, on a dit, 'Non, on vient de passer une nuit, on veut au moins un petit-déjeuner confortable (rires), s'asseoir.' Ben, on a eu de la peine à trouver un espèce de petit recoin au fond de, d'une salle. ***Ils nous ont servi un café***, alors moi, pas du tout à la française, alors le café, hum, pas digéré du tout ça (pfc)

Parfois, un même lieu engage simultanément plusieurs types de classes à reconstituer à partir du contexte. On notera en particulier le rôle des prédications respectives pour distinguer entre les différents référents évoqués par les occurrences de *ils* :

(410) le problème c'est que en Australie sans voiture | \_ | t'es dans la dèche [...] mais ***ils ils vont pas genre comme chez nous à perpète-les-joints avec le train*** quoi [...] donc euh la voiture c'est c'est quelque chose d'assez indispensable | \_ | d'ailleurs euh ***ils peuvent*** | \_ | ***avoir leur permis*** | \_ | à seize ans | \_ | mais ça pose d'énormes problèmes parce que déjà ici à dix-huit ans les gamins ***ils ont pas forcément*** | \_ | la maturité euh | \_ | pour comprendre la puissance d'une voiture [...] enfin voilà quoi en plus euh | \_ | avec les kangourous qui traversent la route rien à voir euh ça aide pas non plus | \_ | t'as meilleur temps de pas aller trop vite | \_ | mais alors maintenant ***ils ont instauré des nouvelles lois*** genre euh ils ont pas le droit d'avoir de trop grosses cylindrées | \_ | heureusement | \_ | euh | \_ | ils ont pas le droit d'être plus que deux dans une voiture après neuf heures le soir | \_ | pour pas qu'ils euh | \_ | enfin pas qu'ils aillent faire la foire [...] bon voilà ***ils ont mis des restrictions*** par rapport à ça (ofrom)

Dans les séquences en gras romain, le pronom semble renvoyer, comme dans la sous-section précédente, aux occupants, autrement dit aux 'Australiens', auxquels sont octroyés des prédicats génériques, *i.e.* qui généralisent un état-de-choses (Carlson 2006). Le locuteur restreint ensuite explicitement cette classe dans les passages soulignés au moyen de plusieurs attributs (âge, dénomination, prédicats, etc.), qu'on peut identifier comme les 'jeunes conducteurs australiens'. Mais dans l'intervalle et de manière inopinée, il est fait référence à

une autre classe via *ils* : la variable n'est en effet unifiaible, au vu du prédicat 'instaurer des lois', ni avec les 'occupants australiens', ni avec les 'jeunes conducteurs'. Il s'agit donc d'inférer la classe qui correspond à un actant impliqué dans un schéma vraisemblablement lié à l'administration du lieu en question. On observe dans la suite du passage que le locuteur alterne de manière imprévisible entre cette classe et celle des 'jeunes conducteurs', seuls les prédicats respectifs permettant de faire la part des choses. Il est intéressant de remarquer que malgré ce qu'on pourrait malencontreusement prendre, en raison de l'absence de redénomination, pour une *chaîne de référence* (Chastain 1975, Corblin 1995, Schnedecker 1997, Schnedecker & Landragin 2014), le réseau référentiel à l'œuvre se construit sur le vif de manière assez instinctive, fondé sur une part d'implicite et sur une sorte de « sélection naturelle » opérée par les prédicats. On est donc loin de l'identification de « chaînons » distincts, établis et contrôlés que suppose la notion de *chaîne de référence* (cf. *supra* Ch.II §3.3).

On pourrait multiplier les exemples avec toutes sortes de lieux dont on peut inférer des scénarios :

(411) voilà ben je me suis inscrit euh | \_ | d'abord au conservatoire de de | # | et puis | \_ | vu mon âge avancé euh | \_ | ***ils ils hésitaient à me prendre*** parce que y avait pas beaucoup de place enfin c'était assez difficile (ofrom)

(412) J'attends la réponse de l'hôpital. J'ai expliqué ta situation. ***Ils vont me rappeler***. As-tu mal à la tête? Bises (Swiss sms corpus)

Dans cette catégorie, la question d'une interprétation collective/distributive offre un éclairage nouveau. En effet, l'emploi du verbe à la 6<sup>e</sup> personne permet de diffuser collectivement et indistinctement la responsabilité d'un éventuel exécuteur sur l'ensemble des agents associés à l'un des scénarios propres à un lieu. En d'autres termes, l'interprétation est compatible avec la situation où un seul individu a pu agir au nom de l'ensemble. En cela, la structure en *ils* + V s'oppose à une structure causative en *faire*+inf. (également appelée factitive) qui procède par dédoublement des agents : celle-là distingue en effet un *instigateur*, « qui promeut l'action » d'un *exécuteur*, « qui la réalise » (Simone 2014 : 347) (e.g. *le professeur a fait nettoyer le tableau par les élèves*). La structure causative est donc une diathèse qui permet de complexifier un procès en augmentant le nombre d'agents. Parmi ses fonctions pragmatiques, on peut noter celle de marquer l'autorité de l'instigateur sur l'exécuteur, et réciproquement, celle de diminuer la responsabilité de l'exécuteur (*ibid.*). A l'inverse, l'emploi de *ils* dans les contextes présents permet de diluer la responsabilité d'un éventuel exécuteur dans une source diffuse. On peut ainsi voir la structure en *ils* comme le pendant de la diathèse causative.

Cette observation nous permet de mieux comprendre pourquoi les locuteurs peuvent recourir à *ils* alors qu'un seul *exécuteur* est envisageable : *ils* projette ainsi la responsabilité sur une entité plus diffuse comprenant potentiellement plusieurs niveaux d'agents. Pour mémoire, de nombreux auteurs remarquent qu'il est possible d'utiliser la 6<sup>e</sup> personne dans des cas où un

seul individu est concerné par la situation décrite (Goudet 1983, Suñer 1983, Jaeggli 1986, Myhill 1997, Siewierska 2008). Siewierska (2008) soutient cependant que cet emploi n'est pas possible en français, contrairement à l'anglais et au néerlandais qui l'autorisent dans la mesure où l'individu est considéré comme le représentant d'un organisme (Myhill 1997 : 807-808) :

(413) **They** accused me of coming to this country just to get rich !

It must be assumed that *they* refers to the entire group of people in whose name the action was done, even if in a particular case only one person may have literally done the action (*ibid.*)

Contrairement à l'avis de Siewierska, cette situation vaut donc aussi pour le français, comme en témoignent bien les exemples de la section.

### 3.2.1.3 Inférence scénario-actant

Dans cette catégorie, les occurrences de *ils* requièrent l'inférence d'une classe correspondant à l'actant d'un scénario. Ici non plus, la commutation avec *les gens* n'aboutit pas systématiquement à une interprétation similaire, car sa sous-spécification invite moins spontanément à l'inférence d'un rôle actantiel du référent. Dans l'exemple ci-dessous, une locutrice raconte les circonstances d'un saut en parachute, impliquant un même référent dont l'identité se précise au fil des prédications et dont les pointages successifs contribuent à augmenter la saillance :

(414) c'était vraiment sympa on a fait aussi un saut en parachute à San Diego | \_ | c'était assez euh | \_ | assez incroyable surtout que | \_ | ouais déjà on arrive là-bas euh | \_ | **ils nous ont presque pas donné d'explications** | \_ | et pis euh tout d'un coup **ils nous mettent un** | \_ | **un espèce de gros euh** | \_ | **un espèce de gros baudrier** avec deux grosses sangles qui passent sur les épaules | \_ | et pis euh | \_ | **ils nous font ah ben on va dans l'avion** machin donc euh on nous on s'y attendait trop pas | \_ | pis après on monte dans l'avion euh il commence à | \_ | à décoller c'était vraiment un petit avion on était quatre dans l'avion mais l'avion il était plein | \_ | on aurait pas pu être une personne de plus | \_ | et pis euh l'avion il monte et pis on on était euh | **il(s) nous donne(nt) les explications** du saut dans l'avion et alors qu'y a | plein de bruit | \_ | ouais c'était assez euh un peu à l'arrache comme ça | \_ | et pis euh après à quatre mille mètres **il ouvre** la porte et pis hop | on saute euh on avait juste des petites lunettes on n'avait rien du tout pour euh | \_ | protéger les oreilles ou comme ça (ofrom) = (101)

Il est intéressant de constater qu'au début de la narration de l'épisode, la locutrice évoque via la 6<sup>e</sup> personne un ensemble, mais qu'elle en extrait par la suite<sup>292</sup> un seul individu, peut-être l'exécuteur du scénario. Mais l'on pourrait émettre une autre hypothèse : celle selon laquelle

<sup>292</sup> Peut-être déjà dès l'avant-dernière occurrence, étant donné l'homophonie de la forme verbale (3<sup>e</sup> ou 6<sup>e</sup> personne). La dernière occurrence est quant à elle clairement réalisée sans liaison, donc très probablement marquée au singulier.

le référent évoqué d'abord sous forme d'ensemble soit composé en tout et pour tout d'un seul membre.

Cette hypothèse remettrait en question la conception traditionnelle du pluriel comme simple addition d'éléments (*cf. supra* Ch.II §2.3) déjà affaiblie par d'autres faits, comme les phénomènes d'accords liés aux noms collectifs, les interprétations des 4<sup>e</sup> (*nous*) et 5<sup>e</sup> (*vous*) personnes, le cas des *pluralia tantum*, etc. Ces faits mettent notamment en évidence la possibilité d'interpréter le pluriel comme un ensemble ( $\approx$  pluriel interne). Or, cette conception n'exclut pas *a priori* que l'ensemble considéré soit constitué d'un seul élément, comme peut-être ci-dessus (voire d'aucun élément, *cf.* la notion d'ensemble vide). Pour certains auteurs (Suñer 1983, Goudet 1983, Siewierska 2008, Cabredo Hofherr 2003, 2014), la possibilité de cette lecture « singulière » témoigne du caractère non marqué du pluriel<sup>293</sup>. Pour Suñer (1983) et Siewierska (2008), le pluriel serait à interpréter comme une marque de référence arbitraire ou comme la grammaticalisation d'une forme d'impersonnel. Selon Cabredo Hofherr (2003, 2014), le pluriel serait responsable des traits [+comptable] et [+humain] dans l'usage des pronoms arbitraires. Notre approche, à l'inverse, se distingue de celles-ci par le fait qu'elle révisé la conception traditionnelle du pluriel, dont la marque peut dès lors s'interpréter soit comme une somme, soit un ensemble, auquel cas le nombre d'éléments peut éventuellement équivaloir à 1.

Cette interprétation du pluriel vaut pour les exemples ci-dessous également. Le suivant comporte un discours rapporté portant sur un examen de conduite :

(415) [les interlocuteurs viennent d'évoquer la récente obtention du permis de conduire par une connaissance mutuelle] avec qui que je discutais l'autre jour ah | # | # | qui me disait qu'y a un jeune qui était sûr de passer son | \_ | pis qui qui roulait bien | \_ | mais trop trop sûr de lui pis **ils lui ont / pas donné** parce que | \_ | justement il était trop sûr de lui dans les ronds-points | il y allait trop | \_ | trop | % | \_ | si tu fais pas de bêtises **ils te le donnent hein** | \_ | (ofrom)

Au vu du scénario sous-jacent, on peut assez plausiblement inférer l'identité de l'agent du procès 'donner', tout comme, d'ailleurs, celui de son objet Ø, qu'on présume coréférentiel au SN précédent *son* Ø. A noter que le référent est dans un premier temps évoqué dans un

---

<sup>293</sup> Or, selon Corbett (2000 : 17), un système où le pluriel servirait de nombre non marqué n'est pas attesté à travers les langues, même si certaines langues n'obligent pas à sélectionner le trait de nombre pour la catégorie nominale, en recourant à une forme « générale », qui n'engage pas le locuteur à ce niveau (p. 9 *sqq.*). Par exemple, dans une langue couchitique appelée le bayso, la forme générale *lúban* ('lion(s)') est utilisée dans les cas où le nombre est non pertinent, ne signalant pas s'il est question d'un ou plusieurs animaux de ce type, par opposition au singulier *lúban-titi* ou au pluriel *lúban-jool*. Cette situation est également valable pour le fula, un dialecte de Guinée ainsi que pour l'arabe syrien, à la différence près que la forme générale n'est pas disponible pour tous les types de noms (p. 12-13). Dans d'autres langues, plus nombreuses, une interprétation « générale », *i.e.* qui n'engage pas quant au nombre d'individus, peut se faire au moyen de l'une des formes existantes du système du nombre, en l'occurrence, le singulier. C'est le cas du japonais où le singulier peut indiquer un nombre non spécifié. Ainsi, *inu* (sg.) pourra signifier 'un ou plusieurs chien(s)' tandis que *inu-tati* (pl.) indiquera clairement qu'il est question de plusieurs chiens. Par contre, Corbett ne relève pas le pendant de cette situation, *i.e.* une langue où le pluriel est le cas non marqué.

contexte épisodique (ici via le passé-composé), puis reconvoqué pour une prédication générique.

Ci-dessous, l'extrait illustre le concours d'indices de diverses natures pour l'instanciation de la variable exprimée par *ils* :

- (416) Voici comment Fred a trouvé notre voiture hier après l'opéra...  
[photo d'une voiture dont une vitre est cassée]  
Shit happens! *Ils ont volé un appareil photo* et qq trucs du boulot, notamment des clés #<sup>294</sup> (vachement délicat)! Mais par contre, le GPS qui était collé au pare-brise, mes lunettes à soleil, etc, *ils n'y ont pas touché*... (courriel, 30.01.2012)

La photographie évoque un scénario probable dont on peut inférer le type d'agent. La prédication attribuée à *ils* vient confirmer l'hypothèse interprétative avec laquelle on peut dès lors instancier la variable. Ici aussi, il est possible que le nombre de membres équivale à 1.

L'exemple suivant comporte lui aussi une dimension visuelle, car il s'agit d'un spot publicitaire de prévention des accidents, qui met en scène une statuette brisée reprenant progressivement la silhouette d'un ouvrier :

- (417) L'accident est arrivé comme un éclair<sup>F295</sup> je me suis plus ou moins cassé tout ce qu'on peut casser<sup>F</sup> j'ai eu la trouille quand *ils m'ont dit que je devais arrêter de travailler* pendant cinq mois au moins<sup>F</sup> au boulot *ils ont été drôlement embêtés<sup>S</sup> et ils ont trouvé quelqu'un* pour faire mon travail<sup>F</sup> après la quatrième opération j'étais complètement déprimé<sup>S</sup> c'est seulement ensuite que j'ai pu commencer la réadaptation<sup>F</sup> (TV, spot publicitaire pour une assurance-accidents, intitulé « Attention, fragile »)

A partir du scénario mis en scène et du discours rapporté (un ordre d'arrêt de travail), il est assez aisé d'inférer une classe correspondant à la première occurrence de *ils*. Là aussi, le nombre de membres de l'ensemble est potentiellement égal à 1. On remarque que les occurrences suivantes de *ils* ne prolongent pas cette référence. Elles se trouvent sous la portée d'un circonstant cadratif (*au boulot*) permettant (de même que les prédicats formulés) de restreindre le domaine d'inférence de la nouvelle classe.

Ci-dessous, le réseau de *ils* renvoie invariablement à une même classe, dont l'identité des membres n'est pas pertinente :

- (418) [La locutrice donne son avis sur différents films] y avait aussi le maître du jeu je sais pas si tu | \_ | euh le maître du jeu ou le | \_ | comment ça s'a/ y a une autre euh y a une autre euh y a un autre titre euh | \_ | c'est vraiment un film euh | \_ | le film est vraiment fantastique mais *ils ont complètement changé le le le | \_ | l'histoire* par rapport au bouquin | \_ | mais les deux sont juste incroyables [...] j'aime ces films où quand tu sors de de d/ par exemple du cinéma | \_ | où les quand tu vois les gens tu les croises pis tu

---

<sup>294</sup> Nous anonymisons.

<sup>295</sup> Pour rappel, l'exposant S signale une intonation continuative et F une intonation conclusive.

discutes du film | \_ | qu'on soit dix pis y a dix v/ dix façons d'avoir vécu le film différemment | \_ | et pis de façons de penser euh à ce qu'il s'est passé différentes | \_ | le ben par exemple Matrix le le le la toute première euh le tout premier | \_ | euh | \_ | donc Matrix un | \_ | c'était j'ai trouvé ça génial | \_ | parce que parce que ça te laissait | \_ | ça te laissait en suspens *ils te donnent pas* euh | \_ | de de *ils te donnent pas tout en main* euh | \_ | voilà t'as ta propre imagination c'est cool ça [...] le Da Vinci code pour moi | \_ | *ils auraient pu arrêter les vingt dernières minutes* | \_ | ce film il était juste | \_ | le donc le premier le deuxième c'était nul à chier | \_ | c'était trop visible directement où euh | \_ | le résultat était trop visible | \_ | et le Da Vinci code le un | \_ | *ils auraient coupé les vingt dernières minutes* ce ce film serait été extraordinaire | \_ | pourquoi *ils ont besoin à la fin de dire* | \_ | voilà alors elle c'est la fille de parce que nin nin nin et pis il lui est arrivé ça et pis | \_ | je veux dire | \_ | *ils ils ont* | \_ | % bout d'un moment *ils / \_ / coupent toute euh / \_ / toute imagination* | \_ | pis je dans ce film c'est dommage quoi | \_ | tout le long *ils arrivent à à faire un suspense / \_ / incroyable* | \_ | pis ces vingt dernières minutes ça a ça a tué tout le tout le film (ofrom)

La locutrice use massivement du *ils* pour impliquer une classe dont on infère le rôle à travers les prédicats utilisés et le contexte cinématographique. Après la première occurrence, le rôle est donc valide pour toutes les occurrences ultérieures (cf. l'exemple (410) où il change inopinément), sans qu'il soit nécessaire d'inférer l'identité spécifique, dont on sait qu'elle varie en fonction des différents films évoqués.

Pour tous les exemples de cette catégorie, la lecture semble par défaut collective et permet de diffuser la responsabilité d'un éventuel unique exécuteur sur l'ensemble.

#### 3.2.1.4 Bilan

Les sous-catégories présentées ci-dessus nous semblent pouvoir être avantageusement décrites en termes d'anaphore indirecte. En effet, les extraits mettent en jeu un individu collectif, un lieu ou alors impliquent la reconnaissance d'un scénario, sur lequel peut se fonder l'inférence du référent présupposé par *ils*. S'il est vrai que les membres de la classe évoquée ne sont pas individuellement identifiables (d'où l'effet de sous-détermination), il n'en reste pas moins que l'on peut inférer des propriétés typantes et distinctives de la classe. Bien que les prédicats attribués à une classe issue d'un individu collectif (§3.2.1.1) ou aux occupants d'un lieu puissent être aussi bien collectifs que distributifs, l'inférence d'un *actant*, éventuellement via un lieu (§3.2.1.2), à partir d'un scénario impliqué (§3.2.1.3) favorise une lecture collective, avec un rendement d'expansion de la responsabilité sur l'ensemble. Dans ce cas, une commutation en *les gens* fonctionne mal, car elle dilue le rôle dès lors endossé par le référent.

D'autres indices montrent que *ils* n'est pas ici un simple élément postiche, mais qu'il renvoie à un référent doté d'un rôle à part entière. En effet, il se voit affecté d'attitudes, d'émotions, d'intentions, à travers l'occurrence d'auxiliaires modaux (*devoir, vouloir*) ou exprimant l'état mental ou physique d'un « expérimentateur » (*croire, avoir faim, aimer, avoir peur, souffrir, hésiter, avoir besoin*). On peut également invoquer la possibilité de prolonger la référence sur

un empan discursif assez large (e.g. (409), (414), (416), (418)), à travers de nombreuses occurrences de *ils* qui contribuent à en renforcer la saillance.

### 3.2.2 Anaphore indirecte ou variable non instanciée ?

Il n'est pas rare que l'on puisse hésiter sur l'analyse à imputer au pointeur *ils*. Certains indices permettent en effet d'envisager un procédé d'inférence d'un objet pour *ils*, à l'instar des précédents, mais d'autres orientent vers une mise en retrait du rôle de l'agent dans l'expression d'un procès, interrogeant le fondement d'une résolution référentielle.

#### 3.2.2.1 Ambiguïté d'analyse

L'exemple suivant illustre une analyse ambiguë, qui se traduit par la possibilité d'argumenter pour ou contre un maintien de la fonction instructionnelle d'instanciation de la variable :

- (419) tous les barèmes pour les pratiques sont basés sur euh les meilleurs é- étudiants euh des universités en sport donc tout toutes les meilleures performances euh y euh de chaque euh canton ou université de Suisse forment la moyenne donc ***ils basent les barèmes*** d'après ça donc euh ce qu'il faut savoir c'est que bien souvent euh c'est des barèmes assez élevés (ofrom)

La première analyse consiste à considérer *ils* comme inférable, au même titre que les exemples de la catégorie précédente, en termes d'anaphore indirecte : le contexte académique et le prédicat typant restreignent sensiblement l'extension du pointeur, bien que l'inférence, comme toujours, puisse demeurer incertaine {*les professeurs, les experts, les membres d'une commission, etc.*}.

Mais une autre analyse consiste à renoncer à un traitement inférentiel et à maintenir la variable introduite par *ils* au rang de variable non instanciée. Dans ce cas, *ils* servirait à remplir une place d'agent dont l'implication dans le procès se verrait réduite<sup>296</sup>, d'où la mise en évidence du procès lui-même (et éventuellement de son objet patient ou effectué) (Myhill 1997, Siewierska 2008). Dans l'exemple en question, le locuteur chercherait à mettre en avant les modalités de conception des barèmes plutôt qu'à prédiquer cette manière de faire sur un référent établi (les professeurs ou autres agents potentiels). La diathèse passive utilisée au début de l'extrait (*tous les barèmes pour les pratiques sont basés sur...*), n'obligeant pas à attribuer une valeur référentielle pour l'agent impliqué, tend à renforcer cette hypothèse. Nous examinerons les différences entre les structures actives vs passives *infra* (cf. §3.4.1).

Toute une série de données peuvent être analysées de la sorte :

- (420) on se retrouve quand même dans des endroits, franchement/ J'ai, j'ai une c/ j'ai une amie là qui a eu son concours l'année dernière qui, qui est instit. Donc au début ***ils, ils***

---

<sup>296</sup> Cf. les notions de « baisse d'agentivité » (Maillard 1994), d'« agent defocusing » (Myhill 1997) ou d'« agent demotion » (Siewierska 2008).

*L'ont mise, à Meylan* et tout, donc elle était super contente elle s'est dit 'voilà je vais me retrouver avec des petits bourges et tout, il y a pas de problème'. (pfc)

D'un côté, il y a moyen d'inférer une classe responsable de l'affectation en question à la faveur d'un scénario aisément reconstituable. D'autre part, on peut s'en tenir à l'indication d'affectation concernant l'individu, la participation d'un agent restant en marge.

Dans l'exemple suivant, c'est l'introduction d'un lieu qui permet d'invoquer une analyse en termes d'inférence pour le *ils*, désignant les occupants du lieu. A noter que la mention du lieu remonte à plusieurs minutes et tours de parole en amont. Néanmoins, le lieu sert de cadre référentiel à tout l'épisode narratif :

(421) [à propos d'un voyage au Maroc]

G: Eh ben ma foi, tss le Maroc pour moi c'était bien. Si tu veux savoir comment que je l'ai, euh, je l'ai organisé, c'est une chose. [...]

E2: Et alors la nourriture, vous avez mangé des choses euh spécifiques?

G: Ah ben pour la nourriture euh, c'est tout euh, enfin c'est la tajine qu'**ils appellent**. C'est euh, c'est un, un espèce de euh, enfin tu sais ce que c'est qu'une tajine quoi, c'est un plat en terre. Un grand plat en terre avec un espèce de couvercle en cône. Et puis qu'ils mettent là, ils font cuire là-dedans comme ça. Alors tu manges la tajine aux oeufs, tu manges la tajine au mouton, tu manges la ta/, enfin **ils appellent** ça la tajine. (pfc)

Mais d'un autre côté, il paraît aussi envisageable de considérer l'acte de dénomination de l'objet indépendamment de la source effective (cf. *ça s'appelle la tajine* ou *on appelle ça la tajine*, voir *infra* §3.4), au vu de la routinisation de la combinaison de *ils* avec des verbes de parole à fonction évidentielle (cf. *infra* §3.2.3.2).

Ci-dessous, le scénario impliqué par l'évocation d'un chantier permet d'inférer une classe qui indifférencie plusieurs niveaux d'actants potentiels (exécuteurs, instigateurs) :

L1 : euh + y avait + les bâtiments qui étaient sur le quai qui existaient + y avait encore la halle aux v- + nous on a connu la halle aux vins encore donc les derniers les derniers récalcitrants quand même hein euh + la la majeure partie + était déjà partie mais il restait quelques-uns euh + autre XX c'était un chantier encore hein c'était vraiment

L2 : **ils démolissaient**

L1 : **ils démolissaient** on a euh + on a fait nos études au moment de la démolition + moi je trouve que le + par rapport à ce qu'on a connu le quartier a pas trop trop changé (cfpp)

Quant à l'analyse concurrente, elle voit dans la structure *ils+V* un moyen de mettre l'accent sur le procès de démolition et son aspect progressif (via l'imparfait, cf. *infra* §3.4.1). Par la suite, le procès se voit d'ailleurs reformulé via une nominalisation ad hoc, qui permet également de mettre en retrait le rôle d'agent.

Les exemples en vrac ci-dessous illustrent la même ambiguïté d'analyse :

- (422) on habitait à la rue ici tu vois près de l'hospital à côté du marché de la Boqueria c'était vachement bien c'était un des plus grands marchés d'Espagne | % | \_ | mais non parce que c'est un marché où *ils vendent que de la nourriture* | \_ | pis du coup mais ouais mais euh on est allés acheter des jus de fruits et tout (ofrom)
- (423) Tout est provisoire et tout s'achète, sauf Octave. Car je me suis racheté ici, dans ma prison pourrie. *Ils m'ont autorisé* (contre menue monnaie) à regarder la télé dans ma cellule (Beigbeder, 99F, p. 269)
- (424) [Trois interlocuteurs discutent à propos d'un voisin habitant manifestement en face du lieu où ils se trouvent]  
AB1: C'est allumé ou pas ? Il y est ? Il y est le gars ?  
E: Pourquoi ? Il y a qui ?  
AB1: Il y a un mec très intéressant, qui va partir euh aux Etats-Unis pour faire reporter-photo. Inter/, international. Parce qu'il parle en fait, il parle arabe, il parle anglais, il parle français. *Ils l'ont engagé<sup>297</sup> parce qu'il parle arabe. Donc ils vont l'envoyer dans les pays arabes.* (pfc)
- (425) On mange chacun un sandwich généralement on l'achète sur le bateau parce *qu'ils nous en font des tout frais*,> sur place on le commande puis *ils le font i/ ils font* (pfc)

On remarque donc que l'ambiguïté se produit à chaque fois que le contexte délivre des indices suffisants pour une inférence plausible du référent (cadre référentiel, prédicat typant, scénario impliqué, pointages successifs, etc.), mais que l'on juge que la pertinence communicative porte sur le procès ou son résultat.

### 3.2.2.2 *Quiproquos interprétatifs*

En fait, l'existence de deux scénarios interprétatifs concurrents se reflète à travers certains échanges qui illustrent le décalage de traitement potentiel entre les interlocuteurs. Il arrive ainsi que l'interprète se lance dans la quête du référent sans toutefois parvenir à une hypothèse suffisamment fiable. Plutôt que de laisser la variable non instanciée, il peut solliciter l'aide de son interlocuteur :

- (426) [discussion autour du thème de la « fête des voisins »]  
L1 : enfin nous on a fait on a fait on a ça aussi à Sevran dans le quartier où c'est très très grand hein euh *ils mettent les barbecues* (mm mm) et tout ça euh c'est bien organisé + avec une fête avant pour les enfants + on *ils + ils ramènent des poneys* tu vois *ils* <ouais ah oui> *ils font faire des tours* XX  
L2 : **quand <d'accord> tu dis «ils» <ils> c'est qui «ils» c'est pas les habitants qui**

---

<sup>297</sup> Après plusieurs écoutes attentives, nous avons corrigé une transcription à notre avis erronée (*il est engagé*). Mais les deux versions ne se distinguant finalement que sur un phonème vocalique, il est vrai que le doute subsiste.

**s'organisent <si c'est les> c'est d'autres ?**

**L1: hab- non c'est les habitants du quartier** (cfpp)

La locutrice L1 use d'abord d'un indice de personne [- délocuté] (ou [+inclusif]) (*nous on a fait on a fait on a ça aussi*), puis passe, après hésitation (*on ils*), à un indice [+ délocuté] de 6<sup>e</sup> personne (*ils ramènent des poneys*). C'est peut-être ce changement qui pousse L2 à s'interroger sur l'extension de ce dernier (cf. *quand tu dis « ils » c'est qui « ils » c'est pas les habitants qui s'organisent c'est d'autres ?*). En fait, le choix opéré par L1 pour un pronom [+ délocuté] traduit manifestement l'implication d'exécuteurs tiers pour les animations spécifiques énumérées, tandis que l'instigation de l'événement dans sa globalité émane des habitants dont fait partie L1, d'où l'emploi de *nous*. Quoi qu'il en soit, cet extrait illustre l'usage d'un pointeur vraisemblablement « démotivé » pour L1, dont l'enjeu est de mettre en avant les différentes activités proposées, tandis que L2 cherche à instancier une valeur pour l'agent impliqué.

Il en va de même dans l'extrait de conversation suivant :

(427) E: Tu sais qu'on n'a plus accès au parc ?

BL: Je sais bien oui.

E: C'est triste hein quand même.

BL: <mais c'est comme> ici tu ne peux plus aller au domaine de l'O.N.E. non plus hein.

E: plus du tout ?

BL: non *ils ont mis une barrière* en bois enfin avec un cadenas tu ne sais plus rentrer non plus c'est malheureux.

E: **et qui est-ce qui a mis ça ?**

BL: **ben le propriétaire** tu ne sais pas qui c'est ? tu es pas au courant de tout ça ?

E: ça été revendu ?, ça a été acheté ?

BL: c'est acheté et tout maintenant hein, ah oui c'est un des, des propriétaires d'in/d'interbrew, un des fils je pense.

E: Ah oui.

BL: Un gros manitou. Et euh, il, il fait une villa privée. (pfc)

Cet exemple témoigne de la possibilité de taire l'identité du référent en toute connaissance de cause. On peut invoquer ici le rendement de dilution de l'agentivité par rapport à un scénario éventuellement complexe impliquant différents niveaux d'agents (un instigateur qui fait faire quelque chose par l'intermédiaire d'un/d'exécuteur(s)). L'usage du pointeur *ils* permet de styliser le procès pour mettre en avant son résultat, l'enjeu étant pour le locuteur de justifier sa déception. Mais cela n'empêche pas l'interlocuteur E de vouloir connaître les tenants et aboutissants de la situation – y compris sa source évoquée via *ils*– dont il n'est visiblement pas au courant.

L'extrait suivant met en jeu non pas une demande d'identification, mais une demande de confirmation sur l'interprétation de *ils* :

(428) [Il est question d'une baignade dans une rivière au courant rapide]

L1 : tu fais comment pour sortir<sup>Q</sup>

L2 : Ben *ils te disent* quand il faut sortir<sup>F</sup>

L1 : Ah mais **y a des gens qui surveillent**<sup>Q</sup>

L2 : Non mais **c'est écrit sur des panneaux**<sup>A</sup> (oral, au vol<sup>298</sup>)

A partir de la construction en *ils* utilisées par L2, L1 infère l'existence d'une classe d'individus dont le rôle est de surveiller les nageurs. Cette réaction oblige L2 à rectifier l'origine du procès en le détaillant. La mise au point de L2 confirme bien le processus de confusion référentielle opéré à travers *ils*, qui condense indistinctement un ensemble de circonstances impliquant potentiellement un *instigateur* [+humain] (d'où l'emploi de *ils disent*) et un exécuter [- humain] (les 'panneaux d'avertissement'), ce dernier se voyant complètement dilué dans la source exprimée (*ils*). L'objectif de L1, à travers cette construction en *ils*, est probablement d'exprimer l'existence d'un système de signalisation à l'intention des baigneurs, indépendamment de la source du procès. Cet exemple reflète bien la manière dont les locuteurs, pour des questions de pertinence, manient le degré de granularité (Sturt *et al.* 2004) de leur représentation des faits du monde, avec le risque toutefois que la sous-détermination soit mal interprétée.

Dans l'exemple ci-dessous, l'emploi de *ils* est non seulement motivé par l'imputation d'une responsabilité collective et indifférenciée, mais aussi par une stratégie de cryptage à des fins de sauvegarde de la face :

(429) [la police est en négociation téléphonique avec un preneur d'otage qui soupçonne un complot de l'Etat suite au décès de sa compagne. Le dialogue met en scène le preneur d'otages et une psychologue de la police]

- J'étais psychologue. Plutôt bon, je dois dire. J'avais mon propre cabinet et, financièrement, je m'en sortais très bien. Mais ce n'était pas là l'essentiel. Pour moi, il ne s'agissait pas d'un boulot ordinaire, c'était une véritable vocation. Mon travail, c'était toute ma vie, et *ils me l'ont pris*.

- **Qui ça, ils ?**

- Au début, quand j'ai commencé à poser des questions, mes interlocuteurs ont d'abord été très obligeants. [...] J'ai reçu la visite de deux messieurs, probablement envoyés par les services secrets [...]. Pour commencer, lors d'un contrôle routier, la police a trouvé de la cocaïne dans mon coffre. Sans même écouter mes dénégations, l'ordre des psychologues m'a retiré mon autorisation d'exercer [...] (Sebastian Fitzek, *Ne les crois pas*, ch. 7)

Le preneur d'otages se plaint, via la structure en *ils+V*, de la perte de son travail. Ce faisant, il cherche visiblement à accuser les responsables de la situation. En passant sous silence leur identité, il tente de se prémunir contre d'éventuelles menaces pour sa face. En même temps, sa réponse à l'interrogation de l'interlocutrice met en évidence la complexité des

---

<sup>298</sup> Nous avons tenté de reconstituer les propos en restant au plus près des faits.

circonstances et la diversité des actants impliqués dans la perte de son emploi (police, service secret, ordre des psychologues, etc.). Le titre du roman (*Ne les crois pas*), afin de susciter la curiosité du lecteur, consiste d'ailleurs principalement à entretenir le mystère là autour...<sup>299</sup>

L'extrait ci-dessous illustre encore la coexistence de deux analyses possibles, conscientisée paradoxalement par un même locuteur, à travers une paire question/réponse auto-locutée :

(430) Elle [la mère du narrateur] avait eu l'habitude d'exercer des responsabilités, de prendre des décisions, de dépenser de l'argent sans trop compter, elle était coquette et cultivée et voilà à quoi on l'avait réduite, quarante kilos de chair meurtrie et d'esprit ravagé qui cherchaient encore à paraître quelque chose avec ses lunettes sur le nez et des romans sur sa table de nuit. Un vieux machin délaissé emballé dans une blouse à fleur, un cadavre en décomposition qui essayait de se faire passer pour vivant. C'est ça qu'*ils avaient anéanti*, avec *leur* rapport lapidaire : l'humanité de ma mère. **Qui ça ils ?** Je ne le savais pas. La maison de retraite ? L'association qui gérait sa curatelle ? Le système, quoi. Ou bien alors des individus, curateur, infirmière, médecin qui se retranchaient derrière un système qui déshumanisait les cas sur lesquels ils travaillaient. (Yves Aubrymore, *Qui a tué Frajdla Cinnamone ?*, p. 51)

Cet exemple à caractère dialogique représente le pendant de (427), où un locuteur recourt à *ils* sans réellement avoir de valeur référentielle en tête. La question *Qui ça ils ?* souligne le paradoxe auquel se trouve confronté le sujet, entre l'instruction d'instancier la variable, et l'échec de la procédure, confirmé par l'aveu d'ignorance (*Je ne le savais pas*) et par la diversité des hypothèses formulées. Ici aussi, *ils* permet d'évoquer le caractère multiple et dilué des responsabilités en jeu.

### 3.2.3 Variable non instanciée

Cette section comprend les cas où les coûts cognitifs s'avèrent manifestement trop élevés pour que l'on puisse prétendre à un traitement inférentiel de la variable<sup>300</sup>, bien qu'en théorie, l'hypothèse ne soit jamais complètement réfutable. Nous proposons donc de mettre en avant les indices qui orientent vers l'abandon d'un traitement inférentiel de la variable.

#### 3.2.3.1 Promotion du procès

Dans certains cas, la mise en retrait du rôle d'agent au profit de la promotion du procès ou de son résultat est manifeste au regard des objectifs communicationnels du locuteur. L'exemple

---

<sup>299</sup> Cf. aussi le film intitulé « Ils », auquel nous a rendue attentive C. Schnedecker (c. p.), dont voici le synopsis : « Lucas et Clémentine, un couple trentenaire expatrié en Roumanie, habite [*sic*] depuis peu une maison isolée en banlieue de Bucarest. Elle, professeur de Français, lui, romancier, vivent un bonheur paisible... Pourtant, un soir, dans leur maison, tout va basculer... La pluie battante fait rage à l'extérieur... Le téléphone retentit, des voix lointaines au bout du fil... incompréhensibles. Le couple n'est pas seul... Le cauchemar commence... ILS sont là. » (<http://www.allocine.fr>).

<sup>300</sup> Cette catégorie recouvre en grande partie les emplois appelés (*quasi*-)existentiels par d'autres (Cinque 1988, Cabredo Hofherr (2003, 2014) cf. *supra* §2.4.2. Cependant, comme nous n'adhérons pas à une analyse en termes de quantification existentielle et qu'à cet égard, quelques cas commentés ici s'en distinguent sensiblement, nous renonçons à emprunter le terme.

suisvant montre que l'état de choses dénoté par la construction en *ils* se dote d'une fonction argumentative dans l'orientation du programme discursif en cours :

(431) Je suis parti au milieu de la leçon de géographie parce que j'avais mal à la tête. La dame m'a donné un cachet, j'avais toujours mal, alors elle m'a fait étendre sur un lit. On m'a oublié tellement j'ai dormi après. Quand je me suis réveillé, il faisait nuit dans l'école et les autres, ils avaient déjà fini. J'ai sauté par-dessus la grille d'entrée. Dehors, *ils faisaient la grève du bus et du métro*, et comme j'avais pas d'argent pour téléphoner à ma mère que j'étais en retard, alors je suis allé à pied à la gare de Lyon qui n'est pas loin et j'ai pris le train pour Marseille qui passe par chez ma grand-mère, à Montélimar. (Yves Pagès, *Les Gauchers < Envol lycée*)

On peut noter dans cet extrait quelques traces de contrefaçon d'oral, telles que la dislocation (*les autres, ils...*) ou l'omission du *ne* (*j'avais pas d'argent...*). La grève en vigueur exprimée par la construction en *ils* constitue l'une des causes de la conclusion 'je suis allé à pied' : la situation de grève, plutôt que ses agents, évoque par ailleurs une conséquence implicite déterminante – à savoir l'absence de transports publics opérationnels – pour comprendre le raisonnement argumentatif.

Il en va de même dans la conversation en ligne ci-dessous, où L1, de retour de vacances, exprime son découragement à reprendre le travail, auquel s'ajoute un désagrément supplémentaire :

(432) L1 : Retour au bureau. Bonjour tristesse ! (statut)  
L2 : Pauvre # (commentaire du statut de L1)  
L3 : rhoohhhhh ça promet pour mon retour (commentaire, L3 travaille avec L1)  
L1 : *i z ont coupé l'eau* ce matin de 9.30 à 12.00 (commentaire) (réseau social, écrit depuis le lieu de travail, 26.02.2013)

Malgré le prédicat tyant, susceptible d'orienter vers un traitement plus abouti, l'interprétation ne va manifestement pas au-delà de l'expression du procès et son résultat (l'interruption d'approvisionnement en eau), dont l'enjeu est de contribuer à imaginer l'humeur du locuteur.

Dans le SMS suivant, l'objectif de l'auteur est de renseigner son destinataire sur la situation (une suppression de train) et sa conséquence (une arrivée probablement retardée) :

(433) *Ils ont supprimé mon train*. J'arrive à sonceboz vers 14h05. (Swiss sms corpus)

L'emploi de *ils* semble avantageusement exploité pour mettre en relief le mouvement argumentatif de cause à effet. C'est aussi le cas ci-dessous :

(434) [le locuteur raconte son voyage au Maroc]  
C'était un grand euh, un grand bâtiment, une grande terrasse, si tu veux, qui était couverte

[s'ensuit un épisode racontant les dérangements intestinaux dont souffrait le locuteur]  
Il y avait une douche, je me rappelle, et puis il y avait quand même le W.C. Mais le plus drôle, c'est que comme c'était un compre/, un euh, un compresseur, un euh.

D : Groupe.

G : Un groupe oui, je dis un compresseur, mais c'est un groupe. Un groupe qui faisait électricité. A minuit, quand *ils ont coupé*, il y avait plus d'électricité. Alors dans la chambre pour se lever, ben à tâtons quoi (pfc)

L'«interruption de courant» exprimée via la construction verbale et son résultat («l'absence d'électricité») causent l'embarras du locuteur sujet aux va-et-vient nocturnes. Dans tous ces extraits, où le procès évoqué intervient comme une simple étape d'un raisonnement argumentatif, il paraîtrait incongru de demander des clarifications sur l'identité de l'agent impliqué. On remarque aussi qu'une commutation avec *les gens* induirait une lecture binaire sujet-prédicat, que l'emploi de *ils* tend à éclipser.

On relève également le même genre d'enchaînement argumentatif dans les exemples ci-dessous, où c'est le procès en soi plutôt que la prédication sur des individus qui justifient l'appréciation du locuteur :

(435) et puis le je jour avant on | \_ | on est allé à Québec il y avait le le concert des Pink Floyd | \_ | donc *ils étaient en train de démonter la scène* donc on n'a | \_ | c'est vrai que c'était pas c'était pas très joli il y avait un peu des grues des échafaudages voilà (ofrom)

(436) on est allé jusqu'aux chutes du Niagara | \_ | puis là on a pu les voir euh de jour | \_ | et également de nuit puis c'est vrai que c'était juste euh | \_ | magique | \_ | de nuit c'était on avait l'impression de voir euh | \_ | *ils projettent des lumières* sur ces chutes c'est juste magique (ofrom)

(437) [la conversation porte sur la pratique du vélo en ville de Paris] c' qui est sympathique c'est c'est le dimanche le long des + le long d' la Seine là + une ou deux fois elle l'a fait avec nous oui quand *ils ferment les quais* + là c'est sympa + d'être au milieu des rollers (cfpp)

En d'autres termes, ce sont respectivement le «démontage de la scène», «la projection des lumières» et «la fermeture des quais» qui motivent les évaluations exprimées. Dans l'exemple suivant, la structure en *ils* entre dans un mouvement cette fois concessif :

(438) bah le troisième c'est j'imagine que ça doit être très sympa d'y vivre aujourd'hui euh ça doit être très calme très tranquille euh même *s'ils brûlent les scooters* là depuis peu depuis peu et tout ça (cfpp)

La construction en *ils* représente un argument inefficace contre la conclusion qu'on peut gloser «il fait bon vivre dans le quartier». La pertinence de l'argument, à nouveau, réside dans

le constat des incidents qui s’y produisent plutôt que dans la reconnaissance d’individus responsables des incidents en question.

Il peut être intéressant d’observer la portée de la négation dans les énoncés de ce type :

(439) [CR est en train de montrer des photos]

CR : Ca c'est (X) et (X) à New-York il faisait froid on était coincé dans la neige on a été pris dans une tempête de neige on a pas pu rentrer.

E : C'est vrai ? Ben j'imagine au mois de février à New-York c'est c'est le plein hiver

CR : Donc il y avait person/ du coup **ils avaient pas déblayé la rue** ça a été épouvantable on a, on a é/ on est/ on est parti avec une journée de retard. (pfc)

(440) ouais c'était vraiment réputé pour être des pistes difficiles | \_ | je sais pas si toi tu as eu mais peut-être plus jeune quand-même hein | \_ / **ils ils travaillaient pas les pistes** encore euh c'était vraiment des des champs de bosses tu vois | \_ | puis c'est bien c'était bien raide | \_ | c'était réputé pour ne pas être | \_ | idéal pour des débutants (ofrom) = (4)

Un aspect remarquable est que la négation ne semble pas affecter la relation entre une constante (la classe d’individus c) et son prédicat, comme schématisé ainsi : Neg[déblayer (c, r) ou Neg[travailler (c, p)]. Mais elle semble plutôt porter sur le procès subi par le patient, à la manière d’une structure au passif sans agent exprimé : Neg[déblayé(r)] ou Neg[travaillé(p)]. Autrement dit, les locuteurs respectifs nient que la rue ait été déblayée et que les pistes aient été travaillées pour en promouvoir le résultat et en tirer les conclusions respectives. Nous reviendrons sur la question du passif *infra* (§3.4.1).

### 3.2.3.2 *Verba dicendi*

On trouve régulièrement en français, dans les genres de parole « improvisée » en tout cas, des constructions en *ils* avec un verbe de parole. A noter que cela va à l’encontre des observations de Siewierska (2008), pour qui cet emploi « évidentiel » de la 6<sup>e</sup> personne se verrait restreint, sauf en anglais familier, à des genres discursifs particuliers, comme les mythes, fables ou proverbes. A ses yeux, les langues européennes privilégient l’usage du passif sans agent ou d’autres formes de construction à agent non spécifique, par exemple l’usage du *on* en français ou du *man* en allemand. Les exemples ci-dessous remettent donc en cause cette prédiction. Précisons que ce n’est pas l’indice de personne en soi, mais l’ensemble de la construction qui fonctionne comme une marque d’évidentialité :

(441) c'est sûr que y a un décalage parce que je veux dire euh | \_ | on on on a une autre dynamique scolaire | que si tu es dans un tout petit village euh | \_ | où tu as que des enfants de la région où tu as des enfants qui sont euh | \_ | suivis par les parents | \_ | euh | \_ | après | \_ | si tu prends Genève où de nouveau y a une autre dynamique euh | \_ | ça peut être très différent euh | / ces résultats scolaires | \_ | je pense | \_ | qu'est-ce qui fait que | \_ | qu'est-ce/ je sais pas pourquoi euh | \_ | c'est un peu **c'est vrai qu'ils disent que / \_ / que Fribourg / % / le Valais en général on a assez de bons résultats** au niveau de ces tests Pisa (ofrom)

- (442) [la locutrice répond à la question de l'enquêtrice sur ses projets de séjour à l'étranger après ses études] d'abord je pense | \_ | faire en Suisse | \_ | pour f/ avoir l'expérience | \_ | avoir des bases | \_ | pis souvent p/ pour partir **ils disent que c'est mieux de d'avoir travaillé deux ans** | dans un endroit | \_ | pour être bien formé | \_ | pis après ben on verra ce que l'avenir nous réserve on sait | on sait jamais mais j'aimerais bien voyager ouais (ofrom)
- (443) [au sujet d'un voyage au Népal] la plupart du temps j'étais dans la vallée de Katmadou | \_ | qui est euh | \_ | oui qui est | \_ | très touristique c'est clair | \_ | j'ai jamais fait un voyage | \_ | vraiment | \_ | très très loin dans les euh | \_ | dans les euh lieux inconnus sauf une fois | \_ | oui lorsque j'étais dans le | \_ | **comme ils disent le Far West** | \_ | @ | oui oui oui oui oui oui | \_ | c'était la seule fois (ofrom)

Dans ces extraits, l'enjeu consiste à signaler que l'information est de seconde main, qu'il s'agisse d'étayer un argument, de se distancier des propos, de les présenter comme incertains, de les rattacher à un jargon, etc. La construction verbale a donc une fonction clairement mitigatrice. On a déjà évoqué *supra* (Ch IV §4.5.3) le fait que tout discours rapporté est par nature contrefait (à l'exception peut-être des textes scientifiques où la doxa exige en principe l'exhaustivité en la matière) et cela est particulièrement manifeste dans les genres de parole spontanée, comme ci-dessus, où il n'y a pas d'exigence, outre de pertinence relative, ni sur la reproduction des propos ni sur leur source. Il est évident que le contexte et le prédicat hôte restreignent les possibilités d'interprétation de l'« autorité » invoquée et qu'il est possible de faire des conjectures plus ou moins fiables à son propos. Mais la récurrence du procédé oriente plutôt vers l'hypothèse d'une lexicalisation d'un oui-dire. La comparaison avec la locution *on dit* sera examinée *infra* (§3.4.2).

Il en va un peu de même dans un contexte particulier, celui des prévisions météorologiques, qui met régulièrement en jeu l'occurrence de *ils* avec le verbe *annoncer*, dans divers genres de l'« immédiat » communicatif (Koch & Oesterreicher 1985).

- (444) à part ça ma cousine elle a toujours pas téléphoné pour le pique-nique | alors je suppose qu'il aura pas lieu | \_ | étant donné qu'**ils annoncent du froid** | \_ | **pis** | **du mauvais** (ofrom)
- (445) Ils vont avoir de la neige parce qu'**ils ont annoncé du mauvais temps** là non non huit jours je sais pas quinze jours non ? (pfc)
- (446) Quelle sale temps! J'ai regardé la météo, **ils annoncent pas vraiment d'amélioration** pour cette a-m... (courriel privé, 23.07.10)
- (447) Tu fais très attention si tu dois prendre la route, **ils annoncent la remontée des pluies** vers le Gard dans la journée. Bisous. (88milSMS)

Dans les corpus oraux examinés (PFC, OFROM et CFPP<sup>301</sup>), la séquence *ils annoncent* (ou conjuguée à un autre temps verbal dans OFROM) ne se retrouve que dans ce type de contexte météorologique. On peut signaler que nous n’y avons pas trouvé le verbe conjugué avec *on*. Au vu de la prédisposition de la séquence pour ce contexte discursif<sup>302</sup>, on peut y voir une locution qui vise à exprimer une prévision climatique dont la source n’est même pas traitée, malgré l’évidence à laquelle mènerait une telle inférence. A nouveau, le procédé mitigateur permet de signaler que le locuteur tient l’information d’une source tierce à défaut de la prendre en charge lui-même (cf. *Il va faire froid, beau, mauvais temps*, etc.).

### 3.2.3.3 *Incrimination feinte*

Un rendement différent susceptible d’intervenir dans ces mêmes contextes météorologiques est celui qui consiste à simuler ponctuellement l’existence d’un agent dans un scénario réputé généralement ne pas en contenir (cf. les verbes impersonnels comme *il pleut, il vente, il fait chaud*, etc.). Les exemples suivants, recueillis au vol, personnifient en quelque sorte la responsabilité des changements climatiques, sans que l’on soit pourtant amené à en fournir une valeur :

(448) [la canicule était annoncée pour la veille] Ah aujourd'hui il fait vraiment chaud parce que hier je me suis dit qu'*ils auraient pu faire* mieux question canicule (au vol, 30.06.2015)

(449) [Matin gris et froid] Premier jour de l’automne *ils te le font direct comprendre* (au vol, 21.09.2014)

(450) dimanche on se baignait et aujourd'hui c’est l’hiver non mais *ils sont fous* (au vol, 21.05.2015) = (384)

Dans ces situations, on impute ironiquement l’influence de températures relatives à des agents fictifs, ceux qui auraient le pouvoir de « faire la pluie et le beau temps ». On observe ce rendement dans d’autres contextes également :

(451) *ils ont rajouté des kilomètres* entre Yverdon et Fribourg ou quoi ? (au vol, 10.06.2015)

Ici, le locuteur accuse un responsable du surplus de trajet éprouvé.

Dans un registre similaire, on peut citer la célèbre réplique du dessin animé humoristique *Southpark*, où l’un des personnages, Kenny, meurt à la fin de chaque épisode. L’un de ses acolytes répète alors à chaque reprise :

---

<sup>301</sup> A noter qu’aucune occurrence n’a été trouvée dans CFPP, mais cela n’est pas vraiment étonnant étant donné que le sujet de conversation est contraint, en l’occurrence les différents quartiers de Paris.

<sup>302</sup> Elle n’est pas pour autant absente d’autres contextes, cf. *Je suis à la gare de montpellier j attends le train, ils annoncent 10min de retard. Bisou.* (88milSMS) ou *USA : Aux informations, ils annoncent la mort de Barack Obama au lieu de Oussama Ben Laden, grosse boulette ...* (<http://www.leforumsecret.com>, consulté, le 2 mai 2011)

(452) Oh mon Dieu ! *Ils ont tué Kenny* ! (version française traduite de l'anglais : *Oh my God ! They killed Kenny* !)

Or, ce « refrain » est repris dans toutes sortes de situations, que le personnage soit tué par une classe, un individu unique, un référent [– humain] (e.g. un taureau), ou bien à la suite d'une catastrophe naturelle (une boule de lave volcanique), ou encore d'une inattention de sa part (en jouant au spiroballe, par électrocution, etc.). Certes, il faut tenir compte du genre décalé et ironique de la série télévisée. On remarque cependant que la formule se perpétue quelles que soient les causes de la mort. Quoi qu'il en soit, elle vise surtout à inspirer la pitié sur le sort immuable de la victime.

Ces exemples montrent que l'agent peut être monté de toutes pièces et que malgré son éventuel retrait au profit du procès ou de l'état résultant, sa participation n'est pas complètement évacuée. La fonction d'incrimination n'est pas propre à des scénarios à agent « fictif » : le rendement d'imputation d'une responsabilité se manifeste également avec des procès agentifs ordinaires, comme on l'a déjà vu avec (221), (429) ou (430). Mais le paradoxe qui résulte d'une agentivité que le destinataire sait feinte montre bien qu'il n'y a pas de pertinence à instancier la variable.

#### 3.2.3.4 *Prédicat non typant*

On peut enfin relever un dernier indice plaidant pour un traitement démotivé du pointeur *ils*, à savoir l'expression d'un prédicat non typant (ou non catégorisant). Autrement dit, certaines propriétés prédiquées n'impliquent pas l'appartenance à un type de référent particulier mais peuvent au contraire s'appliquer à des référents divers. Dans l'exemple suivant, la propriété 'poser un grillage' ne sélectionne pas forcément une classe d'individus dont c'est un attribut définitoire :

(453) Merci pour ton message d'encouragements! :D mais tu sais quoi? Je suis scandalisé!  
Faut que je fasse tout le tour pour aller chez toi maintenant :( *ils ont tout grillagé!*  
(88milSMS)

La contingence de la propriété est encore plus claire ci-dessous :

(454) On a vu dans la Liberté qu'*ils ont retrouvé un cadavre* dans l'Aar hier dans l'après-midi dans la zone où on nageait... (courriel, 03.08.2012) = (372)

Dans cet extrait de courriel, l'auteur résume une information dont il a pris connaissance par la presse. Le prédicat verbal ne permet pas de cibler une classe dont le rôle est de 'retrouver un cadavre'. Par ailleurs, nous avons récupéré l'extrait de presse auquel le courriel fait visiblement écho :

(455) Noyade dans l'Aar (titre) : Le corps sans vie d'un asiatique a été repêché hier après-midi dans l'Aar à Berne. C'est un passant qui a donné l'alerte en le voyant descendant la rivière au-dessous du pont de Kirchenfeld, a précisé la police bernoise. Une heure et

demie plus tard, il a pu être sorti de l'eau au-dessous du restaurant de Schwellenmätteli. (*La Liberté* < ats, 03.08.2012)

On constate que le processus de découverte du corps résulte de la participation d'actants multiples (passant, police). On remarque également l'usage du passif à double reprise dans l'extrait source, visant à mettre en avant l'objet patient et le procès, au détriment de l'agent (*cf. infra* §3.4.1), conformément à l'enjeu d'un article de fait divers de ce type. Au vu du caractère « collaboratif » de la récupération du cadavre et de la non pertinence de l'identité de l'agent, l'usage de *ils* en (454) apparaît comme une alternative commode pour indifférencier la source de l'action et en venir au fait, à savoir celui de partager la stupeur suscitée par la découverte macabre.

### 3.2.3.5 Bilan

Cette section a mis jour quelques indices orientant vers une analyse « postiche » du pointeur *ils*, dont le rôle actantiel n'est cependant pas complètement éliminé. On peut ainsi invoquer la promotion de la relation procès-objet au détriment de la relation agent-procès, pour des raisons argumentatives (§3.2.3.1) ; on peut aussi relever l'aspect lexicalisé de certaines séquences, comme l'association de *ils* à des verbes de parole marquant l'hétérogénéité énonciative du propos asserté, dont le rendement mitigateur relègue au second plan la résolution référentielle (§3.2.3.2) ; la simulation du rôle d'agent dans un scénario qui en est d'ordinaire privé met également en évidence le paradoxe d'une tentative d'identification (§3.2.3.3) ; enfin l'attribution de propriétés contingentes, à défaut d'autres indices contextuels, peut également favoriser l'abandon d'une procédure interprétative (§3.2.3.5).

## 3.3 Discussion

Le classement ci-dessus a permis de révéler les points communs et différences existant entre les usages de l'indice de 6<sup>e</sup> personne *ils* à valeur sous-déterminée. Nous avons ramené un certain nombre de cas au phénomène d'anaphore indirecte car on relève des indices contextuels favorables à l'inférence d'un référent dont on peut conjecturer des traits distinctifs, en général, à partir de la présence en M d'une entité collective, d'un lieu ou alors via la reconstitution d'un scénario. Dans ce cas de figure, il est possible d'interpréter la séquence verbale comme une prédication sur la classe concernée, à laquelle – et parfois de manière distributive – peuvent être attribuées des attitudes, réactions ou émotions et qui reste disponible en M pour d'éventuels pointages successifs. D'autres cas, à l'inverse, sont difficilement analysables en termes d'anaphore : le processus d'unification référentielle paraît court-circuité et le pointeur référentiellement démotivé, sans toutefois que soit annulée la relation agentive impliquée par le verbe. Dans cette situation, la séquence en *ils* s'interprète peu vraisemblablement comme une opération binaire de prédication sur un sujet (*cf.* un énoncé dit « catégorique »), sans toute fois s'assimiler à un énoncé « thétique », autrement dit dont l'interprétation est la simple reconnaissance d'un état-de-choses. En effet, si le procès est bien mis en évidence, la relation actantielle avec le sujet demeure toutefois à l'arrière-plan.

Entre ces deux analyses, on trouve toute une série de faits dont les indices sont ambivalents. Certaines situations dialogiques ont toutefois permis de mettre en évidence la coexistence, entre les interlocuteurs, de deux traitements concurrents. Le tableau ci-dessous résume les trois catégories illustrées chacune par un exemple :

Anaphore indirecte	Analyse ambiguë	Variable non instanciée
(395) donc c'est un groupe qui a dû organiser des un jeu de nuit dans la forêt [...] alors <i>ils ont dû</i> / _ / <i>organiser donc des jeux de nuit</i>	(420) j'ai une amie là qui a eu son concours l'année dernière qui, qui est instit. Donc au début <i>ils, ils l'ont mise, à Meylan</i> et tout	(453) Faut que je fasse tout le tour pour aller chez toi maintenant :( <i>ils ont tout grillagé!</i>

La récurrence de cas faisant l'objet d'une analyse ambiguë est peut-être le signe d'une réanalyse<sup>303</sup> ou métanalyse<sup>304</sup> du clitique *ils*, qui peut se voir « recyclé » comme élément postiche non instancié, un peu à la manière de la construction asubjectale *ça*+V, étudiée par Maillard (1989, 1994b) : à l'origine pleinement référentiel, on y recourt volontiers de nos jours lorsqu'il s'agit d'« impersonnaliser » un nouveau verbe<sup>305</sup>, là où *il* « impersonnel » ne se montre plus du tout productif car ne s'associant qu'à une liste finie de verbes (Maillard 1989 : 37). L'auteur relève à l'inverse la « créativité du schème » en *ça* comme l'illustraient à l'époque les expressions *ça boume, ça va barder, ça va chier, ça bouchonne, ça gaze, etc.* (*ibid.* 92-93). Il fait de *ça* un « nouveau régisseur impersonnel » (1994b : 50-52), autrement dit, une sorte d'opérateur de diathèse bloquant la position sujet et par là, la participation même d'un agent. Cette analyse de *ça* pourrait apporter des éléments de réponse à la « restriction » (à confirmer) de *ils* à la position sujet dans le cas où il reste non instancié (*vs* en anaphore indirecte, *cf.* (400) et (399)). Cela dit, on ne peut aller aussi loin avec *ils*, dont l'agent, à défaut d'être référentiellement instancié, reste sémantiquement impliqué, contrairement à celui de *ça*. On note cependant la subtile modification qu'il induit dans la relation actantielle, reléguant au second plan la relation agent-procès au profit de la promotion du procès (et de son éventuel son objet), parfois de l'état qui en résulte.

Si l'on rappelle les rendements discursifs propres à chaque catégorie relevée, on note quelques différences, mais aussi de nombreux points communs. L'une des motivations possible pour un pointage via *ils* spécifique à l'anaphore collective (éventuellement aussi à l'anaphore « locative ») réside dans l'attribution d'une propriété distributive au référent, d'où

<sup>303</sup> Langacker (1977 : 58) propose la définition suivante : « I will define *reanalysis* as change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation. Reanalysis may lead to changes at the surface level, as we will see, but these surface changes can be viewed as the natural and expected result of functionally prior modifications in rules and underlying representations ».

<sup>304</sup> Selon Blinkenberg (1950 : 43), « [l]a métanalyse suppose en effet une même forme analysable de deux façons. C'est la phrase à double sens qui est le point de départ et le pivot du mouvement qui amène le regroupement, comme c'est la phrase à double forme qui en est le point d'arrivée. On n'explique aucune métanalyse sans s'appuyer sur des exemples équivoques, on ne prouve la métanalyse que par des exemples univoques. »

<sup>305</sup> *ça télécharge tout seul !* (forum <http://www.commentcamarche.net/>); *ça roots ?* (≈ *ça va ?*, oral, au vol), *sur le net, ça bogue* (*Libération*, 17.11.2014).

le recalibrage de la collection au format de classe, *cf.* (398) et (400). Un autre objectif guidant l'emploi de *ils* au détriment d'une expression lexicale consiste à éviter la redondance qu'elle pourrait engendrer à travers la répétition d'un N collectif (*cf.* la reformulation dans (395)). On pourrait peut-être même étendre cette stratégie aux cas où la redondance concernerait le radical d'un N de lieu et d'un N de gentilé, voire d'un verbe et d'un N prédicatif :

(456) [...] je passe dix jours en **Belgique**. Le pays des moules, des frites, du chocolat – si, si, **les Belges** y croient ! – et des bières évidemment. (exemple modifié) = (402)

(457) [...] **Les voleurs ont volé** un appareil photo et qq trucs du boulot [...] (exemple modifié) = (416)

Toutefois, dans la plupart des cas, l'alternative d'un SN lexical n'a pas besoin d'être envisagée, l'enjeu reposant sur l'exploitation de la sous-détermination *ils*. En effet, dans les trois catégories, on a relevé le rendement de diffusion indistincte de la responsabilité du procès sur un agent collectif, dont l'identité soit demande à être globalement inférée (§3.2.1), soit demeure en suspens (§3.2.3), qu'elle soit insignifiante ((453), (454)), ignorée (430) ou encore délibérément occultée (429). Dans ces cas-là, l'indistinction peut conduire à reléguer l'actant au second plan dans l'expression du procès qui hérite ainsi lui-même du premier plan.

Bien que nous ayons essentiellement travaillé sur des données orales spontanées, nous avons également relevé le phénomène dans d'autres genres de « l'immédiat » (Koch & Oesterreicher 1985) comme les SMS, les courriels privés, les réseaux sociaux, etc. Dans les genres discursifs réputés plus travaillés, les occurrences de *ils* à valeur sous-déterminée ne sont pas pour autant absentes, comme en témoignent les extraits littéraires reproduits, mais en priorité dans des contrefaçons de l'oral, des dialogues ou alors dans des récits à la 1<sup>ère</sup> personne, reflétant les pensées du narrateur-personnage. Notons que malgré notre contact privilégié avec les écrits scientifiques, nous n'y avons pour l'heure, sans grande surprise, relevé aucun emploi de ce type (*cf.* le recours à d'autres procédés ci-après §3.4). Il semble donc que les normes propres aux genres discursifs aient une grande influence sur la distribution des emplois de *ils*. Néanmoins, l'inventaire en amont montre qu'il est simplificateur de réduire ceux-ci à une stratégie d'économie propre au style « relâché », au vu des rendements diversifiés dont ils sont porteurs.

A cet égard, les rendements décrits ne sont pas l'apanage de *ils*, comme nous l'avons suggéré à plusieurs reprises. Il existe en effet d'autres ressources linguistiques, visiblement mieux tolérées dans les genres plus centrés sur les intérêts de l'interprète, offrant certains rendements similaires et affectant de la sorte la distribution de *ils*. Nous proposons à présent de mettre en perspective ces procédés avec les constructions en *ils*.

### 3.4 *ILS dans le paradigme des constructions à agent sous-déterminé*

Plusieurs auteurs ont relevé des similitudes entre l'emploi de *ils* et *on*<sup>306</sup> d'une part (Goudet 1983, Siewierska 2011, Cabredo Hofherr 2014, Creissels *à par.*), et *ils* et le passif d'autre part (Yule 1982, Gundel *et al.* 1993, [Reichler-]Béguelin 1993a, Myhill 1997, Blevins 2003, Siewierska 2008, Cabredo Hoffherr 2014). En effet, tous sont susceptibles de manifester la sous-détermination d'un agent, ainsi que l'illustrent les exemples ci-dessous :

(458) il y a eu énormé/ certainement une civilisation, dans le sous-sol de Gières. Parce que, quand, *ils ont construit le groupe scolaire* là, le groupe sur la place, le groupe scolaire qui s'appelle René Cassin maintenant, il y a eu, *ils ont découvert des vestiges* (pfc) = (468)**Erreur ! Source du renvoi introuvable.**

(459) [...] quand on a construit le groupe scolaire [...] on a découvert des vestiges (exemple modifié)

(460) [...] quand le groupe scolaire a été construit [...] des vestiges ont été découverts (exemple modifié)

A travers la manipulation des énoncés, on constate que chaque procédé permet l'occultation de l'identité de l'agent engagé dans le procès. Dans l'exemple authentique suivant, cette fonction commune est particulièrement mise en évidence dans l'exploitation des trois structures, dont on peut contextuellement inférer la coréférence de l'agent :

(461) [début d'un article sur des cas de mobbing dans une école] « *Mon fils a été frappé* et *on lui a uriné* dessus. Ensuite, *ils l'ont fait chanter* », s'indigne Gabriella Schlegel (presse, *20minutes.ch*, 29.10.2015 < Béguelin 2015)

Néanmoins, chaque construction comporte des spécificités à différents niveaux, qui influent sur sa mise en œuvre. Dans les sections qui suivent, nous proposons de confronter l'usage de *ils* avec les deux autres procédés afin de cerner une éventuelle zone de recouvrement et de dégager les facteurs et circonstances qui motivent le choix pour l'un au détriment des autres.

#### 3.4.1 *ILS vs le passif*

Les constructions en *ils* dont la valeur reste non instanciée sont souvent présentées comme fonctionnellement équivalentes aux constructions passives sans complément d'agent (Myhill 1997, Blevins 2003, Siewierska 2008b) :

(462) They've stolen my bag (Siewierska 2008 : 31)

(463) My bag has been stolen (*ibid.*)

On peut tout d'abord remarquer que la comparaison ne concerne que les verbes transitifs à deux actants, à savoir un agent et un objet patient. Bien que la construction transitive soit

---

<sup>306</sup> Ou ce qui est présenté comme leurs formes « correspondantes » dans d'autres langues.

privilegiée dans nos données en §3.2.3, l'usage de *ils* est a priori moins contraint, comme le montrent les emplois à « agent fictif » où il se combine à la copule *être* (450), à une construction factitive (449) ou à un verbe modal (448).

La diathèse au passif peut être décrite comme un processus transformationnel ou dérivationnel à partir d'un verbe transitif qui subit une dérivation suffixale générant un adjectif (participe passé) (Berrendonner 2000). Le patient se voit promu à la place d'argument sujet tandis que l'agent se mue en actant interne au procès. L'agent peut néanmoins être exprimé par un complément oblique (e.g. *par x, de x*), mais c'est loin d'être la norme. D'ailleurs, dans le cadre des constructions qui nous intéressent, la restitution explicite d'un agent ne permet plus de mettre en œuvre un procédé de sous-détermination :

(464) quand le groupe scolaire a été construit *par eux* [...] des vestiges ont été découverts *par eux* (exemple modifié)

Blevins (2003) considère que le processus de détransitivisation de la diathèse passive réduit la valence lexicale du verbe. On peut invoquer la notion de diathèse *récessive*, selon la terminologie de Tesnière (1959 : 272), qui « diminue d'une unité le nombre des actants ». En fait, cette conception doit être nuancée. Pour Berrendonner (2000 : 47), si le passif « absorbe » l'agent dans le lexème prédicatif, cet agent n'en reste pas moins présent, quoiqu'indéterminé. Selon Muller (2000 : 51-52), le passif ne fait que reléguer « au rang de relation facultative » la relation agent-action, sans pour autant l'oblitérer. Creissels (2006 : 9) va dans le même sens : l'agent peut « être complètement occulté, mais sa participation même à l'événement reste impliquée »<sup>307</sup>. Certaines formes passives, en particulier les « passifs d'état » (Riegel *et al.* 2009 : 736), présentent à cet égard un comportement remarquable :

(465) Ils [les fruits et légumes] contiennent beaucoup de vitamines et de minéraux, bien davantage que lorsqu'ils sont *cuits* (Nitshe, C., *Le décodeur minceur*, 2012, p. 44)

L'énoncé présente la même forme qu'une structure adjectivale de type attributive (par ex. *frais, crus*, ici le participe passé résultatif *cuits*). Cependant, on peut considérer ce genre de constructions comme passives dans la mesure où elles dénotent le résultat du procès achevé 'ils ont été cuits'. Si la construction se voit suivie d'un complément d'agent, on retrouve l'aspect inaccompli :

(466) [Il s'agit d'une expérience rapportée dans une étude] Deux gigots d'un poids égal *sont cuits par un cuisinier intelligent*, dans les mêmes conditions de temps et de chaleur. (Loverdo, J., *Le froid artificiel et ses applications industrielles : commerciales et agricoles*, 1903, p. 336)

Dans cet exemple, l'aspect inaccompli de la construction active correspondante (*un cuisinier cuit deux gigots*) est conservé. Dans ce cas, le verbe *être* correspond aux auxiliaires d'autres

---

<sup>307</sup> Creissels démontre cela au moyen de l'énoncé *La porte a été ouverte*, qui accepte le modifieur *volontairement*, mais non le modifieur *toute seule*, à l'inverse de l'énoncé *La porte s'est ouverte*.

langues exprimant le sens de *(de)venir* (all. *werden*, it. *venire* ou *andare*) : il indique la modification en devenir subie par l'objet patient (Muller 2000).

Dans le cas du passif d'état, *être* ne s'interprète pas comme un processus en devenir mais signale l'aboutissement du processus : on constate cependant que l'interprétation résultative n'est pas toujours bien distincte d'une interprétation purement attributive. Dans l'énoncé (465), on peut supposer que l'état résulte d'une action comprenant un agent implicite. Par contre, dans l'exemple suivant, il est difficile de présumer l'intervention d'un agent (préalable ou en cours) :

(467) la mer est salée (Riegel *et al.* 2009 : *ibid.*)

Lorsqu'on ne peut reconstituer un agent, par transposition à partir d'un verbe à l'actif (*on a salé la mer*), d'un passif processif (avec ajout d'un complément d'agent) ou qu'on ne peut considérer que l'état résulte d'un procès accompli (*la mer a été salée*), on ne traite généralement pas la construction en termes de diathèse passive, mais comme une simple construction adjectivale ou attributive (comme dans *Il semblait déçu* < *ibid.*). Comme on le verra, les indices ne sont pas toujours univoques et le débat autour de ces constructions, qui dépasse le propos de ce travail, est loin d'être clos.

Cela dit, le fait qu'un passif puisse avoir une interprétation inaccomplie ou résultative (stative) offre un éclairage intéressant sur nos données. On peut en effet supposer que le choix du passif est motivé dans certains cas par des questions d'aspect et d'organisation de la structure actantielle. Certains extraits d'oral fournissent des alternances significatives entre l'emploi du passif et l'emploi du *ils*. Revenons à l'exemple introductif, qui comprend une reformulation au passif (soulignée) d'un procès exprimé à l'actif avec *ils* :

(468) E : Parce qu'il y a pas eu de galeries creusées, il y a pas eu/ ?

AS : Non, il y a eu énormément/ certainement une civilisation, dans le sous-sol de Gières. Parce que, quand, ils ont construit le groupe scolaire là, le groupe sur la place, le groupe scolaire qui s'appelle René Cassin maintenant, il y a eu, ils ont découvert des vestiges, *ils avaient arrêté les travaux. Ils avaient arrêtés les travaux, les travaux ont été arrêtés pendant deux mois.* il y avait des galeries, des, des briques, des trucs comme ça, il y a certainement eu une civilisation. (pfc)

Cette « reformulation » permet d'exprimer sans doute plus clairement l'état résultant de l'arrêt des travaux que la forme transitive *ils avaient arrêté les travaux*, signalant l'accomplissement du procès (dont le résultat n'est qu'implicite) : c'est en effet le résultat statif du procès qui hérite de l'aspect duratif exprimé par le circonstant (*pendant deux mois*). En outre, du point de vue de la progression thématique (Combettes 1983), on repère une progression linéaire, où le rhème de l'énoncé précédent (*les travaux*) devient le thème du suivant. On observe ainsi une réorganisation de la structure informationnelle, en particulier à travers la promotion de l'objet patient en position de sujet, mettant en relief la relation qu'il entretient avec le procès.

Dans l'exemple suivant déjà cité *supra*, la reformulation est opérée en sens inverse, c'est-à-dire qu'une construction au passif se voit reformulée via une construction à l'actif avec *ils* :

(469) donc faut savoir que | \_ | c'est tous les cours deux sont plus basés | \_ | euh performance | \_ | donc euh par exemple euh pour euh athlétisme | \_ | il y a le cent mètres qui est chronométré | \_ | ou | \_ | c'est les tous les barèmes pour les pratiques sont basés | \_ | sur euh les meilleurs é- étudiants euh des universités en sport | \_ | donc tout toutes les meilleures performances euh y euh | de chaque euh canton ou université de Suisse | \_ | forment la moyenne donc ***ils basent les barèmes d'après ça*** | \_ | donc euh ce qu'il faut savoir c'est que bien souvent euh c'est des barèmes assez élevés (ofrom) = (419)

L'extrait porte sur la question des barèmes des examens en sport, à propos desquels on apprend la manière dont ils sont calculés, et de fait, leur sévérité (*c'est des barèmes assez élevés*). Au vu du développement argumentatif, 'les barèmes' constitue ainsi le centre organisateur courant dans M. Le locuteur présente d'abord le calcul des barèmes comme un fait établi au moyen du passif (interprétation stative). S'ensuit un éclaircissement sur ce qu'il entend par *les meilleurs étudiants des universités en sport*. Puis il recourt à la reformulation avec *ils*, qui fait écho à l'agent implicite de la construction passive, visant à récapituler globalement comment sont conçus les barèmes. L'aspect inaccompli du verbe à l'actif permet d'insister ainsi non plus tant sur le résultat (le fait établi) que sur la manière dont sont calculés les barèmes, par une sorte de métalepse (la conception étant une étape antérieure au résultat). On voit ainsi comment il est possible d'envisager la situation sous divers angles aspectuels selon les besoins de l'argumentation.

Le choix entre le passif et l'emploi de *ils* est donc visiblement motivé par des critères liés à l'aspect du verbe (résultat vs processus) et de la structure informationnelle (type de progression, réorganisation des rôles sémantiques). A cet égard, Lambrecht (1994) fait l'hypothèse qu'il existe une loi communicative selon laquelle il n'est pas optimal d'introduire un référent et de prédiquer sur lui dans le cadre d'une même énonciation, ce qui expliquerait en partie pourquoi les SN indéfinis sont évités en position sujet, position typiquement thématique. Dans cette perspective, la volonté d'introduire des objets nouveaux pourrait orienter vers l'emploi d'une structure active, plaçant ceux-ci en position d'objet direct plus adaptée à la fonction rhématique d'apport informationnel :

(470) juste tu ne peux plus tourner chez moi maintenant ***ils ont mis des poteaux*** (ofrom)

(471) ***ils nous ont presque pas donné d'explications*** = (101)

(472) Tu sais bien que cette année avec la sécheresse avec la canicule, les montagnes euh, les neiges étaient très, très fondues, les glaciers ont fondu et caetera. ***Ils avaient, ils ont dû mettre des flics*** pour que les gens ne montent pas au Mont Blanc (pfc)

Il va de soi qu'aucune de ces constructions n'est impossible au passif, mais on remarque que la « promotion » de l'objet nouveau délivré en position sujet chamboulerait passablement la

structure informationnelle d'origine. Pour les deux derniers exemples, une diathèse au passif engendrerait en outre une surcharge cognitive, respectivement dans la production d'un sujet quantifié étendu (473) et dans l'expression de trois auxiliaires successifs (474) :

(473) Presque pas d'explications nous ont été données (exemple modifié)

(474) des flics ont dû être mis pour que les gens ne montent pas au Mont Blanc (exemple modifié)

La mise en perspective des constructions active et passive révèlent ainsi qu'elles ne sont de loin pas des variantes libres malgré certains traits et effets similaires<sup>308</sup>.

On peut encore noter certaines différences supplémentaires entre les constructions étudiées. Tandis que la forme *ils* exhibent des marques morphologiques, l'agent non exprimé du passif apparaît à cet égard beaucoup moins spécifié. En particulier, il se montre insensible au trait [+délocuté] propre aux indices de 3<sup>e</sup> et 6<sup>e</sup> personne :

(475) donc bon les journées de ski donc euh ben on a pris des abonnements directement à la semaine | \_ | comme ça ben on | c'était quelque chose de plus avantageux pour euh nous euh les organisateurs du camp | \_ | euh le ski et donc euh **les enfants sont répartis** dans des groupes de six à | \_ | à neuf euh élèves six à neuf enfants (ofrom)

En (475), aucune information explicite ne signale l'inclusion ou l'exclusion du locuteur à l'égard du rôle d'agent dans le procès exprimé, par opposition à *ils ont réparti les enfants*. La construction au passif, non marquée sur ce point, peut donc recouvrir une situation où le locuteur a participé ou non à la répartition effective. Cet exemple montre en outre que le nombre de l'agent reste lui aussi indéterminé, la diathèse recouvrant une action potentiellement individuelle ou collective.

On peut également songer à la question de l'interprétation humaine de l'agent, souvent présentée comme inhérente à l'emploi de *ils* à valeur sous-déterminée (voir *supra* §2.2.2.3).

<sup>308</sup> Siewierska (2010) dégage une parenté morphologique entre les constructions dans certaines langues, où la marque de 3<sup>e</sup> personne du pluriel est présentée comme étant à la source du passif : dans les langues nilotiques, le passif est perçu comme à l'origine de la marque de 6<sup>e</sup> personne ; quant aux langues bantoues, la 6<sup>e</sup> personne est encore visible en synchronie dans les verbes au passif (le morphème *a-* ci-dessous) :

Kimbundu (Givón 1976 : 180)

a- mu- mono  
3PL 3SG saw

“They saw him”

Ou :

Nzua a- mu- mono kwa meme  
Nzua PASS- 3SG saw by me

“Nzua was seen by me”

De surcroît, dans des langues qui ne possèdent pas de passif, les indices de 6<sup>e</sup> personne sont souvent utilisés pour traduire le passif des langues européennes (Siewierska 2010) :

Libum (Nforgwei 2004 : 278)

A fa rlii muu ene jisos  
3PL give name child that Jesus

“He was named Jesus/ They named him Jesus”

Le passif apparaît pour sa part non marqué à cet égard, comme le montrent les exemples ci-dessous, qu'il régit ou non un complément d'agent :

(476) Vers 19H50, une personne a grimpé sur le toit d'un train qui repartait de la gare de Mantes-la Jolie vers Paris. **L'homme a été électrocuté par une caténaire et projeté** sur les voies. (<http://www.lefigaro.fr/flash-actu/>, 27.03.2013)

(477) Une triple fracture du bassin, neuf côtes et trois vertèbres cassées : Didier Castella **a été grièvement blessé** le 23 juillet lors d'une excursion en quad en Colombie britannique, à l'est du Canada. (presse, *La Gruyère*, 04.08.2016)

On remarque que l'emploi d'un *ils* engagerait préférentiellement une interprétation humaine de l'agent :

(478) Une triple fracture du bassin, neuf côtes et trois vertèbres cassées : **ils ont grièvement blessé Didier Castella** le 23 juillet lors d'une excursion en quad en Colombie britannique, à l'est du Canada. (exemple modifié)

En effet, le passif oriente moins sensiblement vers une interprétation humaine (ou animée). Cela nuance les observations de Blevins (2003 : 475) selon lesquelles aussi bien l'impersonnel, le passif sans complément d'agent et les constructions verbales actives à la 6<sup>e</sup> personne impliquent des agents par défaut humains.

Par ailleurs, on peut s'interroger sur la nature inhérente ou non du trait [+humain] pour *ils* : plutôt qu'en faire un trait inhérent à l'emploi de *ils*, [Reichler-]Béguelin (1993b) (*cf. supra* §2.2.2.3) met l'interprétation humaine au compte de la prédominance en général des référents humains pour le pronom de 6<sup>e</sup> (l'hypothèse de « l'anthropocentrisme des discours »). Sur un état antérieur de la base OFROM (décembre 2012), signalons que nous avons relevé seulement 46 interprétations [-humain] du pronom *ils* (toutes occurrences de *ils* confondues<sup>309</sup>) sur un total de 822 occurrences, soit 5,6 % contre 94,4% de références [+humaine].

Si l'on recherche les formes passives dans OFROM<sup>310</sup> et qu'on examine le type d'agents implicites, les résultats sont *a priori* variés, bien que nous n'ayons pas fait de comptage exhaustif :

---

<sup>309</sup> Y compris les *ils* anaphoriques « classiques » (*les jeunes aimaient bien ça avec le pain c/ c/ | \_ | ils ils apprécient hein la nutella*), les redoublements du sujet (*les trois cons ils attendaient leur tour*), etc. Par contre, nous n'avons pas inclus les occurrences répétées dans les bribes (comme la 2<sup>e</sup> occurrence dans *ils ils apprécient*) ou qui interviennent dans des structures abandonnées (anacoluthes) (*donc euh ils | \_ | ils | \_ | c'étaient des masters plus orientés [...]*).

<sup>310</sup> Nous avons recherché les formes du lemme *être* à tous les temps simples suivies d'un participe passé à moins de deux secondes d'intervalle. Mais la requête comporte du bruit parce qu'elle retient tous les verbes à une forme composée avec l'auxiliaire *être* (*e.g. être allé*, etc.). Pour les temps composés, nous avons recherché *été* suivi d'un autre participe passé. Là aussi, un grand tri manuel s'est avéré indispensable, en raison de nombreux phénomènes d'homonymie (n.m. *été*) dus à des problèmes d'annotation en amont. Pour cette requête, nous avons passé en revue les 200 premières occurrences, produites par 58 locuteurs différents.

- (479) donc il faudrait que le peuple valaisan puisse se rallier derrière un entraîneur | qui est du coin | \_ | et pis là du coup comme *il a été viré* on se retrouve avec Docastel | qui est déjà passé à Sion (ofrom)
- (480) c'est une série de vidéos | tournées dans les dans des parcs publics à New York | \_ | et *qui ont été postées* sur le portail internet Youtube dès le milieu des années deux mille | qui marque véritablement la naissance du street workout dans sa forme actuelle (ofrom)
- (481) alors ça ça fait partie aussi de mon travail | dans la mesure du possible si ce n'est pas trop compliqué | \_ | je peux *j'ai été amenée* à à faire aussi ce genre de choses donc (ofrom)
- (482) c'était le plus jeune je pense de l'assemblée à peu près | \_ | et pis franchement euh | \_ | # | *elle a été traumatisée* parce que | \_ | quand euh | \_ | quand ils lui ont demandé quand euh quand il est arrivé et que | \_ | et qu'il a fait euh elle lui a demandé de quel pays euh il venait il a fait des États-Unis | \_ | hein | \_ | pis ça pouvait pas être autre chose (ofrom)

Dans les deux premiers exemples, les prédicats sélectionnent typiquement des agents humains, mais ce n'est pas le cas dans les deux autres exemples dont l'agent demeure sous-spécifié. Les passifs d'état semblent particulièrement révélateurs de l'absence de marquage du trait d'humanité :

- (483) tu vois elle elle est quand même dans le monde du service | et elle a des *horaires qui sont complètement décalés* | elle comprend pas que le matin je me lève à six heures et demi pour aller travailler (ofrom)
- (484) il y a des enfants qui jouent dans la dans la boue | \_ | et puis le taxi gentiment comme en Inde | \_ | euh prend son chemin | \_ | par cette foule | \_ | tout gentiment | \_ | parce que tu veux pas euh | \_ | rouler sur des gens | \_ | ça euh là | maintenant *je suis euh habitué* à ça mais la première fois *j'étais / \_ / complètement choqué* (ofrom)

On peut d'ailleurs relever la difficulté de distinguer ci-dessus entre passif d'état et construction attributive : dans l'exemple (483), on peut envisager un agent non humain responsable du décalage, comme le travail de la personne en question, avec une lecture résultative du passif ; mais on peut également y voir une construction attributive au vu de la substitution possible de *décalés* par de simples adjectifs (*irréguliers, fous*, etc.) ou de la copule *être* avec *sembler*. Quant à (484), le participe passé *habitué* à se laisse volontiers substituer par des adjectifs (*familier de, coutumier de*), mais il n'est pas forcément incompatible avec l'implication d'un agent (par exemple l'expérience, la réalité, etc.). Il en va de même pour la construction *j'étais choquée* subséquente. Quoi qu'il en soit, contrairement à l'emploi de *ils* où l'interprétation [+humain] est prédominante, la nature de l'agent des formes passives semble *a priori* plus aléatoire sur ce point. Néanmoins, une étude plus minutieuse sur l'emploi du passif serait nécessaire pour étayer ces hypothèses.

Cette brève confrontation des constructions active et passive met ainsi en évidence des différences d'emploi liées à des paramètres aspectuels (accompli/inaccompli, résultatif/attributif) et informationnels (structure, progression thématique, relation et ordre des actants et leur rapport avec le procès, etc.). Au niveau de la nature de l'agent, nous avons montré que celui impliqué par le passif était moins spécifié que celui impliqué par *ils*, s'en distinguant par son indifférence au nombre, à l'inclusion du locuteur et manifestement au trait d'humanité. Nous n'avons pas comparé exhaustivement l'influence du genre de parole ou du registre de langue. A cet égard, plusieurs auteurs mentionnent une différence à propos de l'anglais : l'emploi de *they* « impersonnel » ayant tendance à être perçu comme familier (Siewierska 2008), ils suggèrent que l'emploi du passif serait privilégié dans des registres soutenus et écrits (Yule 1982, Myhill 1997). Pour le français, on a relevé *supra* (§2.1) le jugement que portent les grammairiens à l'égard de *ils* à valeur sous-déterminée (style familier, voire populaire). Ce qu'on peut avancer, c'est qu'une semblable attitude ne se rencontre pas, à notre connaissance, à l'égard du passif. L'une des raisons est sans doute l'instruction d'unification référentielle fortement attachée au pronom de 3<sup>e</sup> et 6<sup>e</sup> personne, qui n'est pas requise pour l'agent impliqué du passif. Il n'est donc pas surprenant que les genres influencés par des exigences de « style » et d'interprétabilité soient peu féconds en *ils* à valeur sous-déterminée. Mais cela ne signifie pas que le passif soit réservé aux modes de la « distance » (Koch & Oesterreicher 1985), comme en témoignent les occurrences usuelles de notre bref aperçu provenant de la base OFROM. Nous pouvons à présent nous tourner vers un autre moyen d'exprimer la participation d'un agent sous-déterminé.

### 3.4.2 ILS vs ON

Dans certaines circonstances que nous allons caractériser (critères syntaxiques, sémantiques et pragmatiques), *ils* et *on* peuvent apparaître en concurrence. Les grammairiens, on s'en souvient, notent régulièrement ce rapprochement en considérant *ils* comme la variante familière ou populaire (Haase 1898, Brunot & Bruneau 1949, Sandfeld 1970). Le pronom *on* fait l'objet d'une littérature abondante<sup>311</sup> que nous ne pouvons examiner dans le détail ici. Néanmoins, nous proposons une synthèse, forcément simplifiée, des principaux apports de ces travaux, à des fins contrastives avec l'emploi de *ils*.

Du point de vue syntaxique, *on* n'apparaît que comme clitique sujet. Sur le plan sémantique, outre son trait [+animé], *on* est particulièrement sous-spécifié, ce qui lui vaut une flexibilité remarquable d'usage. Riegel *et al.* (2009 : 364) relèvent l'adéquation de la définition de « vague sujet » proposée par les verbicrucistes. Wilmet (2007 : 58) invoque son caractère « omnipersonnel » dont le contenu peut s'étendre « de 'quelqu'un, qui que ce soit, n'importe qui' à 'tout le monde' » ; les extrémités de ce continuum sont illustrées respectivement par (485) et (486) :

---

<sup>311</sup> Voir entre autres Muller (1970), Boutet (1984), François (1984), Blanche-Benveniste (1987) et (2003), Viollet (1988), Leeman (1991), Rabatel (2001), Guerin (2006), Fløttum *et al.* (2007), Béguelin (2014c).

(485) *On* a sonné. (Riegel *et al.* 2009 : 364)

(486) En Bavière, *on* boit beaucoup de bière. (*ibid.*)

Le clitique *on* a en particulier la possibilité de recouvrir la référence de toutes les personnes de la conjugaison (Moignet 1965, Riegel *et al.* 2009 : 364) :

(487) Excusez-nous d'arriver en retard. *On* a eu une panne. (Riegel *et al.* : *ibid.*)

(488) Alors, *on* fait la forte tête ? (*ibid.*)

(489) - Comment ça va ?  
- *On* fait aller. (*ibid.* : 365)

(490) Je les avais prévenues, mais *on* n'a pas voulu m'écouter. (*ibid.*)

La diversité des emplois de *on* montre d'emblée le caractère restreint de la zone éventuelle de recouvrement avec *ils*. La flexibilité référentielle de *on* n'a semble-t-il d'égal dans aucune autre langue européenne (Fløttum, Jonasson & Norén 2007) et se répercute sur la morphosyntaxe à travers la diversité des accords possibles des adjectifs ou participes passés (Béguelin 2014c). Elle met en outre à l'épreuve les linguistes sur le débat d'un traitement polysémique *vs* homonymique de *on*<sup>312</sup>. En tout cas, *on* invoque généralement la nature « indéfinie » du pronom pour les emplois (485) (*on* tantôt appelé *spécifique, existentiel*) et (486) (*on* dit *générique, gnomique* ou *universel*) et un emploi de 4<sup>e</sup> personne pour (487), que les grammaires normatives tendent à stigmatiser (Béguelin (2014c). Certains emplois exploitent simultanément diverses caractéristiques de *on*, comme celui-ci :

(491) [à propos de participants à un camp de ski que le locuteur organise] une bonne euh partie est également un peu triste de se quitter parce que ben **on** retrou- | **on** retourne chez les parents **on** se retrouve tout seul sans les amis donc c'est c'est une autre ambiance que d'être euh entouré (ofrom).

En effet, les interprétations semblent se brouiller : du point de vue du locuteur, *on* peut renvoyer à chacun des enfants pris comme un individu anonyme spécifique, mais en même temps, il reflète le point de vue interne de l'enfant sur le mode du discours indirect libre, discours par ailleurs présenté comme stéréotypique. Selon les distinctions établies ci-dessus, cet emploi de *on* se situe donc au carrefour des emplois spécifique, de 4<sup>e</sup> personne (= *nous*) et générique...

L'un des facteurs de la souplesse référentielle de *on* se situe dans son absence de marquage sémantique quant à l'inclusion ou l'exclusion du locuteur, par opposition aux autres indices personnels (Creissels *à par.*). Le pronom *on* se différencie sur ce point à la fois clairement de *nous*, toujours inclusif (Blanche-Benveniste 2003), et des 3<sup>e</sup> et 6<sup>e</sup> personnes, par définition

---

<sup>312</sup> L'examen des potentialités de coréférence des différents emplois du pronom *on* par Creissels (*à par.*) penche pour un traitement homonymique.

[+délocuté]. C'est le contexte, le cas échéant, qui est responsable d'une interprétation incluant ou non le locuteur. Mais dans bien des contextes, ce point reste indécidable :

(492) [le locuteur parle de son métier de monteur en images] maintenant **on** travaille sur des logiciels de montage euh donc c'est du montage virtuel à l'époque **on** faisait du montage euh sur bandes (ofrom)

S'il y a des raisons de penser que le locuteur s'inclut, étant donné son métier, dans la référence du premier *on*, la seconde occurrence est plus floue, selon qu'elle représente la classe atemporelle des 'monteurs en images' ou par contraste (*maintenant vs à l'époque*), celle des 'monteurs du passé'. En fait, il vaut mieux considérer que l'interprétation reste en suspens de ce point de vue et n'a même pas à être traitée. Néanmoins, en laissant cette information indéterminée, le locuteur prend le risque d'une inférence erronée de la part de son interlocuteur. Ci-dessous, l'hésitation du locuteur peut être interprétée comme la prise de conscience d'une ambiguïté à ce niveau :

(493) il a suscité pas mal de haine pendant sa pendant son existence mais aussi euh beaucoup de dévouement bon c'est vrai qu'on le **on le on certains le considèrent** un peu comme un boucher (ofrom)

Finalement, le locuteur opte pour l'emploi de *certain*s, pronom [+délocuté] le dissociant clairement de la référence. Un comportement assez proche peut être observé dans l'exemple suivant :

(494) moi je me rappelle dans le bureau à ma grand-mère ils ét- tous ces vieux papiers dont **on a fait un résumé** dont j'ai un résumé encore à la main **qui a été fait** de la famille # dedans qui appartient à ma grand-mère (ofrom)

Le locuteur reformule la proposition en *on* par une structure au passif sans agent<sup>313</sup> (*j'ai un résumé [...] qui a été fait de la famille #*) : la reformulation au passif peut ainsi contribuer à réduire le risque d'une interprétation inclusive<sup>314</sup>. On peut aussi rappeler cet exemple où la locutrice corrige l'usage d'un *on* par un *ils* :

(495) L1 : enfin **nous on a fait on a fait on a ça** aussi à Sevran dans le quartier où c'est très très grand hein euh **ils mettent les barbecues** (mm mm) et tout ça euh c'est bien organisé + avec une fête avant pour les enfants + **on ils + ils ramènent** des poneys tu vois ils <ouais ah oui> ils font faire des tours = (426)

L'exemple suivant illustre la situation inverse, où *ils* se voit rectifié par *on* visiblement parce que la locutrice décide de s'inclure dans la référence :

---

<sup>313</sup> Le syntagme prépositionnel *de la famille #* apparaît en contexte comme le complément du nom *résumé* et non comme un complément d'agent (*≠ par la famille #*)

<sup>314</sup> Rappelons toutefois que l'agent implicite du passif est en théorie également non marqué du point de vue de la délocation. Mais il existe sans doute des facteurs de pertinence qui rendent plus ou moins plausible une inférence dans un sens ou un autre, du genre, l'agent d'un passif sans complément est interprété inclusivement seulement s'il y a des indices clairs en ce sens.

(496) et du coup y a ce problème dans la région de | # | \_ | et euh on avait un terrain à | # | \_ | euh qu'était appartenait à la commune qu'était euh voilà qui ét/ qui | \_ | sur lequel ils voulaient faire quelque chose | \_ | et pis ben le village ayant perdu sa poste | il y a quelques années son petit magasin de justesse | \_ | **ils l'ont ils l'ont on on l'a récupéré** mais euh | \_ | mais par un un mec genre primo qui ouvre pas comme avant (ofrom)

Un fait remarquable avec *on* est la possibilité de voir se succéder en discours des suites d'occurrences non coréférentielles (Berrendonner 1981, Blanche-Benveniste 2003), vivement critiquées par les puristes (Béguelin 2014c)<sup>315</sup>, dont l'alternance entre interprétations inclusive ou exclusive ne pose néanmoins aucun problème :

(497) **on** le renvoie comme ça et **on** nous le renvoie comme ça (au guichet d'une banque, à propos d'un chèque, < Blanche-Benveniste, 2003 : 43)

(498) L'huître n'est pas si malheureuse que nous, **on** l'avale sans qu'elle s'en doute ; mais pour nous, **on** vient nous dire que nous allons être avalés et **on** nous fait toucher au doigt et à l'œil que nous serons digérés éternellement. (Montesquieu, Correspondance, 1754 < Béguelin 2014c)

Nous avons montré *supra* ((410) et (417)) que des enchevêtrements de *ils* non coréférentiels (forcément [+délocuté]) ne constituaient pas non plus des obstacles à l'interprétation, celle-ci étant naturellement orientée par la restriction des prédications successives.

La zone de recouvrement entre *ils* et *on* peut s'observer à travers l'alternance de *on* et *ils* potentiellement coréférentiels, comme en témoignent (461) *supra* et l'exemple suivant :

(499) moi je peux pas j' arriverais pas + parce que c'est encore mon mon Montreuil à moi bien qu'**on<sub>i</sub> nous<sub>j</sub> l'ait pris** [...] c'est plus mon Montreuil à moi + ben le le fait qu'**ils<sub>i</sub> aient fait des HLM** y avait tous les gens qu'**on<sub>k</sub> pouvait imaginer** + mais là à c' moment-là **on<sub>i</sub> nous<sub>j</sub> a pris** moi- pris mon Montreuil parce qu'**on<sub>i</sub> a pris les champs** + alors Montreuil c'était des champs c'était des petits pavillons bon faits de brique et de broc hein parce qu'**on<sub>j</sub> n'avait pas grand-chose** mais c- y avait encore des jéa- des jardins et des champs + que maintenant y a plus rien [...] **ils<sub>i</sub> ont pris tout tout notre coin** [...] **ils<sub>i</sub> ont fait des** dans j'sais plus quelle année mais dans quelles années c'était + tout plein de plein de HLM en bas du cimetière + dans la rue Gaston Leriaux tout ça là + (cfpp)

La locutrice exprime dans cet extrait le sentiment de s'être fait déposséder de son quartier avec la construction de HLM et l'arrivée de nouveaux habitants. Elle en impute ainsi la responsabilité, à travers l'alternance de *on<sub>i</sub>* et *ils<sub>i</sub>*, à un agent tiers. Malgré un doute potentiel sur la coréférence des deux premières occurrences *on* et *ils*, la reformulation à la fin de

<sup>315</sup> Béguelin (2014c) cite quelques extraits parlants dont celui-ci : « Inadvertances gênantes, aussi, que celles consistant à utiliser dans une même phrase des *on* se rapportant à des sujets différents. Il vaut mieux ne pas imiter Théophile Gautier, qui, dans *Jean et Jeannette*, a écrit : *On [= la servante ou le maître d'hôtel, probablement] vint dire à Mme de Kerkeradec qu'elle était servie, et l'on [= les convives] passa dans la salle à manger.* ». (Pierre-Valentin Berthier & Jean-Pierre Colignon, *Le français écorché*, Paris, Belin)

l'extrait *ils ont pris tout tout notre coin* (cf. les autres emplois du verbe *prendre* avec *on*) plaide pour un traitement coréférentiel. On remarque ici encore l'intercalation non problématique d'autres occurrences de *on<sub>j</sub>* non coréférentielles (parfois coréférentielles à *nous<sub>j</sub>*).

Au niveau sémantique, on peut mettre en regard *ils* et *on* sur le trait du nombre. Le clitique *on* apparaît non marqué de ce point de vue, susceptible de représenter au moins un individu (Berrendonner 1981, Creissels *à par.*). Ceci constitue une différence sensible avec *ils*, ce dernier étant marqué en nombre (indiquant l'existence d'un ensemble, lui-même potentiellement composé d'un élément). C'est notamment ce qui explique qu'on ne recoure pas spontanément à *ils* en français lorsqu'on vise à décrire une situation impliquant potentiellement un individu isolé (à moins de vouloir mettre en avant son appartenance à un ensemble), là où *on* se montre approprié :

(500) **On sonne** à la porte. Personne n'est sur scène. **On sonne** de nouveau. (didascalie d'une scène, Reutermann, L., *Un toit pour trois*, p. 71)

Ces contextes appelés *existentiels ancrés* (Cabredo Hofherr 2014) mettent en évidence un autre trait sémantique sur lequel se distinguent *ils* et *on* : si *on* est par nature apte à introduire un individu sous-déterminé, les choses sont plus nuancées pour *ils* : bien que l'instruction d'unifier sa variable se trouve en quelque sorte suspendue, elle n'en devient pas pour autant l'instruction d'introduire une telle variable. L'introduction effective de la variable résulte selon nous de l'abandon de la procédure d'unification plutôt que d'une indication inhérente à *ils*.

Parmi les contextes communs à *ils* et *on*, on retrouve leur combinaison avec des verbes de parole (cf. *supra* §3.2.3.2). La construction verbale indique que les propos proviennent d'une source indirecte non divulguée ou qu'ils ont été obtenus par oui-dire :

(501) donc euh on commence euh à des heures s | \_ | style neuf heures | \_ | et pis euh à partir de là on | \_ | on commence par euh **comme on dit** digitaliser les images | \_ | les images (ofrom)

(502) [à propos de Napoléon] peut-être même serait-il mort euh | \_ | puisqu'il était quand même euh il faisait assez peu de cas | \_ | **on l'a accusé** d'avoir f- de faire assez peu de cas de la vie des autres mais il faisait euh bon aussi assez peu de cas | \_ | de sa propre vie (ofrom)

(503) c'est vraiment la réputation qui poursuit les deux euh les deux lycées | \_ | euh **on dit** que le Jean Piaget est | plus facile que le Denis-de-Rougemont | \_ | donc le Denis-de-Rougemont se targue un peu d'avoir euh | \_ | une euh une influence et une | \_ | un | \_ | une meilleu/ un meilleur rang (ofrom)

On remarque d'emblée que le sujet *on* ne dit rien sur l'éventuelle association du locuteur au point de vue qu'il exprime, tandis qu'avec *ils*, les paroles sont sans conteste imputées à autrui,

même anonyme. Ainsi, le contexte de (501) laisse ouverte la possibilité d'inférer que le locuteur, monteur en images, partage l'usage du jargon signalé. La formule *comme on dit* permet également à celui-là de prévenir toute réaction potentielle à l'emploi de termes spécialisés. L'exemple (502) illustre l'occurrence de *on* dans une clause rectificative visant sans doute à mitiger le fait à asserter (*faire assez peu de cas de la vie des autres*). Le locuteur peut ainsi se désengager de l'assertion en l'attribuant à une source anonyme. On peut rapprocher ce rendement du procédé de « désénonciation » caractéristique du texte scientifique (Ouellet 1984), où *on* contribue à l'élaboration d'un *ethos objectif* (Béguelin 2015), l'auteur s'effaçant discrètement derrière une source scientifique anonyme... La valeur mitigatrice est également bien perceptible dans l'exemple (503), où le verbe introducteur réduit la subjectivité de la comparaison polémique subséquente sans exclure l'adhésion du locuteur, qui conserve toutefois la possibilité de s'en défendre en cas de reproche. On perçoit à travers ces quelques exemples (*cf.* aussi *supra ex.* (491) le rapport subtil et ambigu de *on* à l'énonciation, qui permet d'exploiter de nombreux effets de points de vue<sup>316</sup>.

C'est pour finir également cette flexibilité énonciative qui rend possible une extension potentiellement illimitée de la référence de *on*, invoquée dans les énoncés gnomiques (504) (Creissels à *par.*), par opposition à *ils* qui exclut forcément de sa référence le locuteur (505) (Kleiber 1992b, Cabredo Hofherr 2014 : 11) :

(504) **on** ne vit qu'une fois (*ibid.* : 15)

(505) on se projette tout le temps et tout et bien là-bas | \_ | euh pas **ils** vivent le moment présent (ofrom)

### 3.4.3 Bilan

Dans cette section, nous avons mis en perspective trois procédés dont le point commun est de marquer une forme de sous-détermination de l'agent. Nous avons repéré les zones de recouvrement potentiel pour y observer les spécificités de chacun et pour y relever les facteurs susceptibles de motiver le choix de la mise en œuvre de l'un au détriment des autres. Nous avons tout d'abord constaté que la concurrence est restreinte à l'expression de procès impliquant les rôles d'agent et de patient. Concernant le type d'agent possible, nous proposons un tableau récapitulatif des traits sémantiques oppositifs propres aux trois constructions :

---

<sup>316</sup> Voir à cet égard Maingueneau (2000) et Béguelin (2014c) qui relèvent des effets de brouillage narratif subtils dans des textes littéraires.

**Tableau 2 :** Tableau synthétique des traits oppositifs de *ils*, *on* et  $\emptyset$ <sup>317</sup>

Agent	<i>ILS</i> + V	<i>ON</i> +V	$\emptyset$ (agent implicite de V-PASSIF)
[+ensemble]	+		
[+humain]		+	
[+délocuté]	+		

Cet aperçu met d'emblée en évidence le caractère plus spécifié de *ils* par rapport aux autres. Ce n'est donc véritablement que lorsque le contexte est compatible avec les traits de *ils* qu'il y a concurrence potentielle (format d'ensemble, rôle d'agent et trait délocuté).

A ces différences de spécification sémantique s'ajoutent des éléments liés à la valeur aspectuelle du procès et à la structure informationnelle des énoncés. Le passif est particulièrement apte à marquer l'état et l'aspect accompli résultant d'un procès, tandis que les autres procédés ne peuvent qu'en marquer l'aspect accompli/inaccompli. Dans une structure au passif en outre, la promotion de l'objet en position sujet bouleverse l'orientation des relations et l'importance des rôles (« patient-oriented as opposed to agent-oriented situations », Siewierska 2011 : 68). Cette réorganisation est aussi en jeu dans la manière dont le locuteur délivre l'information.

Nous avons également évoqué chemin faisant quelques éléments de réponse à la distribution apparemment inégale des différents procédés en fonction du genre discursif. A nos yeux, la rareté de *ils* dans les genres orientés vers l'optimisation de l'interprétation s'explique par le fait que *on* et le passif sans complément d'agent sont par définition dédiés à l'introduction d'un agent sous-spécifié, alors que le pronom de 3<sup>e</sup> personne (et par extension, de 6<sup>e</sup> personne) est considéré comme fondamentalement anaphorique dans l'opinion collective. Le défaut de résolution référentielle constaté dans certains cas, dès lors « contraire » à la valeur de base du pointeur, représente sans doute un motif pour ramener l'emploi au style « relâché » (Gundel *et al.* 2000) ou « familier » (Sandfeld 1970). Dans les genres voués à la réalisation d'un produit permanent comme les textes scientifiques, *on* et le passif sont à l'inverse susceptibles d'introduire de manière « légitime » une source dont la sous-détermination fournit paradoxalement un gage d'objectivité tout en n'effaçant pas complètement l'instance énonciatrice.

<sup>317</sup> Par analogie aux analyses oppositives en sémantique componentielle, une case vide indique ici le caractère non marqué de l'expression pour le trait indiqué.

#### 4. Conclusion

Dans ce dernier chapitre, nous nous sommes intéressée à un emploi de la 6<sup>e</sup> personne considéré comme marginal du fait de sa divergence de comportement par rapport aux emplois anaphoriques canoniques. En effet, il se heurte non seulement aux descriptions textuelles de l'anaphore pronominale car il n'est pas dépendant d'un antécédent segmental, mais aussi au critère du maintien de l'attention des approches cognitives. Il n'est donc pas étonnant qu'il soit traité en marge du phénomène de l'anaphore, soit comme un procédé non anaphorique (alors qualifié d'indéfini, d'impersonnel, d'aréférentiel, d'arbitraire, etc.), soit comme un cas très particulier d'anaphore qui présente des contraintes d'emploi strictes (emploi dit *collectif* au sens kleibérien). A l'opposé, une alternative suggérée par Yule (1982) et Béguelin (1993b), sur la base de faits d'oral notamment, consiste à réviser le critère de saillance et de résolution référentielle de la conception courante de l'anaphore.

Etant donné ces divergences d'opinion, nous avons tenté de confronter le phénomène à la notion d'anaphore. Certains cas s'y apparentent assez bien, dans la mesure où *ils* invite à unifier sa variable avec un objet inférable, quoique parfois peu élaboré, à partir d'indices disponibles dans M. Pour d'autres au contraire, l'instruction semble demeurer en suspens, soit par défaut d'indices, soit parce que la construction en *ils* ne s'analyse manifestement pas en termes binaires sujet-prédicat : la prise en compte de l'orientation argumentative du programme discursif auquel elle participe suggère une interprétation en retrait du rôle d'agent par rapport à l'expression du procès lui-même. Dans ces situations, il est peu vraisemblable qu'un traitement inférentiel soit entrepris pour *ils*, bien que la procédure ne soit jamais exclue. Sans l'instruction d'unification, on peut se demander s'il est encore pertinent de parler d'anaphore. Mais au vu de la difficulté, dans bon nombre de cas, de trancher entre une analyse en termes d'anaphore indirecte et de pointeur référentiellement démotivé, nous avons fait l'hypothèse que ce dernier était le résultat d'une réanalyse du pronom, sur la base d'une analogie avec l'emploi de *ça* « asubjectal ». Contrairement à ce dernier, véritable opérateur de diathèse récessif susceptible d'annuler l'agentivité du verbe, *ils* se borne à modifier l'importance des relations et rôles actantiels dans l'expression du procès. Quoi qu'il en soit, l'hypothèse de la réanalyse pourrait expliquer pourquoi il arrive que deux participants d'un échange traitent la même séquence de manière différente.

L'analyse en termes de variable non instanciée rapproche *ils* de l'usage de *on* et du passif, manifestant des similarités fonctionnelles évidentes dans certaines circonstances. Nous avons vu toutefois que l'agent n'était pas spécifié de la même manière selon la forme utilisée et que leur distribution était motivée par des paramètres informationnels, aspectuels, pragmatiques et liés aux genres discursifs. En particulier, la quasi-absence de *ils* référentiellement démotivé dans les écrits les plus travaillés (à moins d'un effet d'oralité) s'explique vraisemblablement par l'écart mal toléré qu'il y manifeste par rapport à sa fonction anaphorique notoire...



# Conclusion générale

Le propos de cette thèse était de dresser un bilan critique sur la problématique de la référence et de proposer un examen des manifestations linguistiques de la sous-détermination pour contribuer à une meilleure compréhension des procédés référentiels et à une description en adéquation avec les usages des locuteurs. En guise de conclusion, nous proposons une synthèse des résultats les plus importants de l'étude et nous esquissons pour finir quelques pistes de recherche et d'application susceptibles de prolonger la réflexion sur le sujet.

## 1. Synthèse des principaux résultats

Dans le premier chapitre, nous avons dégagé les postulats sous-jacents de la sémantique vérifonctionnelle reflétant une conception du monde essentiellement fondée sur l'intuition du linguiste visant à établir une « grammaire du réel » (Apothéloz & Béguelin 1995). Dans cette optique, la langue est perçue comme prédéterminée par la structure du monde, et les expressions référentielles comme des moyens d'identifier correctement les objets qui nous entourent. Ces principes bien implantés dans la tradition sémantique se heurtent cependant aux comportements effectifs des sujets parlants. L'approche que nous avons adoptée s'efforce à l'inverse de dégager ce que les données langagières réelles révèlent sur la manière dont les usagers conçoivent le monde et gèrent le discours. Nous avons ainsi renversé la perspective pour aborder les aspects incontournables de la référence, comme les types de référents, à l'égard desquels les usages dévoilent des procédés d'indifférenciation, ou encore les notions d'anaphore et de deixis : les faits observés mettent en évidence la nécessité de prendre en compte une combinaison de facteurs divers (pragmatiques, informationnels, cognitifs, interactionnels, etc.) dans le recrutement d'une expression référentielle là où les modèles dominants n'invoquent qu'une seule dimension (texte/situation ou maintien référentiel/nouveauté).

Le modèle du discours fribourgeois que nous avons exploité représente un cadre théorique qui s'est avéré adapté à nos besoins. Il accorde aux faits attestés, qu'ils soient courants ou plus rares, oraux ou écrits, une égalité de traitement, sans préjuger de leur acceptabilité. Les référents y sont conçus comme des constructions cognitives (ou *objets-de-discours*) évoluant au fil du discours dans un espace de connaissances partagées (ou *mémoire discursive*) géré sur le vif par les participants de l'échange. Le modèle propose également de considérer les opérations référentielles comme guidées par des stratégies et objectifs communicationnels variés déployés par les locuteurs. C'est donc tout naturellement ce cadre qui a servi de toile de fond à l'étude proposée.

Le deuxième chapitre avait pour but de cerner un élément central des procédés référentiels, incarné par le fonctionnement du pronom de 3<sup>e</sup> personne *il*. En effet, au vu de sa sous-spécification sémantique, il est considéré comme une ressource particulièrement représentative du fonctionnement de l'anaphore référentielle. Les modèles en vigueur voient

généralement dans l'anaphore pronominale le reflet d'une dépendance segmentale et interprétative entre le pronom et un antécédent présent en amont dans le contexte linguistique. C'est d'ailleurs la conception dominante dans les grammaires ou usuels de linguistique, ainsi que dans de nombreux domaines appliqués de la linguistique où le repérage textuel est un critère déterminant. Au vu des faits récalcitrants évidents, parmi lesquels la difficulté de cerner un segment antécédent, une autre approche propose de déplacer le critère sur une dimension cognitive, vouant au pronom un rôle de marqueur de continuité référentielle. Si le pronom fonctionne souvent de la sorte, nous n'y voyons cependant pas un trait encodé dans le sens du pronom ; la saillance d'un objet constitue selon nous seulement une raison parmi d'autres pour lesquelles un locuteur choisit un pronom de 3<sup>e</sup> personne au détriment d'une autre expression. Le cas des pronoms indirects est particulièrement réfractaire à l'analyse en termes de continuité référentielle. Il montre que des paramètres contextuels divers motivent leur emploi selon les intérêts respectifs des protagonistes, comme des procédés de cryptage du référent, de reformatage, d'hétérogénéité énonciative, d'indistinction, etc.

Le troisième chapitre était consacré à la définition de la notion de sous-détermination et à un inventaire des moyens d'expression à travers lesquels elle se manifeste en discours. Nous avons vu que la recherche d'univocité du sens constituait un principe de rigueur non seulement pour les tenants du bon usage, mais aussi, de manière sous-jacente, dans certaines démarches scientifiques cherchant à pallier le flou potentiel de la langue. Cette position a toutefois déjà été remise en cause par des auteurs s'intéressant à des phénomènes d'ambiguïté ou de vague lexical qui se révèlent parties intégrantes de la langue.

La notion de sous-détermination est étroitement liée à celle de pertinence : des raisons de pertinence (calculées en termes de coûts pour les effets recherchés) poussent en effet le locuteur à varier le degré de détermination des objets-de-discours évoquées. Nous avons défini la sous-détermination comme la rencontre d'obstacles dans l'attribution d'une valeur référentielle à une variable. En cause, le défaut d'attributs distinctifs disponibles (dénomination, format, propriétés typantes, etc.) pour identifier l'objet. Nous nous sommes donc intéressée aux désignateurs qui se heurtent à des difficultés d'instanciation de leur variable relevant de trois sortes : i) la suspension temporaire de l'instanciation de la variable (cataphore) ; (ii) l'inférence avortée, car trop aléatoire, d'une valeur ; (iii) la délimitation incertaine d'une classe-objet évoquée en amont dans le discours. A la suite de cette modélisation, nous avons dressé un inventaire des moyens linguistiques qui mettent en œuvre ces différents cas de figure, fondé sur la nature lexicale ou non des désignateurs. A travers cet inventaire et l'analyse des exemples, nous avons distingué deux types de facteurs à l'origine de la sous-détermination : les facteurs d'ordre accidentel (lacune lexicale, ignorance) vs les facteurs d'ordre stratégique (économie de moyens, enjeux informationnels, interactionnels ou encore liés à l'hétérogénéité discursive).

Les quatrième et cinquième chapitres consistaient en deux études empiriques explorant en détail le fonctionnement de deux marqueurs dont la sous-détermination référentielle est

avantageusement exploitée à des fins pragmatiques. Le chapitre IV consacré à *tout ça* complète les nombreux travaux sur *ça* qui ont mis en avant son caractère sous-spécifié et son défaut de trait de catégorisation expliquant sa souplesse référentielle. Le pointeur se montre en effet capable d'opérer des références de diverses natures à des objets dont nous avons illustré l'hétérogénéité. Nous avons proposé de voir l'association de l'opérateur de quantification *tout* à *ça* comme marquant l'exhaustivité d'un ensemble qui ne correspond pas tant à une classe homogène et dénombrable, mais plus généralement à une classe-objet, c'est-à-dire un ensemble flou et hétérogène lié à l'évocation d'un objet. L'examen des emplois de *tout ça* a mis en évidence une valeur résomptive régulièrement à l'œuvre, qui n'exclut cependant pas l'allusion à des éléments laissés implicites.

Le marqueur *tout ça* en position finale de liste est particulièrement productif à l'oral spontané où il commute avec d'autres particules d'extension de liste. Nous avons suggéré que le fonctionnement pragmatique de *tout ça* n'invite pas à l'inférence des éléments implicites restants, mais qu'il présente d'autres rendements, comme la création d'effets d'intersubjectivité, d'atténuation (ou à l'inverse d'emphase), la participation à la gestion narrative d'un récit, l'approximation de paroles rapportées, etc. Nous avons également émis quelques réserves à l'égard de la notion de grammaticalisation, régulièrement invoquée pour l'analyse des « marqueurs discursifs », qui envisage les emplois des formes étudiées sur des parcours évolutifs prédictifs et uni-directionnels. Plutôt que de recourir à une explication déterministe, il nous est apparu plus pertinent d'examiner les circonstances d'occurrence du marqueur en fonction des indices en présence et des objectifs communicationnels à l'œuvre.

Nous nous sommes pour finir intéressée, dans le dernier chapitre, à l'emploi du pronom conjoint *ils* à valeur sous-déterminée, également bien attesté à l'oral spontané. Son emploi, tantôt décrit comme *indéfini*, *impersonnel* ou *collectif*, se distingue du fonctionnement canonique de l'anaphore et l'un des enjeux du chapitre était de situer celui-là par rapport à celle-ci. En fait, nous avons relevé bon nombre de cas qui peuvent être traités en termes d'anaphore indirecte, car certains indices référentiels permettent d'inférer un objet, quoique peu déterminé, pour la variable introduite. D'autres cas y sont plus réfractaires, une inférence apparaissant beaucoup plus aléatoire, si bien qu'il y a tout lieu de penser qu'elle n'est même pas envisagée et que la variable reste intacte. A cet égard, un aspect remarquable de la construction *ils+V* est qu'elle n'a pas à être interprétée en termes de sujet-prédicat, mais plutôt, comme la mise en évidence du procès verbal éclipsant le rôle d'agent, sans toutefois supprimer celui-ci. Entre ces deux analyses, de nombreuses données peuvent être traitées des deux manières : des échanges témoignent de traitements concurrents entre les interlocuteurs pour une même séquence. Ces situations d'ambiguïté nous ont poussée à émettre l'hypothèse d'une réanalyse du pronom *ils* référentiel en un pointeur démotivé, par analogie à l'emploi de *ça* en construction asubjectale, qui a également des répercussions sur l'organisation actantielle de l'énoncé.

L'emploi de *ils* à valeur sous-déterminée a enfin été mis en regard de *on* et du passif dans des contextes comparables afin de dégager les spécificités de chacun. Bien qu'ils permettent tous trois de situer un agent en retrait d'un procès, ils présentent des différences au niveau sémantique qui contribuent à expliquer leur distribution, y compris dans les genres discursifs. En particulier, contrairement à *ils*, *on* et l'agent  $\emptyset$  du passif ne sont marqués ni en nombre, ni sur le trait [+délocuté]. En outre, ils servent par définition à introduire une variable sous-déterminée, contrairement à *ils* dont ce n'est pas la fonction, mais plutôt la conséquence d'une démotivation référentielle. Cet écart de *ils* par rapport à son comportement anaphorique ordinaire représente un élément d'explication plausible sur les jugements de valeur dont il fait l'objet.

## 2. Pistes de recherche

A l'issue de ce travail, nous souhaitons évoquer quelques directions de recherche et points en suspens pour un prolongement de la réflexion en vue de recherches futures. En sélectionnant deux sujets d'étude empirique bien délimités, nous sommes consciente d'avoir laissé de côté une part importante de phénomènes de sous-détermination dont nous avons montré un aperçu dans l'inventaire du chapitre III. En particulier, les pronoms clitiques régimes *le*, *en* et *y* mériteraient qu'on s'attarde plus sérieusement sur leur fonctionnement référentiel, souvent perçu comme un renvoi « propositionnel » ou à un phénomène de suppléance au seul niveau « lexical ». L'étude référentielle de ces pointeurs pourrait bénéficier à notre sens d'apports significatifs par l'examen des faits oraux.

D'autre part, le rapport entre l'indéfinitude et la définitude en français nécessiterait un examen plus approfondi. Nous avons en effet, pour des raisons d'homogénéité, exclu de notre étude les pronoms ou déterminants « indéfinis », alors qu'ils sont fondamentalement marqueurs de sous-détermination. Si notre intérêt portait délibérément sur les désignateurs, pour étudier le « conflit » qu'ils incarnent (la difficulté de trouver une valeur pourtant requise), on constate que les motivations de l'emploi d'expressions indéfinies recouvrent en grande partie celles des pointeurs étudiés. Bien que nous ayons abordé la question avec la mise en perspective de *ils vs on* (et du passif), une recherche minutieuse est à continuer en ce sens, examinant notamment la piste de la nature qualifiée d'« existentielle » de la 6<sup>e</sup> personne (Cabredo Hofherr 2003, 2014). A cet égard, une comparaison avec d'autres marqueurs d'indéfinitude pourrait être prolongée (*quelqu'un*, (*il y a*) *des gens (qui)*, *certain*, etc.).

Il nous semble que les projets d'envergure de constitution de vastes bases de données en français ORFEO<sup>318</sup> et CRFC<sup>319</sup>, bientôt accessibles au public, pourront apporter des éclairages nouveaux non seulement sur les faits de sous-détermination, particulièrement manifestes dans les productions spontanées, mais aussi sur les procédés référentiels en général.

---

<sup>318</sup> <http://www.projet-orfeo.fr>.

<sup>319</sup> Siepmann, Bürgel & Diwersy (2016).

Au niveau de la recherche appliquée en traitement automatique des langues, cette étude esquisse quelques éléments sur la manière d'intégrer un certain nombre de paramètres contextuels, qui ne sont pas pris en compte dans les méthodes de résolution actuelles. On pourrait par exemple repérer les extraits de discours directs, qui, en tant que manifestations de l'hétérogénéité énonciative, représentent un terrain propice à l'émergence de pronoms indirects susceptibles de biaiser les rapports de coréférence. Ensuite, outre l'exclusion déjà en vigueur des *il* « impersonnels » et une attention déjà portée au cas de *ce*, *ça*, il vaudrait la peine de repérer les « aphorismes lexicalisés » avec pronom complément (cf. *supra* Ch. III §5.2.2.3) du type *en rajouter*, *s'y mettre*, *la fermer*, *s'en sortir*, etc. Enfin, pour ce qui concerne plus particulièrement les corpus oraux, un traitement spécifique pourrait être réservé aux *là* « de clôture » (Ch. III, §5.2.4) ainsi qu'au cas de *ils*, en recherchant par exemple dans le voisinage de ce dernier la présence d'un éventuel N collectif, d'un N (ou complément) de lieu. Ces quelques pistes permettraient peut-être d'éviter que la machine se lance dans la résolution des formes en italiques ci-dessous :

(506) le domaine *là* | électronique *là* c'était à la | \_ | # | \_ | où y a le maintenant le | \_ | comment *ça* s'appelle | \_ | le centre euh | \_ | de recherche *et tout ça* | \_ | euh | \_ | *ils* ont démolì l'usine pour construire un autre bâtiment à la place (ofrom) = (225)

Pour terminer, au vu du manque de ressources didactiques pour l'enseignement des procédés de cohérence discursive et de cohésion du texte, il serait opportun de proposer des outils adéquats aux enseignants dans ce domaine, qui puissent sensibiliser les apprenants à la prise en compte de facteurs de divers niveaux qui interviennent dans le choix d'une expression référentielle.

Afin que les élèves prennent conscience des différences de contraintes normatives qui pèsent sur l'emploi des désignateurs en fonction des genres discursifs, il nous paraîtrait utile de les confronter concrètement à des types de production distincts. Une application pertinente en ce sens consisterait à comparer, dans une phase d'observation, les désignateurs provenant d'extraits de genres très différents, par exemple d'ouvrages scientifiques *vs* d'oral spontané. Pour aller plus loin, elle pourrait s'étendre à des extraits narratifs ou de presse, qui mettent en place certaines stratégies particulières (e.g. respectivement les *incipits in medias res* ou l'ajout d'informations inédites via un SN anaphorique). Quoi qu'il en soit, il nous semble indispensable de contextualiser et de formuler les exigences propres à chaque type de produit en vue du déploiement de stratégies de désignation appropriées, en particulier dans l'écriture scolaire.

L'ouvrage de [Reichler-]Béguelin (1990) met ainsi en évidence la nécessité de « cotextualiser » la référence dans ce genre d'écrits, autrement dit d'introduire explicitement les référents dans le savoir partagé, là où des procédés allusifs et implicites passent sans problème à l'oral spontané. D'autres principes comme la proscription des redondances lexicales de même que, dans le domaine des pronoms, la prévention d'ambiguïtés et le respect

des « contraintes d'accord » (genre, nombre) sont de mise, reflétant une « conception idéalisée » de l'anaphore du point de vue de l'écrit qui transpose la nature cognitive des mécanismes en jeu à une dépendance séquentielle (*ibid.* : 79-80). L'ouvrage propose de nombreuses activités de rédaction notamment sur la base de documents authentiques comme la recherche de reformulations para-synonymiques ou hyperonymiques qui tiennent compte de la progression du texte, l'amélioration de « maladroites » (redondances, reprises lexicales problématiques, pronoms présentant des « discordances » morphologiques), le rappel par nominalisation, etc.

Pour aller plus loin et prendre en compte la dimension cognitive et argumentative des procédés de référence, nous proposerions de fournir des ressources pour l'oral également, dans le cadre d'activités du type « exposé » sur un sujet ciblé devant la classe, avec un support adéquat (présentation powerpoint par exemple). L'exploitation des procédés de référence s'ancrerait ainsi fondamentalement dans la situation d'énonciation, tout en participant à la structuration discursive du propos. La présence du public et d'un support à commenter (plan, schéma, tableau) implique non seulement la maîtrise des désignateurs appropriés, mais aussi les dimensions gestuelle, spatiale et visuelle qui sont de fait intrinsèquement liées à ceux-ci.

Enfin, pour ce qui concerne aussi bien l'oral que l'écrit académique, on pourrait imaginer, au-delà des exercices de nominalisation bien connus (*il a perdu son portefeuille ; cette perte...*, *ibid.*), un accent ciblé sur des stratégies résomptives à plus grande échelle, c'est-à-dire sur des étapes entières de raisonnement, domaine en général étudié à travers l'usage des connecteurs logiques. En effet, la maîtrise de rappels par des N « capsules » (*cf. supra* Ch. III §5.1.2), par exemple *hypothèse, aspect, objectif, question, problème, enjeu, solution, argument*, etc.) permet non seulement de développer des compétences métadiscursives de retour synthétique sur le propos, mais aussi de véhiculer un positionnement argumentatif qu'il s'agit de maîtriser en contexte académique.

A travers ces quelques suggestions d'application concrète, nous avons esquissé certaines exploitations possibles des résultats de cette thèse, qui serviront, nous l'espérons, à renouveler la réflexion sur les procédés de référence en général.

# Bibliographie

- Achard-Bayle, G. (2001), *Grammaire des métamorphoses. Référence, identité, changement, fiction*, Bruxelles, De Boeck Supérieur.
- Alonso Ovalle, L. (2000), « Is the 'arbitrary interpretation' a semantic epiphenomenon ? », *University of Massachusetts Occasional Papers* 23, 155-183.
- Amsili, P., Denis, P., & Roussarie, L. (2005), « Anaphores abstraites en français : représentation formelle », *Traitement Automatique des Langues* 46(1), 15-39.
- Apothéloz, D. (1995a), *Rôle et fonctionnement de l'anaphore dans la dynamique textuelle*, Genève, Droz.
- Apothéloz, D. (1995b), « Nominalisations, référents clandestins et anaphores atypiques », *TRANEL* 23, 143-173.
- Apothéloz, D. & Béguelin M.-J. (1995), « Construction de la référence et stratégies de désignation », *TRANEL* 23, 227-271.
- Apothéloz, D. & Béguelin, M.-J. (1999), « Interpretations and functions of demonstratives NPs in indirect anaphora », *Journal of Pragmatics* 31(3), 363-397.
- Apothéloz D. & Pekarek Doehler, S. (2003), « Nouvelles perspectives sur la référence », *Verbum* 25(2), 109-139.
- Ariel, M. (1988), « Referring and accessibility », *Journal of Linguistics* 24, 65-87.
- Ariel, M. (1990), *Accessing noun phrase antecedents*, London/New York, Routledge.
- Ariel, M. (2001), « Accessibility theory : an overview », in T. Sanders *et al.* (éds.), *Text Representation : Linguistic and Psycholinguistic Aspects*, Amsterdam, John Benjamins, 29-87.
- Aristote (1959), *De l'interprétation* (trad. J. Tricot), Paris, Vrin.
- Arnauld, A. & Nicole, P. (1662=1874), *Logique de Port-Royal*, Paris, Hachette.
- Arnavielle T. (1984), « Note à propos de l'article de J. Goudet : "Les marques pronominales de la 3ème personne dans le système du français contemporain" (L'Information grammaticale, n°19, octobre 1983) », *L'Information grammaticale* 22(1), 48.
- Arrivé, M., Gadet, F. & Galmiche, M. (1986), *La grammaire d'aujourd'hui*, Paris, Flammarion.
- Asher, N. (1993), *Reference to abstract objects in discourse*, Dordrecht, Kluwer.
- Auchlin, A. & Simon, A. C. (2004), « Gabarits prosodiques, empathie(s) et attitudes », *Cahiers de l'Institut de Linguistique de Louvain* 30, 181-206.
- Auer, P. (2009), « Context and contextualization », in J. Verschueren & J.-A. Östman (éds.), *Handbook of pragmatics 1995 : Key notions for pragmatics*, Amsterdam/Philadelphia, John Benjamins, 86-101.
- Auger, J. (1995), « Les clitiques pronominaux en français parlé informel : une approche morphologique », *Revue québécoise de linguistique* 24(1), 21-60.
- Auroux, S. (1992), « La catégorie de l'adjectif et les déterminants : l'apport de Beauzée », *Histoire Epistémologie Langage*, 14(1), 159-179.
- Authier-Revuz, J. (2003), « Le fait autonymique : langage, langue, discours. Quelques repères », in J. Authier-Revuz, M. Doury & S. Reboul-Touré (éds.), *Parler des mots. Le fait autonymique en discours*, Paris, Presses Sorbonne Nouvelle, 67-96.
- Avanzi, M., Béguelin, M.-J. & Diémoz, F. (2012-2015), *Présentation du corpus OFROM – corpus oral de français de Suisse romande*, Université de Neuchâtel, <http://www.unine.ch/ofrom>

- Azoulay, A. (1978), « Article défini et relations anaphoriques en français » *Recherches linguistiques françaises* 7, 5-46.
- Bally, Ch. (1932), *Linguistique générale et linguistique française*, Berne, Francke.
- Barbérís, J.-M. (1987), « Deixis spatiale et interaction verbale : un emploi de "là" », *Cahiers de praxématique* 9, 23-48.
- Barbérís, J.-M. (1998), « Identité, ispéité dans la deixis spatiale : *ici* et *là*, deux appréhensions concurrentes de l'espace ? », *L'Information grammaticale* 77, 28-32.
- Barlow, M. (1988), *A situated theory of agreement*, PhD dissertation, Stanford University.
- Bartning, I. (2006), « Une formule bien utile. Le cas de *c'est* en français parlé L1 et L2 », in M. Riegel *et al.* (éds), 2006, *Aux carrefours du sens. Hommages offerts à Georges Kleiber pour son 60e anniversaire*, Louvain, Peeters, 175-190.
- Barton, S. B. & Sanford, A. J. (1993), « A case study of anomaly detection : Shallow semantic processing and cohesion establishment », *Memory and Cognition* 21(4), 477-487.
- Bat-Zev Shyldkrot, H., Adler S. & Asnes, M. (2010), *Approximation et précision* (éds.), *L'Information grammaticale* 125.
- Bat-Zev Shyldkrot, H., Adler S. & Asnes, M. (2014), *Précis et imprécis : études sur l'approximation et la précision*, Paris, Honoré Champion.
- Bat-Zev Shyldkrot, H., Adler S. & Asnes, M. (2016), *Nouveaux regards sur l'approximation et la précision*, Paris, Honoré Champion.
- Beeching, K. (2007), « La co-variation des marqueurs discursifs *bon, c'est-à-dire, enfin, hein, quand même, quoi* et *si vous voulez* : une question d'identité ? », *Langue française* 154, 78-93.
- Béguelin<sup>320</sup>, M.-J. (1997a), « Stratégies référentielles et variation », *Langue française* 115, 101-110.
- Béguelin, M.-J. (1997b), « Anaphores pronominales en contexte d'hétérogénéité énonciative : effets d'(in)cohérence », in L. Tasmowski-De Ryck, W. De Mulder & C. Vetter (éds.), *Relations anaphoriques et (in)cohérence*, Amsterdam, Rodopi 31-54.
- Béguelin, M.-J. (1998), « L'usage des SN démonstratifs dans *les Fables* de La Fontaine », *Langue française* 120(1), 95-109.
- Béguelin, M.-J. (2002), « Construire l'énonciation », in M. Carel (éd.), *Les facettes du dire. Hommage à Oswald Ducrot*, Paris, Kimé, 25-37.
- Béguelin, M.-J. (2012), « Le statut de l'écriture », in D. Druetta (éd.), *Claire Blanche-Benveniste (1935-2010). A l'école de l'oral*, Sylvain-les-Moulins, Éditions Gerflint, 39-54.
- Béguelin, M.-J. (2013), « Note sur l'emploi du pronom de troisième personne chez Gustave Flaubert », in A. Gendre, A. Kamber, L. Petris & C. Skupien Dekens (éds.), *Des mots rayonnants, des mots de lumière. Mélanges de littérature, d'histoire et de linguistique offerts au professeur Philippe Terrier*, Genève, Droz, Recueil de travaux publiés par la Faculté des Lettres et Sciences humaines 58, 297-308.
- Béguelin, M.-J. (2014a), « Ce que nous enseignent les « aphorismes » lexicalisés », in M. Fossard & M.-J. Béguelin, *Nouvelles perspectives sur l'anaphore. Points de vue linguistique, psycholinguistique et acquisitionnel*, Berne, Peter Lang, 137-168.
- Béguelin, M.-J. (2014b), « Deux points de vue sur le changement linguistique », *Langages* 196, 13-36.
- Béguelin, M. J. (2014c), « La concurrence entre *nous* et *on* en français », in M. C. Janner, M. A. Della Costanza, & P. Sutermeister (éds.), *Noi - Nous - Nosostros. Studi Romanzi - Études romanes - Estudios románicos*, Berne, Peter Lang, 73-96.

<sup>320</sup> Voir aussi sous [Reichler-]Béguelin pour les parutions antérieures à 1997.

- Béguelin, M.-J. (2015), « La polyvalence référentielle de *on* : déficience ou atout ? », conférence donnée lors du *Deuxième Colloque International DIA du français actuel*, Université de Neuchâtel, 4-6 novembre 2015.
- Béguelin, M.-J. (2016), « Quand la référence se fait évanescence : *ne pas en démordre, gros comme ça*, et autres locutions à pointeurs démotivés », in L. Sarda, D. Vigier & B. Combettes (éds.), *Connexion et indexation : ces liens qui tissent le texte*, Lyon, ENS Éditions, 99-119.
- Béguelin, M.-J., Corminboeuf, G. & Johnsen, L. A. (2014), « Réflexions sur le statut de la réanalyse », *Verbum* 34(1), 3-16.
- Béguelin, M.-J. & Corminboeuf, G. (à par.), « *Ou comme ça, machin* et autres marqueurs d'indétermination dans les listes », *Discours* 20.
- Bello, A. (1847), *Grammática de la lengua castellana* (8<sup>e</sup> éd. 1970), Buenos Aires, Editorial Sopena Argentina.
- Benveniste, E. (1966), *Problèmes de linguistique générale 1*, Paris, Gallimard.
- Benzitoun, Ch. & Cappeau, P. (2010), « Description sur corpus. Quelques réflexions autour des données et des instruments pour le français (parlé) à travers la description de *cause* et *causer* », in F. Neveu *et al.* (éds.), *Congrès Mondial de Linguistique Française – CMLF 2010*, 1383-98.
- Berrendonner, A. (1981), *Éléments de pragmatique linguistique*, Paris, Éditions de Minuit.
- Berrendonner, A. (1983), « Connecteurs pragmatiques et anaphores », *Cahiers de linguistique française* 5, 215-246.
- Berrendonner, A. (1985), « Faux aléthique et faux helvétique », *Protée* 13(1), 73-80.
- Berrendonner, A. (1990a), « Attracteurs », *Cahiers de Linguistique française* 11, 149-158.
- Berrendonner, A. (1990b), « Pour une macro-syntaxe », *Travaux de linguistique* 21, 25-36.
- Berrendonner, A. (1991), « Variation sur l'impersonnel », in M. Maillard (éd.), *L'impersonnel. Mécanismes linguistiques et fonctionnements littéraires*, Grenoble, Ceditel, 125-132.
- Berrendonner, A. (1993), « Sujets zéro », in S. Karolak & T. Muryn (éds.), *Complétude et incomplétude dans les langues romanes et slaves : actes du VI Colloque international de Linguistique romane et slave*, Cracovie, 29 sept. – 3 oct. 1991, Cracovie, Wydawnictwo Naukowe WSP, 17-45
- Berrendonner, A. (1994), « Anaphores confuses et objets indiscrets » in C. Schnedecker *et al.* (éds.), *L'anaphore associative. Aspects linguistiques, psycholinguistiques et automatiques*, Paris, Klincksieck, 209-230.
- Berrendonner, A. (1995), « Anaphore associative et méréologie », in D. Miéville & D. Vernant (éds.), *Stanislaw Lesniewski aujourd'hui, Recherches sur la philosophie et le langage* 16, 237-255.
- Berrendonner, A. (1997), « Schématisation et topographie imaginaire du discours », in D. Miéville & A. Berrendonner (éds.), *Logique, discours et pensée. Mélanges offerts à J.-B. Grize*, Berne, P.Lang, 219-238.
- Berrendonner, A. (2000), « Que reste-t-il de nos actants ?... Les passifs impersonnels en français », *Cahiers de l'ILSL* 12, 43-53.
- Berrendonner, A. (2002a), « Types », in M. Carel (éd.), *Les facettes du dire, Hommage à O. Ducrot*, Paris, Kimè, 39-54.
- Berrendonner, A. (2002b), « Les deux syntaxes », *Verbum* 24 (1-2), 23-35).
- Berrendonner, A. (2005), « Intensions et extensions », in A. Auchlin *et al.* (éds.), *Structure et discours, mélanges offerts à Eddy Roulet*, Québec, Nota Bene, 151-166.
- Berrendonner, A. (2008), « Dislocation et conjugaison en français contemporain », *Cahiers de praxématique* 48, 85-110.

- Berrendonner, A. (2014), « Anaphores louches et dualités », in M. Fossard & M.-B. Béguelin, *Nouvelles perspectives sur l'anaphore. Points de vue linguistique, psycholinguistique et acquisitionnel*, Berne, Peter Lang, 169-186.
- Berrendonner, A. (2016), « Attentes et insertions parenthétiques », *Langue française* 192, 37-51.
- Berrendonner, A. & [Reichler-]Béguelin, M.-J. (1989), « Décalages : les niveaux de l'analyse linguistique », *Langue française* 81, 99-125.
- Berrendonner, A. & [Reichler-]Béguelin, M.-J. (1995), « Accords associatifs », *Cahiers de praxématique* 24, 21-42.
- Berrendonner, A. & [Reichler-]Béguelin, M.-J. (1996), « De quelques adjectifs à rendement anaphorique : *premier, dernier, autre* », *Studi italiani di linguistica teoretica e applicata* 25(3), 475-502.
- Berthoud, A.-C. (1999), « Recatégorisation des objets au fil du discours », in A. Deschamps & J. Guillemin-Flescher (éds), *Les opérations de détermination : qualification/quantification*, Paris, Ophrys, 185-198.
- Bertrand, R. & Chanet, C. (2005), « Fonctions pragmatiques et prosodie de *enfin* en français spontané », *Revue de sémantique et pragmatique* 17, 41-68.
- Bidaud, S. (2015), *La vicariance en français et dans les langues romanes : (italien, espagnol, portugais)*, Thèse de doctorat, Université de Bourgogne.
- Bilger, M. (1989), « Les réalisations en *et tout (ça)* à l'oral », *Recherches sur le français parlé* 9, 97-109.
- Bilger, M. (2003), « Coordination : entre l'ordre et le désordre », in B. Combettes (éd.), *Ordre et distinction dans la langue*, Paris, Champion, 29-40
- Blanche-Benveniste, C. (1990a), « Un modèle d'analyse syntaxique 'en grille' pour les productions orales », *Anuario de Psicologia* 47, 11-28.
- Blanche-Benveniste, C. (1990b), « Grammaire première et grammaire seconde : l'exemple de *en* », *Recherches sur le français parlé* 10, 51-73.
- Blanche-Benveniste, C. (1987), « Le pronom ON : propositions pour une analyse », in *Mélanges offerts à Maurice Molho*. Vol. III : Linguistique, Paris, Les Cahiers de Fontenay, 15-30.
- Blanche-Benveniste, C. (1997), *Approches de la langue parlée en français*, Paris, Ophrys.
- Blanche-Benveniste, C., (2003), « Le double jeu du pronom ON », in M. Berré, A. Van Slijcke & P. Hadermann (éds.), *La syntaxe raisonnée. Mélanges de linguistique générale et française offerts à Annie Boone à l'occasion de son 60<sup>e</sup> anniversaire*, Bruxelles, De Boeck Supérieur, 41-56.
- Blanche-Benveniste, C. (2011), « Les beautés de l'énumération », in G. Corminboeuf & M.-J. Béguelin (éds.), *Du système linguistique aux actions langagières. Mélanges en l'honneur d'Alain Berrendonner*, Bruxelles, De Boeck-Duculot, 161-172.
- Blanche-Benveniste, C., Bilger, M., Rouget, C. & van den Eynde, K. (1990), *Le français parlé*, Etudes grammaticales, Paris, CNRS-Editions.
- Blanche-Benveniste, C. & Chervel, A. (1966), « Recherches sur le syntagme substantif », *Cahiers de lexicologie* 9, 3-37.
- Blanche-Benveniste, C., Deulofeu, J., Stefanini, J., & Van den Eynde, K. (1984), *Pronom et syntaxe. L'approche pronominale et son application au français*, Paris, SELAF.
- Blanche-Benveniste, C., Rouget, Ch. & Sabio, F., (2002), *Choix de textes de français parlé : 36 extraits*, Paris, Champion.
- Blevins, J. (2003), « Passives and impersonals », *Journal of Linguistics* 39, 473-520.

- Blinkenberg, A. (1950), *Le problème de l'accord en français moderne*. Essai d'une typologie, Copenhague, Munksgaard.
- Bloomfield, L. (1933), *Language*, New York, George Allen & Unwin.
- Boivin, M. (1992), *Le marqueur ça : énonciation et discours*, Mémoire de maîtrise, Université de Québec à Chicoutimi.
- Bonnard, H. (1950), *Grammaire française des lycées et collèges*, Paris, Sudel.
- Boone, A. (1987), « Les constructions "il est linguiste"/"c'est un linguiste" », *Langue française* 75, 94-106.
- Bosch, P. (1983), *Agreement and Anaphora : a Study of the Role of Pronouns in Syntax and Discourse*, London, Academic Press.
- Bosch, P. (1986), « Pronouns Under Control ? A Reply to Liliane Tasmowski and Paul Verluoyten », *Journal of Semantics* 5(1), 65-78.
- Boutet, J. (1984), « La référence à la personne en français parlé : le cas de *on* », *Langage et Société* 38, 19-50.
- Branca-Rosoff, S., Fleury, S., Lefevre F., Pires, M. (2012), *Discours sur la ville. Corpus de Français Parlé Parisien des années 2000 (CFPP2000)* <http://cfpp2000.univ-paris3.fr/>
- Brault, G. (2001), *Les adverbes spatiaux : le cas de là-bas*, Thèse de doctorat, Université de Strasbourg.
- Brinton, L. J. (1996), *Pragmatic Markers in English. Grammaticalization and Discourse Functions*, Berlin, Mouton de Gruyter.
- Brinton, L. J. (2008), *The Comment Clause in English : Syntactic Origins and Pragmatic Development*, Cambridge, Cambridge University Press.
- Brown, G. & Yule, G. (1983), *Discourse Analysis*, Cambridge, Cambridge University Press.
- Brugmann, K. (1904), *Demonstrativpronomina der Indogermanischen Sprachen*, Leipzig, Teubner.
- Brunot, F. (1922), *La Pensée et la langue*, Paris, Masson.
- Brunot, F. & Bruneau, Ch. (1933), *Précis de grammaire historique de la langue française* (2<sup>e</sup> éd.), Paris, Masson.
- Bühler, K. (1934=2009), *Théorie du langage : la fonction représentationnelle* (trad. D. Samain), Marseille, Agone.
- Bürgi, A. (1999), « Le pronom *ça* en français vaudois », *Vox Romanica* 18, 149-171.
- Buroker, J. V. (1993) « The Port-Royal semantics of terms », *Synthese*, 96(3), 455–75.
- Burston, J. & Monville-Burston, M. (1981), « The use of demonstratives and personal pronouns as anaphoric subject of the verb *être* », *Linguisticae Investigationes* 5(2), 231-257.
- Bybee, J. (2010), *Language, usage and cognition*, Cambridge, Cambridge University Press.
- Cabredo Hofherr, P. (2003), « Arbitrary readings of 3pl pronominals », in M. Weisberger, (éd.), *Proceedings of the Conference Sinn un Bedeutung 7 (Sub7)* (Arbeitspapiere des Fachbereichs Sprachwissenschaften 114), Universität Konstanz, 81-94. <http://ling.uni-konstanz.de/pages/conferences/sub7>
- Cabredo Hofherr, P. (2014), *Les pronoms impersonnels humains : syntaxe, sémantique, morphologie*, Mémoire de synthèse en vue d'une habilitation pour diriger des recherches, Université Paris 8.
- Cadiot, P. (1988), « De quoi *ça* parle? A propos de la référence de *ça*, pronom-sujet », *Le français moderne* 56(3-4), 174-192.

- Cadiot, P. (1991), « Le paradoxe de la référence dédoublée du pronom *ça* », in M. Maillard (éd.), *L'impersonnel*, Grenoble, CEDITEL, 151-156.
- Caffi, C. (2007), *Mitigation*, Amsterdam, Elsevier.
- Cappeau, P. & Schnedecker, C. (2014), « *Des gens* bien différents à l'écrit et à l'oral. Hypothèses sur la pronominalisation des SN en *gens* », *Verbum* 36(1), 55-74.
- Cardinaletti, A. & Starke, M. (1994), « The typologie of structural deficiency. On the three grammatical classes », *Working Papers in Linguistics* 4(2), 41-109.
- Carlier, A. (1996), « 'Les gosses *ça* se lève tôt le matin' : l'interprétation générique du syntagme nominal disloqué au moyen de *ce* ou *ça* », *Journal of French Language Studies* 6(2), 133-162.
- Carlson, G. (2006), « Generic, gabituals and iteratives », in K. Brown *The encyclopedia of language and linguistics* (2<sup>e</sup> éd.), Amsterdam, Elsevier.
- Carnap, R. (1947), *Meaning and necessity*, Chicago, University of Chicago Press
- Cinque, G. (1988), « On *si* constructions and the theory of arb », *Linguistic Inquiry* 19, 521-582.
- Chafe, W. L. (1976), « Givenness, contrastiveness, definiteness, subjects, topics, and point of view », in C. Li (éd.), *Subject and Topic*, New York, Academic Press, 25-55.
- Chanet, C. (2001), « 1700 occurrences de la particule *quoi* en français parlé contemporain : approche de la "distribution" et des fonctions en discours », *Marges linguistiques* 2, 56-80.
- Channell, J. (1994), *Vague language*, Oxford, Oxford University Press.
- Charolles, M. (1990), « L'anaphore associative. Problèmes de délimitation », *Verbum* 13(3), 119-148.
- Charolles, M. (1991), « L'anaphore : problèmes de définition et de classification », *Verbum* 14(2-3-4), 203-215.
- Charolles, M. (1994), « Anaphore associative, stéréotype et discours », in C. Schnedecker *et al.* (éds.), *L'anaphore associative. Aspects linguistiques, psycholinguistiques et automatiques*, Paris, Klincksieck, 67-92.
- Charolles, M. (1997), « Identité, changement et référence pronominale », in G. Kleiber, C. Schnedecker & J.-E. Tyvaert (éds.), *La continuité référentielle, Recherches linguistiques* 20, Université de Metz, Klincksieck, 71-95.
- Charolles, M. (2002), *La référence et les expressions référentielles en français*, Paris, Ophrys.
- Charolles, M. (1999), « Associative anaphora and its interpretation », *Journal of Pragmatics* 31, 311-326.
- Charolles, M. & Choi-Jonin, I. (1995), « Les SN indéfinis peuvent-ils fonctionner comme anaphores associatives », *Cahiers de praxématique* 24, 43-68.
- Charolles, M. & Schnedecker, C. (1993), « Coréférence et identité : le problème des référents évolutifs », *Langages* 112, 106-126.
- Chastain, C. (1975), « Reference and context », in K. Gunderson (éd.), *Language, Mind and Knowledge*, Minneapolis, University of Minnesota Press, 194-269.
- Chomsky, N. (1965), *Aspects of the theory of syntax*, Cambridge, MA, MIT Press.
- Chomsky, N. (1981), *Lectures on government and binding*, Dordrecht, Foris.
- Clark, H. (1977), « Bridging », in P. N. Johnson & P. C. Wasow, *Thinking*, Cambridge, Cambridge University Press, 411-420.
- Clark, H. & Haviland, S. (1977), « Comprehension and the given-new contract », in R. Freedie (éd.), *Discourse Production and Comprehension*, Hillsdale, N. J., Lawrence Erlbaum Associates, 1-40.
- Colombat, B. (1993), « La catégorie du nombre en latin et en grec », *Faits de Langues* 2, 29-36.

- Combettes, B. (1986), « Introduction et reprise des éléments d'un texte », *Pratiques* 49, 69-84.
- Combettes, B. (1983), *Pour une grammaire textuelle. La progression thématique*. Bruxelles-Paris, Duculot.
- Condoravdi, C. (1989), « Indefinite and generic pronouns », *Proceedings of WCCFL* vol. 8, 71-84.
- Consten, M., Knees, M., & Schwarz-Friesel, M. (2007), « The function of complex anaphors in texts. Evidence from corpus studies and ontological considerations », in M. Schwarz-Friesel *et al* (éds.), *Anaphors in text. Cognitive, formal and applied approaches to anaphoric reference*, Amsterdam/Philadelphia, John Benjamins Publishing Company, p. 81-102.
- Conte, M.-E. (1996), « Anaphoric encapsulation », in W. De Mulder & L. Tasmowsky (éds.), *Coherence and Anaphora, Belgian Journal of Linguistics* 10, 1-10.
- Corbett, G. (1991), *Gender*, Cambridge, Cambridge University Press.
- Corbett, G. (2000), *Number*, Cambridge, Cambridge University Press.
- Corbett, G. (2006), *Agreement*, Cambridge, Cambridge University Press.
- Corblin, F. (1983), « Les désignateurs dans les romans », *Poétique* 54, 199-211.
- Corblin, F. (1985), « Remarques sur la notion d'anaphore », *Revue québécoise de linguistique* 15(1), 173-195.
- Corblin, F. (1987a), Indéfini, défini et démonstratif. Constructions linguistiques de la référence, Genève, Droz.
- Corblin, F. (1987b), « Ceci et cela comme formes à contenu indistinct », *Langue française* 75, 75-93.
- Corblin, F. (1987c), « Les chaînes de référence naturelles », *T.A. Information* 1, 5-21.
- Corblin, F. (1991), « Sujet impersonnel et sujet indistinct : IL et ça », in M. Maillard (éd.), *L'impersonnel. Mécanismes linguistiques et fonctionnements littéraires*, Grenoble, Ceditel, 139-150.
- Corblin, F. (1995), *Les formes de reprise dans le discours. Anaphores et chaînes de référence*, Rennes, Presses universitaires de Rennes.
- Corblin, F. (2005), « Les chaînes de la conversation et les autres », in J.-M. Gouvard (éd.), *De la langue au style*, Presses universitaires de Lyon, 233-254.
- Corminboeuf, G. (2011), « Dualités », in M.-J. Béguelin & G. Corminboeuf, *Du système linguistique aux actions langagières. Mélanges en l'honneur d'Alain Berrendonner*, Bruxelles, De Boeck-Duculot, 473-486.
- Corminboeuf, G. (2014), « L'identification des relations de discours implicites : le cas de l'adversation », in *Actes du CMLF, 4<sup>ème</sup> Congrès Mondial de Linguistique Française*, publié par EDP Sciences ([www.linguistiquefrancaise.org](http://www.linguistiquefrancaise.org)), 2367-2382. <http://dx.doi.org/10.1051/shsconf/20140801049>
- Corminboeuf, G. (2016), « "Comme ça", marqueur d'approximation », in F. Lefevre & G. Dostie, *A l'articulation du lexique, de la grammaire et du discours : marqueurs grammaticaux et marqueurs discursifs*, Paris, Champion, 1-20.
- Cornish, F. (1991), « Non-discrete reference, discourse construction, and the French neuter clitic pronouns », *Journal of French Language Studies* 1(2), 123-138.
- Cornish, F. (1999), *Anaphora, discourse, and understanding : Evidence from English and French*, Oxford, Clarendon.
- Cornish, F. (2001), « L'anaphore pronominale indirecte : une question de focus », in W. de Mulder, C. Vet & C. Veters (éds.), *L'Anaphore pronominale et nominale. Etudes pragma-sémantiques*, New York/ Amsterdam, Rodopi, 1-25.
- Cornish, F. (2005), « Degrees of indirectness : two types of implicit referents and their retrieval via unaccented pronouns », in R. Mitkov, T. McEnery & A. Branco (éds.), *Anaphora processing:*

*linguistic, cognitive and computational modelling*, Amsterdam-Philadelphie, John Benjamins, 199-220.

Cornish, F. (2009), « *Text and discourse as context* : Discourse anaphora and the FDG contextual component », in E. Keizer & G. Wanders (éds.), *Working papers in functional discourse grammar (WP-FDG-82)*, *The London Papers I*, 97-115.

Cornish, F. (2010a), « Indexicaux, discours et mémoire discursive : ce que les premiers révèlent du second et de la troisième », *Linx* 62-63, 111-133.

Cornish, F. (2010b), « Anaphora : Text-based or discourse-dependent ? Functionalist vs. formalist accounts », *Functions of Language* 17(2), 207-241.

Cornish, F. (2017), « Indexicals and L2 learners' metadiscursive awareness. », in R. Breeze & C. Sancho Guinda (éds.), *Essential competencies for English-medium university teaching*, Switzerland, Springer, 65-82.

Cornish, F. (à par.), « Indexicaux : orientation cognitive, point de vue, et création de discours », in M. Jadir, *Linguistique et discours*, Berne, Peter Lang.

Cornish, F., Garnham, A., Cowles, H.W., Fossard, M. & André V. (2005), « Indirect anaphora in English and French : a cross-linguistic study of pronoun resolution », *Journal of Memory and language* 52, 363-376.

Cornish, F. & Salazar Orvig (2016), « A critical look at the notion 'pro-form'. Evidence from indexical markers, spoken discourse and (French) child language », *Language Sciences* 54, 58-76.

Creissels, D. (1991), « Approche des constructions asubjectales », improprement désignées comme "impersonnelles" », in M. Maillard (éd.), *L'impersonnel. Mécanismes linguistiques et fonctionnements littéraires*, Grenoble, Ceditel, 47-57.

Creissels, D. (1995), *Éléments de syntaxe générale*, Paris, P.U.F.

Creissels, D. (2006), *Syntaxe générale, une introduction typologique 2 : la phrase*, Paris, Hermes Sciences-Lavoisier.

Creissels, D. (à par.), « Impersonal pronouns and coreference : the case of French *on* », in S. Manninen, K. Hietaam, E. Keiser & V. Vihman (éds.), *Passives and impersonals in european languages*.

Croft, W. (1991), *Syntactic categories and grammatical relations : The cognitive organization of information*, Chicago, University of Chicago Press.

Croft, W. (2013), « Agreement as anaphora, anaphora as coreference », in D. Bakker & M. Haspelmath (éds.), *Language across boundaries: Studies in memory of Anna Siewierska*, Berlin, De Gruyter Mouton, 107-129.

Cruse, D. (1986), *Lexical semantics*, Cambridge, Cambridge University Press.

Crystal, D. & Davy, D. (1975), *Advanced conversational English*, London, Longman.

Culbertson, J. (2010), « Convergent evidence for categorical change in French : from subject clitic to agreement marker », *Language* 86, 85-132.

Curat, H. (1988), « Pluriel interne et système morphologique du nombre en français », *Revue québécoise de linguistique* 17(1), 29-52.

Damourette, E. & Pichon, J. (1911-1946), *Des mots à la pensée. Essai de grammaire de la langue française*, 7 tomes, Paris, Editions d'Artrey.

Darmesteter, A. (1877), *De la création actuelle de mots nouveaux dans la langue française et des lois qui la régissent*, Paris, F. Vieweg/A. Franck.

Davidson, D., (1967), « The logical form of action sentences », in N. Resher (éd.), *The Logic of Decision and Action*, Pittsburgh, Pittsburgh University Press 81-95.

- Delais-Roussarie, E. *et al.* (2015) « Developing a ToBI system for French », in S. Frota & P. Prieto (éds.), *Intonational variation in Romance*, Oxford, Oxford University Press, 63-100.
- Delattre, P. (1965), « Les dix intonations de base du français », *French Review* 40, 1-14.
- Delic, (2004), « Présentation du Corpus de référence du français parlé », *Recherches sur le français parlé* 18, 11-42.
- Demol, A. (2010), *Les pronoms anaphoriques il et celui-ci*, Bruxelles, De Boeck-Duculot.
- Demonet, M.-L. (2002), « À plaisir ». *Sémiotique et scepticisme chez Montaigne*, Orléans, Paradigme.
- De Mulder, W. (2000), « Démonstratifs et accessibilité », *Verbum* 1, 103-125.
- Dik, S.C. (1997), *The Theory of functional grammar : Part 1 : The structure of the clause*, in K. Hengeveld, (éd.), Amsterdam, Walter de Gruyter.
- De Carvalho, P. (1993), « Aspects du nombre dans les langues anciennes », *Faits de langue* 2, 97-110.
- De Cat, C. (2007), *French dislocation, interpretation, syntax, acquisition*, Oxford, Oxford University Press.
- Demol, A. (2010), *Les pronoms anaphoriques il et celui-ci*, Bruxelles, De Boeck.
- De Mulder, W. (1998), « Du sens des démonstratifs à la construction d'univers », *Langue française* 120(1), 21-32.
- Denturck, E. (2008), *Etude des marqueurs discursifs. L'exemple de quoi*, Mémoire de Master, Université de Gand.
- Desclées, J.-P. (1986), « Implication entre concepts ; la notion de typicalité », *Travaux de linguistique et de littérature* 24(1), 179-202.
- Désoyer, A., Landragin, F., Tellier, I., Lefeuvre, A. & Antoine, J.-Y. (2015), « Les coréférences à l'oral : une expérience d'apprentissage automatique sur le corpus ANCOR », *Traitement Automatique des Langues, ATALA, Traitement automatique du langage parlé* 55 (2), 97-121.
- Dines, E. (1980), « Variation in discourse – and stuff like that », *Language in Society* 1, 13-31.
- Donnellan, K. S. (1966), « Reference and definite descriptions », *The philosophical review* 75(3), 281-304.
- Dostie, G. (2004), *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*, Bruxelles, De Boeck Duculot.
- Dostie, G. & Push, C. (2007), « Présentation. Les marqueurs discursifs », *Langue française* 154, 3-12.
- Dubois, J. (1965), *Grammaire structurale du français : nom et pronom*, Paris, Librairie Larousse.
- Dubois, S. (1992), « Extension particles, etc. », *Language Variation and Change* 4, 179-203.
- Dubois, S. (1993), « Les particules d'extension dans le discours : analyse de la distribution des formes et patati et patata », *Revue québécoise de linguistique théorique et appliquée* 11(1- 4), 21-47.
- Ducrot, O. (1972), *Dire et ne pas dire : principes de sémantique linguistique*, Paris, Hermann.
- Ducrot, O. & Todorov, T. (1979), *Dictionnaire encyclopédique des sciences du langage*, Paris, Seuil.
- Durand, J., Laks, B. & Lyche, Ch. (2002), « La phonologie du français contemporain : usages, variétés et structure », in C. Pusch & W. Raible (éds), *Romanistische Korpuslinguistik- Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*, Tübingen, Gunter Narr Verlag, 93-106.
- Durand, J., Laks, B. & Lyche, Ch. (2009), « Le projet PFC: une source de données primaires structurées », in Durand J. *et al.* (éds), *Phonologie, variation et accents du français*, Paris, Hermès, 19-61.

- Ehlich, K. (1982), « Anaphora and deixis : same, similar, or different ? », in R. J. Jarvella & W. Klein (éds.), *Speech, place and action : Studies in deixis and related topics*, Chichester, John Wiley, 315-338.
- Elmiger, D. (2008), *La féminisation de la langue en français et en allemand. Querelle entre spécialistes et réception par le grand public*, Paris, Honoré Champion.
- Erickson, F. (1992), « They know all the lines: rhythmic organization and contextualization in a conversational listing routine », in P. Auer & A. Di Luzio (éds.), *The Contextualization of language*, Amsterdam/New York, John Benjamins, 365-397.
- Erikson, T. A. & Mattson, M. E. (1981), « From words to meaning : A semantic illusion », *Journal of Verbal Learning and Verbal Behavior* 19, 583-596.
- Erkù, F. & Gundel, J. (1987), « The pragmatics of indirect anaphors », in J. Verschueren & M. Bertuccelli-Papi (éds.), *The Pragmatic perspective : Selected papers from the 1985 International Pragmatics Conference*, Amsterdam, John Benjamins, 533-545.
- Fauconnier, G. (1984), *Espaces mentaux*, Paris, Éd. de Minuit.
- Fernández, J. M. M. (1994), *Les particules énonciatives dans la construction du discours*, Paris, Presses Universitaires de France.
- Ferré, G. (2011), « Analyse multimodale des particules d'extension « et tout ça, etc. » en français », in H.-Y. Yoo & E. Elisabeth-Roussarie, *Interface Discours Prosodie (IDP09)*, Paris, 157-171.
- Ferreira, F., Bailey, K.G.D. & Ferraro, V. (2002), « Good-enough representations in language comprehension », *Current directions in psychological science* 11(1), 11-15
- Filik, R., Sanford, A. J. & Leuthold, H. (2008), « Processing pronouns without antecedents: evidence from event-related brain potentials », *Journal of Cognitive Neuroscience* 20(7), 1315-1326.
- Fillenbaum, S. (1974), « Pragmatic normalization : Further results for some conjunctive and disjunctive sentences », *Journal of Experimental Psychology* 102(4), 574-578.
- Fillmore, Ch. J. (1968), « The case for case », in E. Bach & R. Harms (éds.), *Universals in linguistic theory*, New York, Holt, Rinehart and Winston, 1-88.
- Fillmore, Ch. J. (1971), « Some problems for Case Grammar », in R.J. O'Brien (éd.), *22th Annual round table. Linguistics: developments of the sixties – viewpoints of the seventies*, Washington D.C., Georgetown University Press, 35-56.
- Flaux, N. & De Mulder, W. (1997), « Indéfinis et référence : introduction », *Langue française* 116, 3-7.
- Flaux, N., Glatiny M. & Samain D. (éds.) (1996), *Les noms abstraits. Histoires et théories*, Villeneuve d'Ascq, Presses universitaires du Septentrion.
- Flaux, N. & Van de Velde, D. (2002), *Les noms en français. Esquisse de classement*, Paris, Ophrys.
- Flowerdew, J. (2003), « Signalling nouns in discourse », *English for specific purposes* 22, 329-346.
- Flowerdew, J. & Forest, R. W. (2015), *Signalling nouns in English: A corpus-based discourse approach*, Cambridge, Cambridge University Press.
- Fox, B. (1987), *Discourse structure and anaphora*, Cambridge, Cambridge University Press.
- Fradin, B. (1984), « Anaphorisation et stéréotypes nominaux », *Lingua* 64, 325-369.
- Francis, G. (1994), « Labelling discourse: An aspect of nominal-group lexical cohesion », in M. Coulthard (éd.), *Advances in written text analysis*, London-New York, Routledge, 83-101.
- François, J. (1984), « Analyse énonciative des équivalents allemands du pronom indéfini *on* », in G. Kleiber (éd.), *Recherches en pragma-sémantique*, Paris, Klincksieck, 37-73.

François, J. & Broschart, J. (1994), « La mise en ordre des relations actanciennes : les conditions d'accès des rôles actantiels aux fonctions de sujet et d'objet », *Langages* 113, 7-44.

Frantext *Base textuelle*, ATILF - CNRS & Université de Lorraine, <http://www.frantext.fr>.

Fraser, T. & Joly, A. (1980), « Le système de la deixis. Endophore, cohésion discursive en anglais », *Modèles linguistiques* 2(2), 22-51.

Fraurud, K. (1992), *Processing noun phrases in natural discourse*, PhD dissertation, Stockholm University.

Frazier, L. & Rayner, K. (1990), « Taking on semantic commitments : Processing multiple meanings vs. multiple senses », *Journal of Memory and Language* 29(2), 181-200.

Frazier, L., Patch, J. & Rayner, K. (1999), « Taking on semantic commitments, II : collective versus distributive readings », *Cognition* 70, 87-104.

Frege, G. (1892 = 1971), *Écrits logiques et philosophiques* (trad. C. Imbert), Paris, Éditions du Seuil.

Fuchs, C. & Le Goffic, P. (1983), « Ambiguïté, paraphrase et interprétation », *Modèles Linguistiques* 5, 109-136.

Fuchs, C. (1996), *Les ambiguïtés du français*, Paris, Ophrys.

Furukawa, N. (1977), *Le nombre grammatical en français contemporain*, Tokyo, France Tosho.

Galmiche, M. (1983), « Les ambiguïtés référentielles ou les pièges de la référence », *Langue française* 57(1), 60-86.

Gardes-Tamine, J. (1998), *La grammaire 2. Syntaxe*, Paris, Armand Colin.

Garnham, A. & Oakhill, J. (1993), « Modèles mentaux et compréhension du langage, in M.-F. Ehrlich *et al.* (éds.), *Les modèles mentaux. Approche cognitive des représentations*, Masson, Paris, 23-46.

Gary-Prieur, M.-N. (1994), *Grammaire du nom propre*, Paris, PUF.

Gary-Prieur, M. N. (2011), *Les déterminants du français*, Paris, Ophrys.

Gaatone, D. (1991), « Il y a impersonnel et impersonnel : sur les contraintes et les finalités de la construction impersonnelle », in Maillard M. (éd.), *L'impersonnel. Mécanismes linguistiques et fonctionnements littéraires*, Grenoble, Ceditel, 133-138.

Givón, T. (1976), « Topic, pronoun and grammatical agreement », in Ch. Li (éd.), *Subject and topic*, New York, Academic Press, 151-188.

Givón, T. (1983), « Topic continuity in discourse: an introduction », in T. Givón (éd.), *Topic continuity in discourse : a quantitative cross-language study*, Amsterdam/Philadelphia, John Benjamins, 5-41.

Givón, T. (1984), *Syntax : A functional-typological introduction*, Amsterdam, John Benjamins.

Givón, T. (1992), « The grammar of referential coherence as mental processing instructions », *Linguistics* 30, 5-55.

Gollut, J.-D. & Zufferey, J. (2000), *Construire un monde. Les phrases initiales de La Comédie humaine*, Paris, Delachaux et Niestlé.

Goudet, J. (1983), « Les marques pronominales de la troisième personne dans le système du français contemporain », *L'information grammaticale* 19, 11-15.

Greene, S.B., Gerrig, R.J., McKoon G. & Ratcliff, R. (1994), « Unheralded pronouns and management by common ground », *Journal of Memory and Language* 33, 511-526.

Greimas, A. J. (1966), *Sémantique structurale*, Paris, Larousse.

Grevisse, M. & Goosse, A. (2011), *Le Bon usage* (14<sup>e</sup> éd.), Bruxelles, De Boeck-Duculot.

- Grevisse, M. & Lenoble Pinson, M. (2009), *Le français correct : Guide pratique des difficultés*, Bruxelles, De Boeck.
- Grice, H.P. (1975), « Logic & conversation », in P. Cole & J. L. Morgan, *Syntax and Semantics 3: Speech Acts*, New York, Academic Press, 41-58.
- Grize, J. B. (1974), « Argumentation, schématisation et logique naturelle », *Revue européenne des sciences sociales* 12(32), 183-200.
- Grize, J.-B. (1982), *De la logique à l'argumentation*, Genève et Paris, Droz.
- Grize, J.-B. (1990), *Logique et langage*, Paris, Ophrys.
- Grize, J.-B. (1993), « Logique naturelle et représentations sociales », *Textes sur les représentations sociales* 2(3), 151-159.
- Grize, J.-B. (1996), *Logique naturelle et communications*. Paris, Presses Universitaires de France.
- Grobet, A. (2002), *L'identification des topiques dans les dialogues*, Bruxelles, Duculot.
- Groupe de Fribourg (2012), *Grammaire de la période*, Berne, Peter Lang.
- Guerin, E. (2006), *Introduction de la notion de variation situatiolectale dans la grammaire scolaire par la caractérisation de deux opérateurs pragmatiques : on et ça*, Thèse de Doctorat, Paris X Nanterre.
- Guerin, E. (2007), « "Moi je comprends pas et ça ça m'intrigue". Ce que peut nous apprendre la double occurrence de ça dans les énoncés oraux présentant un phénomène de dislocation », *Linx*, 57, 27-36.
- Guilbaud, G. (1977), « Mathematics and approximation », in H. Athen & H. Kunle (éds.), *Proceedings of the Third International Congress on Mathematics Education*, Karlsruhe, 125-134.
- Guillaume, G. (1919), *Le problème de l'article et sa solution dans la langue française*, Paris, Hachette.
- Guillaume, G. (1964), *Langage et science du langage*, Paris, Nizet.
- Guillaume, G. (1985), *Leçons de linguistique 1945-46(VI)*, R. Valin, W.H. Hirtle & A. Joly (éds.), *Grammaire particulière du français et grammaire générale I*, Lille, Presses universitaire de Lille ; Québec, Presses de l'Université Laval.
- Guillet, A. (1978), « Remarques sur la notion de pluriel », *Cahiers de linguistique* 8, 205-219.
- Guillot, C. (2006), « Démonstratif et deixis discursive : analyse comparée d'un corpus écrit de français médiéval et d'un corpus oral de français contemporain », *Langue française* 152(4), 56-69.
- Gundel, J. K. & Fretheim, T. (2004), « Topic and focus », in L. R. Horn & G. Ward (éds.), *The handbook of pragmatics*, Oxford, Blackwell, 175-196.
- Gundel, J. K., Hedberg, N & Zacharski, R. (1993), « Cognitive status and the form of referring expressions in discourse », *Language* 62(2), 274-307.
- Gundel, J. K., Hedberg, N. & Zacharski, R. (2000), « Statut cognitif et formes des anaphoriques indirects », *Verbum* 22(1), 79-102.
- Gundel, J.K., Hegarty, M. & Borthen, K. (2003), « Cognitive status, information structure, and pronominal reference to clausally introduced entities », *Journal of Logic, Language and Information* 12, 281-199.
- Haase, A. (1898), *Syntaxe française du XVIIe siècle* (trad. M. Obert), Paris, A. Picard et fils.
- Hagège, C. (1978), « Du thème au thème en passant par le sujet : une théorie cyclique », *La Linguistique*, 14(2), 3-38.
- Halliday, M. A. K. & Hasan, R. (1976), *Cohesion in English*, London, Longman.

- Halmøy, O. (2006), « De chose et d'autres : la série *truc, machin, bidule*. Eléments de distribution », in M. Riegel, C. Schnedecker, P. Swiggers, I. Tamba, *Aux Carrefours du sens : hommages offerts à Georges Kleiber pour son 60e anniversaire*, Louvain, Peeters, 511-529.
- Harris, Z. (1976), *Notes du cours de syntaxe*, trad. M. Gross, Paris, Seuil.
- Haspelmath, M. (1997), *Indefinites pronouns*, Oxford, Oxford University Press.
- Haviland, S. E. & Clark, H. H. (1974), « What's new ? Acquiring new information as a process in comprehension », *Journal of Verbal Learning and Verbal Behavior* 13, 512-521.
- Hawkins, J. A. (1977), « The pragmatics of definiteness, Part I », *Linguistische Berichte* 47: 1-27.
- Hawkins, J. A. (1978), *Definiteness and indefiniteness: a study in reference and grammaticality predictions*, London, Croom Helm.
- Hawkins, J. A. (1984), « A note on referent identifiability and co-presence », *Journal of Pragmatics* 8: 649-659.
- Heim, I. (1983), « File change semantics and the familiarity theory of definiteness », in R. Bäuerle, Ch. Schwarze & A.V. Stechow (éds.), *Meaning, use, and interpretation of Language*, Berlin, de Gruyter, 164-189
- Hjemslev, L., (1968), *Prolégomènes à une théorie du langage*, Paris, Minuit.
- Himmelman, N. (1996), « Demonstratives in narrative discourse: a taxonomy of universal uses », in B.A. Fox, (éd.), *Studies in anaphora*, Amsterdam/Philadelphia, John Benjamins, 205-254.
- Huyghe, R. (2015), « Les typologies nominales : présentation », *Langue française* 185, 5-27.
- Jackendoff, R. (1990), *Semantic structures*, Cambridge, MIT-Press.
- Jakobson, R. (1939), « Signe zéro », in *Mélanges de linguistique offerts à Charles Bally*, Genève, Georg et Cie, 143-152.
- Jaeggli, O. (1982), *Topics in Romance syntax*, Dordrecht, Foris.
- Jaeggli, O. (1986), « Arbitrary plural pronominals », *Natural Language & Linguistic Theory* 4(1), 43-76.
- Jakobson, R. (1971), « Signe zéro », *Selected Writings II*, La Haye, Walter de Gruyter (1ère éd. 1939).
- Jakobson, R. (1971), « The gender pattern of Russian », *Selected Writings II*, La Haye, Walter de Gruyter (1ère édition 1959).
- Jefferson, G. (1990), « List construction as a task an resource », in G. Psathas, (éd.), *Interaction competence*, Lanham, University Press of America, 63-72.
- Jespersen, O. (1924), *Mass-words: The philosophy of grammar*, London, Allen and Unwin.
- Johnsen, L. A. (2008), « Procédés référentiels dans les parenthèses », *Verbum* 30(1), 85-102.
- Johnsen, L. A. (2010), « Les pronoms "neutres" et leur référence à des procès en français parlé », *Linx* 62-63, 153-178.
- Johnsen, L. A. (2011), « Un éclairage sur le fonctionnement référentiel de tout ça en fin de liste », in G. Corminboeuf & M.-J. Béguelin (éds.), *Du système linguistique aux actions langagières. Mélanges en l'honneur d'Alain Berrendonner*, Bruxelles, De Boeck-Duculot, 487-505.
- Johnsen, L. A. (2013), « Anaphore pronominale indirecte en micro- et macro-syntaxe », in M. Barbazan (éd.), *Énonciation et texte au cœur de la grammaire. Regards croisés : neuro/psycholinguistique, psychologie cognitive, linguistique, didactique*, Rennes, Presses universitaires de Rennes.
- Johnsen, L.A. (2014a), « Parenthèses ana-cataphoriques et figures du discours », in C. Barbet (éd.), *Linguistique et stylistique des figures*, Bruxelles, Peter Lang, 123-145.

- Johnsen, L. A. (2014b), « L'emploi du *ils* « collectif » en français : entre interprétation anaphorique et interprétation existentielle », in Fossard M. & Béguelin M.-J. (éds.), *Nouvelles perspectives sur l'anaphore : points de vue linguistique, psycholinguistique et interactionnel*, Actes du colloque des 4-5 avril 2012, Université de Neuchâtel, Berne, Peter Lang, 99-136.
- Johnsen, L.A. & Avanzi, M. (à par.), « Examen des contours prosodiques de listes en français parlé ».
- Johnson-Laird, P. N. (1983), *Mental models : Towards a cognitive science of language, inference, and consciousness*, Cambridge, Cambridge University Press.
- Johnson-Laird, P. N. (2010), « Mental models and human reasoning », *Proceedings of the National Academy of Sciences* 107(43), 18243-18250.
- Johnson-Laird, P. N. & Garnham, A. (1980), « Descriptions and discourse models », *Linguistics & Philosophy* 3, 371-393.
- Jonasson, K. (1991), « Les noms propres métaphoriques : construction et interprétation », *Langue française* 92, 64-82.
- Jonasson, K. (1994), *Le nom propre. Constructions et interprétations*, Bruxelles, De Boeck.
- Jucker, A. H., Smith, S. W. & Lüdge, T. (2003), « Interactive aspects of vagueness in conversation », *Journal of Pragmatics* 35, 1737-1769.
- Kahane, S. & Pietrandrea, P. (2012), « La typologie des entassements en français », in F. Neveu *et al.*, *Congrès Mondial de Linguistique Française – CMLF 2012*, EDP Science, 1809-1828.
- Kaltenböck, G., Mihatsch, W. & Schneider, S. (2010), *New approaches to hedging*, Bingley, Emerald.
- Kamp, H. (1981), « A theory of truth and semantic representation », in J. Groenendijk, T.M.V Janssen & M. Stokhof, *Formal Methods in the Study of Language, Proceedings of the Third Amsterdam Colloquium*, Amsterdam, 277-322.
- Kamp, H. & Reyle, U. (1993), *From discourse to logic*, Dordrecht, Kluwer Academic Publisher.
- Karttunen, L. (1976), « Discourse referents », in J. D. McCawley, *Syntax and Semantics 7: Notes from the Linguistic Underground*, New York, Academic Press, 363-386.
- Kempson, R. (1986), « Definite NPs and context-dependence : a unified theory of anaphora », in T. Hyers *et al.* (éds.), *Reasoning and Discourse Processes*, London, Academic Press, 209-239.
- Kerbrat-Orecchioni, C. (1980), *L'énonciation : de la subjectivité dans le langage*, Paris, A. Colin.
- Kerbrat-Orecchioni, C. (2005), *Le discours en interaction*, Paris, Armand Colin.
- Kerbrat-Orecchioni, C. (2010), « Les formes nominales d'adresse dans les conversations familiales », in C. Kerbrat-Orecchioni (éd.) *S'adresser à autrui: les formes nominales d'adresse en français*, Université de Savoie, coll. "Langages", 7-30.
- Kęsik, M. (1989), *La cataphore*, Paris, PUF.
- Kibrik, A. A. (2011), *Reference in discourse*, Oxford, Oxford University Press.
- Kitagawa Ch. & Lehrer A. (1990), « Impersonal uses of personal pronouns », *Journal of Pragmatics* 14, 739-759.
- Kleiber, G. (1981), *Problèmes de référence : descriptions définies et noms propres*, Paris, Klincksieck.
- Kleiber G. (1984), « Sur la sémantique des descriptions démonstratives », *Linguisticae Investigationes* VIII (1), 63-85.
- Kleiber, G. (1987), « Mais à quoi sert donc le mot *chose* ? Une situation paradoxale », *Langue française*, 73,109-128.
- Kleiber, G. (1990a), « Quand il n'a pas d'antécédent », *Langages* 97, 24-50.

- Kleiber, G. (1990b), « Sur l'anaphore associative : article défini et adjectif démonstratif », *Rivista di Linguistica* 2(1), 155-174.
- Kleiber, G. (1990c), « Marqueurs référentiels et processus interprétatifs : pour une approche 'plus sémantique' », *Cahiers de linguistique française* 11, 241-258.
- Kleiber, G. (1990d), « Article défini et démonstratif : approche sémantique vs approche cognitive », in G. Kleiber et J.-E. Tyvaert (éds.), *L'anaphore et ses domaines*, Paris, Klincksieck, 199-227.
- Kleiber, G. (1990e), L'article LE générique. La généricité sur le mode massif, Genève, Droz.
- Kleiber, G. (1992a), « Anaphore-deixis: deux approches concurrentes, in M.-A. Morel & L. Danon-Boileau (éds.), *La deixis*, Paris, PUF, 613-626.
- Kleiber, G. (1992b), « Ils ont encore augmenté les impôts ou Sur le Ils collectif », in L. Tasmowski-De Ryck & A. Zribi-Hertz (éds.), *De la musique à la linguistique. Hommages à Nicolas Ruwet*, Gand, Communication et cognition, 327-344.
- Kleiber, G. (1992c), « Anaphore associative et inférences », in J.-E. Tyvaert (éd.), *Lexique et inférence(s)*, Actes du VIIème Colloque International de Linguistique, Paris, Klincksieck, 175-201.
- Kleiber, G. (1994a), *Nominales*, Paris, Armand Colin.
- Kleiber, G. (1994b), *Anaphores et pronoms*, Louvain-la-Neuve, Duculot.
- Kleiber, G. (1997a), « Sens, référence et existence : que faire de l'extra-linguistique ? », *Langages* 127, 9-37.
- Kleiber, G. (1997b), « Les anaphores associatives actantielles », *Scolia* 10, 89-120.
- Kleiber, G. (1998), « TOUT et ses domaines : sur la structure TOUT+déterminant+N », in A. Englebert, M. Pierrard, L. Rosier & D. Van Raemdonck (éds.), *La ligne claire. De la linguistique à la grammaire*, Bruxelles, Duculot, 87-98.
- Kleiber, G. (1999), « Anaphore associative et relation partie-tout : condition d'aliénation et principe de congruence ontologique », *Langue française* 122, 70-100.
- Kleiber, G. (2001a), *L'anaphore associative*, Paris, PUF.
- Kleiber, G. (2001b), « Anaphore associative, lexique et référence ou Un automobiliste peut-il rouler en anaphore associative ? », in W. De Mulder, C. Vet & C. Veters (éds.), *Anaphores pronominales et nominales*, Amsterdam, Rodopi, 27-42.
- Kleiber, G. (2003), « Adjectifs démonstratifs et point de vue », *Cahiers de Praxématique* 41, 33-54.
- Kleiber, G. (2004), « Anaphores associatives : du large à l'étroit », in C., Leclère et al. (éds.), *Lexique, Syntaxe et Lexique-Grammaire. Papers in Honour of Maurice Gross*, Amsterdam/Philadelphia, John Benjamins, 287-301.
- Kleiber, G. (2011), « La quantification universelle en trio : tous les, chaque et tout », *Studii de lingvistică* 1, 139-157.
- Kleiber, G., Schnedecker C. & Ujma L. (1994), « L'anaphore associative, d'une conception à l'autre », in C. Schnedecker et al. (éds.), *L'anaphore associative. Aspects linguistiques, psycholinguistiques et automatiques*, Paris, Klincksieck, 5-64.
- Koch, P. & Oesterreicher, W. (1985), « Sprache der Nähe – Sprache der Distanz Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte », *Romanistisches Jahrbuch* 36, 15-43.
- Kripke, S. A. (1972), « Naming and necessity », in D. Davidson & G. Harman (éds.), *Semantics for natural languages*, Dordrecht, Reidel, 253-335.
- Laezlinger, C. (2003), *Initiation à la syntaxe formelle du français*, Berne, Peter Lang.

- Lambrecht, K. (1981), *Topic, antitopic and verb agreement in non-standard French*, Amsterdam, J. Benjamins.
- Lambrecht, K. (1994), *Information structure and sentence form*, Cambridge, Cambridge University Press.
- Lammert, M. (2010), *Sémantique et cognition : les noms collectifs*, Genève-Paris, Droz.
- Lammert, M. (2015), « Les pluralia tantum sous l'angle du collectif », *Langue française*, 185(1), 73-84.
- Lammert, M. & Lecolle, M. (2014), « Les noms collectifs en français : une vue d'ensemble », *Cahiers de lexicologie* 105(2), 203-222.
- Landragin, F. (2004), « Saillance physique et saillance cognitive », *CORELA* 2(2). <http://corela.edel.univ-poitiers.fr/index.php?id=603>
- Langacker, R. (1969), « On pronominalization and the chain of command », in D.A Reibel & S.A. Schane (éds.), *Modern Studies in English: readings in transformational grammar*, Englewood Cliffs, Prentice Hall, 160-186.
- Langacker, R. W. (1977), « Syntactic reanalysis », in C. Li, (éd.), *Mechanisms of syntactic change*, Austin, University of Texas Press, 57-139.
- Lasnik, H (1976), « Remarks on coreference », *Linguistic Analysis* 2(1), 1-22.
- Laurent, N. (2016), « Présentation », *Langue française* 190, 5-14.
- Lauwers, P. (2014), « Les pluriels "lexicaux". Typologie quantifiée des déficits de dénombrabilité », *Langue française* 183(3), 117-132.
- Lauwers, P., Willems, D. (2011), « Coercion : Definitions and challenges, current approaches, and new trends », *Linguistics* 49(6), 1219-1235.
- Lazard, G. (1995), « La définition des actants », in F. Madray-Lesigne & J. Richard-Zappella (éds.), *Lucien Tesnière aujourd'hui. Actes du Colloque International CNRS URA1164 – Université de Rouen, 16-18 novembre 1992*, Louvain-Paris, Peeters, 151-158.
- Lebas, F. (1997), « Conséquences théoriques des frontières de la polysémie. Application au pronom *il* », *Langue française* 113, 35-48.
- Le Bihan, M. (1974), *Le nom propre. Etude de grammaire et de rhétorique*, Thèse de doctorat de 3<sup>e</sup> cycle, Université de Rennes.
- Lecolle, M. (2013), « Noms collectifs humains : un point de vue de sémantique lexicale sur l'identité dans le rapport individu/groupe », *¿ Interrogations ?* 16. <http://www.revue-interrogations.org/NOMS-COLLECTIFS-HUMAINS-UN-POINT> (Consulté le 28 juillet 2015).
- Leeman, D. (1991), « On thème », *Linguisticae Investigationes* 15, 101-113.
- Leeman, D. (1994), *Les fautes de français existent-elles ?*, Paris, Seuil.
- Legallois, D. (2006), « Quand le texte signale sa structure : la fonction textuelle des noms sous-spécifiés », *Corela*, mis en ligne le 27 octobre 2006, consulté le 27 janvier 2016. URL : <http://corela.revues.org/1465>.
- Legallois, D. (2008), « Sur quelques caractéristiques des noms sous-spécifiés », *Scolia* 23, 109-127.
- Le Goffic, P. (1981), *Ambiguïté linguistique et activité du langage*, Thèse de Doctorat d'Etat, Université de Paris VIII.
- Le Goffic, P. (1982), « Ambiguïté et ambivalence en linguistique », *DRLAV* 27, 83-105.
- Le Guern, M. (2003), *Les deux logiques du langage*, Paris, Champion.
- Le Querler, A. (1996), *Typologie des modalités*, Caen, Presses Universitaires de Caen.

- Lesniewski, S. (1989), *Sur les fondements de la mathématique. Fragments* (trad. G. Kalinowski), Paris, Hermès.
- Levinson, S. C. (1986), « Pragmatics and the grammar of anaphora : a partial pragmatic reduction of binding and control phenomena », *Journal of Linguistics* 23(2), 379-434.
- Levinson, S. C. (2004), « Deixis », in L. R. Horn & G. L. Ward (éds.), *The handbook of pragmatics*, Oxford, Blackwell, 97-121.
- Lewis, D. (1979), « Score-keeping in a language game », *Journal of Philosophical Logic* 8, 339-359.
- Lyons, J. (1977), *Semantics* (Vol. I & II), Cambridge, Cambridge University Press.
- Lyons, J. (1981), *Language, meaning and context*, London, Fontana.
- Maillard, M. (1974), « Essai de typologie des substituts diaphoriques (supports d'une anaphore et/ou d'une cataphore) », *Langue française* 21, 55-71.
- Maillard, M. (1985), « L'impersonnel français de "il" à "ça" », in J. Chocheyras (éd.), *Autour de l'impersonnel*, Grenoble, Ellug, 63-118.
- Maillard, M. (1989), "Comment ça fonctionne" ou L'étude de fonctionnement de ça en français moderne dans la perspective d'une linguistique génétique, Thèse de doctorat, Univ. Paris X.
- Maillard, M. (1991), « Vers une théorie unitaire de l'impersonnel ? », in M. Maillard, *L'impersonnel. Mécanismes linguistiques et fonctionnements littéraires*, Grenoble, Ceditel, 227-254.
- Maillard, M. (1994a), « Présentation : Les problèmes généraux posés par l'impersonnel », *L'information grammaticale* 62, 3-5.
- Maillard, M. (1994b), « Concurrence et complémentarité de *il* et *ça* devant les prédicats impersonnels en français contemporain ou comment distinguer une phrase asubjectale d'une phrase à sujet indistinct », *L'information grammaticale* 62, 48-52.
- Maillard, M. & Almeida, E. (2000), « Un modèle nodal pour une description cohérente de l'impersonnel en français et en portugais », *Cahiers de l'ILSL* 12, 173-202.
- Maingueneau, D. (2000), « Instances frontières et angélisme narratif », *Langue française* 128, 74-95.
- Marandin, J.-M. (1986), « *Ce* est un autre. L'interprétation anaphorique du syntagme démonstratif », *Langages* 81, 75-89.
- Mari, A. (2006), « Collectif vs distributif » in D. Godard, L. Roussarie et F. Corblin (éds.), *Sémantopédie: dictionnaire de sémantique*, GDR Sémantique & Modélisation, CNRS, <http://www.semantique-gdr.net/dico/>.
- Marchello-Nizia, Ch. (1989), « Le neutre et l'impersonnel », *Linx* 21, 173-179.
- Martin, R. (1986), « Les usages génériques de l'article *le* et la pluralité », in J. David & G. Kleiber (éds.), *Déterminants, syntaxe et sémantique*, Paris, Klincksieck, 187-202.
- Martinet, A. (1967), *Éléments de linguistique générale* (2<sup>e</sup> éd.), Paris, Colin.
- Mertens, P. (2004) « Un outil pour la transcription de la prosodie dans les corpus oraux », *Traitement Automatique des langues* 45 (2), 109-130.
- Meyers, M.W. (1990), « Current generic pronoun usage : an empirical study », *American Speech* 65(3), 228-237.
- Michard, C. (1996), « Genre et sexe en linguistique : les analyses du masculin générique », *Mots* 49, 29-47.
- Mihatsch, W. (2006), « *Machin, truc, chose* : la naissance de marqueurs pragmatiques », in M. Drescher & B. Job (éds.), *Les marqueurs discursifs dans les langues romanes. Approches théoriques et méthodologiques*, Frankfurt : Peter Lang, 153-172.

- Mill, J. S. (1843=1988), *Système de logique déductive et inductive*, Liège, Mardaga (trad. L. Peisse 1866).
- Miller, P. H. (1992), *Clitics and constituents in phrase structure grammar*, New York, Garland.
- Miller, P. & Monachesi, P. (2003), « Les pronoms clitiques dans les langues romanes », in D. Godard (éd.), *Les langues romanes : Problèmes de la phrase simple*, Paris, Éditions du CNRS.
- Milner, J.-C. (1982), *Ordres et raisons de langue*, Paris, Seuil.
- Minsky, M. (1975), « A framework for representing knowledge », in P. H. Winston (éd.), *The psychology of computer vision*, New York, McGraw-Hill, 211-277.
- Moeschler, J. & Reboul, A. (1994), *Dictionnaire encyclopédique de pragmatique*, Seuil.
- Moeschler, J. & Zufferey, S. (2010), *Initiation à la linguistique française*, Paris, Armand Colin.
- Moignet, G. (1965), « Le problème du nombre en français », *Bulletin de la Faculté des Lettres de Strasbourg* 43(6), 463-479.
- Moignet, G. (1981), *Systématique de la langue française*, Paris, Klincksieck.
- Moussy, C. (1991), « The structure of the signifié : the usefulness and limitations of analysis into relevant features (with applications to Latin) », in R. Coleman (éd.), *New studies in Latin linguistics*, Amsterdam, Benjamins, 63-73.
- Muller, C. (1970), « Sur les emplois personnels de l'indéfini *on* », *Revue de linguistique romane* 34, 48-55.
- Muller, C. (2000), « Le passif processif et ses concurrents. Définition et quelques particularités », in L. Schøsler (éd.), *Le passif*, Copenhagen, Museum Tusulanum Press, 49-69.
- Muzerelle, J., Schang, E., Antoine, J.-Y., Eshkol, I., Maurel, D., Boyer, A. & Nouvel, D. (2012), « Annotations en chaînes de coréférences et anaphores dans un corpus de discours spontané en français », in *Actes du CMLF, 3<sup>ème</sup> Congrès Mondial de Linguistique Française*, 2497-2516, publié par EDP Sciences ([www.linguistiquefrancaise.org](http://www.linguistiquefrancaise.org)), DOI: <http://dx.doi.org/10.1051/shsconf/20120100252>
- Myhill, J. (1997), « Toward a functional typology of agent defocusing », *Linguistics* 35, 799-844.
- Noailly, M. (1997), « Les mystères de la transitivité », *Langages* 127, 96-109.
- Nølke, H. (1997), « Anaphoricité et focalisation : le cas du pronom personnel disjoint » in W. De Mulder, L. Tasmowski-De Ryck & C. Vettters (éds.), *Relations anaphoriques et (in)cohérence*. Amsterdam, Rodopi, 55-67.
- Ogden, C. K., & Richards, I. A. (1923), *The meaning of meaning : A study of the influence of language upon thought and of the science of symbolism*, Londres, Routledge & Kegan Paul.
- Ouellet, P. (1984), « La désénonciation : les instances de la subjectivité dans le discours scientifique », *Protée* 12(2), 43-54.
- Overstreet, M. (1999), *Whales, candlelight, and stuff like that : General extenders in English discourse*, Oxford, Oxford University Press.
- Overstreet, M. (2005), « *And stuff, und so*: Investigating pragmatic expressions in English and German », *Journal of Pragmatics* 37, 1845-1864.
- Panaccio, C. (1996), « Des signes dans l'intellect », *Cahiers d'épistémologie* 9603, 1-30.
- Panckhurst R., Détrie C., Lopez C., Moïse C., Roche M., Verine B. (2014), « 88milSMS. A corpus of authentic text messages in French », produit par l'Université Paul-Valéry Montpellier 3 et le CNRS, en collaboration avec l'Université catholique de Louvain, financé grâce au soutien de la MSH-M et du Ministère de la Culture (Délégation générale à la langue française et aux langues de France) et avec la participation de Praxiling, Lirmm, Lidilem, Tetis, Viseo. ISLRN : 024-713-187-947-8.
- Pariante, J.-C. (1973), *Le langage et l'individuel*, Paris, A. Colin.

- Paveau, M.-A. & Rosier, L. (2009), « Grammaire de la liste », in I. Evrard, M. Pierrard, L. Rosier & D. Van Raemdonck (éds.), *Le sens en marge. Représentations linguistiques et observables discursifs*, Paris, L'Harmattan.
- Payne, T. E. (1997), *Describing morphosyntax : A guide for field linguists*, Cambridge, Cambridge University Press.
- Peirce, C. S. (1931-1958), *Collected papers of Charles Sanders Peirce*, (vol. 1-8) in C. Hartshorne P Weiss & A. W. Burks (éds.), Cambridge MA, Harvard University Press.
- Pekarek, S. (1999), « Linguistic forms and social interaction : why do we specify referents more than is necessary for their identification ? », in J. Verschueren (éd.), *Pragmatics in 1998*, Antwerp, International Pragmatics Association, 437-448.
- Porquier, R. (1972), « L'emploi de " ça " en français parlé », *Le Français dans le monde* 12(91), 9-16.
- Pottier, B. (1963), *Recherches sur l'analyse sémantique en linguistique et en traduction mécanique*, Série A. Linguistique appliquée et Traduction automatique II, Publications linguistiques de la Faculté des Lettres & Sciences Humaines de l'Université de Nancy.
- Prince, E. (1981), « Toward a taxonomy of given-new information », in P. Cole (éd.), *Radical pragmatics*, New York, Academic Press, 223-255.
- Pustejovsky, J. (1995), *The generative lexicon*, Cambridge, MA, The MIT Press.
- Quine, W. V. O. (1960), *Word and object*, Cambridge MA, MIT Press.
- Rabatel, A. (2001), « La valeur de *on* pronom indéfini / pronom personnel dans les perceptions représentées », *L'Information grammaticale* 88, 28-32.
- Rastier, F. (2004), « Ontologie(s) », *Revue d'intelligence artificielle* 18(1), 15-40.
- Read, S. (2011), « Medieval theories : properties of terms », *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Zalta E. N. (éd.) URL = <<http://plato.stanford.edu/archives/spr2011/entries/medieval-terms/>>.
- Récanati, F. (1983), « La sémantique des noms propres : remarques sur la notion de 'désignateur rigide' », *Langue française* 57, 106-118.
- [Reichler-]Béguelin, M.-J. (1988), « Anaphore, cataphore et mémoire discursive », *Pratiques* 57, 15-45.
- [Reichler-]Béguelin, M.-J. (1989), « Anaphores, connecteurs et processus inférentiels », in Ch. Rubattel (éd.), *Modèles du discours. Recherches actuelles en Suisse romande*, Berne, Peter Lang, 303-336.
- [Reichler-]Béguelin, M.-J. (1990), *Écrire en français. Cohésion textuelle et apprentissage de l'expression écrite* (2<sup>e</sup> éd.), avec M. Denervaud et J. Jespersen, Neuchâtel-Paris, Delachaux & Niestlé.
- [Reichler-]Béguelin, M.-J. (1993a), « Anaphores associatives non lexicales : incomplétude macrosyntaxique? », in S. Karolak & T. Muryn (éds.), *Complétude et incomplétude dans les langues romanes et slaves : actes du VI Colloque international de Linguistique romane et slave, Cracovie, 29 sept. – 3 oct. 1991*, Cracovie, Wydawnictwo Naukowe WSP, 327-379.
- [Reichler-]Béguelin, M.-J. (1993b), « Faits déviants et tri des observables », *TRANEL* 20, 89-109.
- [Reichler-]Béguelin, M.-J. (1995a), « Alternatives et décisions lexicales dans l'emploi des expressions démonstratives », *Pratiques* 85, 53-87.
- [Reichler-]Béguelin, M.-J. (1995b), « Déterminant zéro et anaphore », *TRANEL* 23, 177-201.
- Reinhart, T. (1986), *The syntactic domain of anaphora*, PhD Dissertation, MIT.
- Rey, A. (1973), *Théories du signe et du sens. Lectures*, t. 1, Paris, Klincksieck.
- Riegel, M. (1985), *L'adjectif attribut*, Paris, PUF.

- Riegel, M., Pellat J.-Ch. & Rioul R. (2009), *Grammaire méthodique du français* (4<sup>e</sup> éd.), PUF, Paris.
- Rivas, A. (1977), *A Theory of clitics*, thèse de doctorat, MIT.
- Rizzi, L. (1986a), « On the status of subject clitics in Romance », in O. Jaegli et C. Silva-Corvalán (éds.), *Studies in Romance linguistics*, Dordrecht, Foris Publications, p. 391-419.
- Rizzi, L. (1986b), « Null objects and the theory of pro », *Linguistic Inquiry* 17, 501-557.
- Roberge, Y. (1995), « Présentation », *Revue québécoise de linguistique* 24(1), 7-10.
- Rosch, E., Mervis, C. B., Gray, W. B., Johnson, D. M. & Boyes-Braem, P. (1976), « Basic objects in natural categories », *Cognitive Psychology* 8, 382-439.
- Roulet, E. (1999), *La description de l'organisation du discours. Du dialogue au texte*, Paris, Didier.
- Rousseau, A. (1996), « Les mots abstraits : une approche interdisciplinaire », in N. Flaux, M. Glatiny & D. Samain (éds.), *Les noms abstraits. Histoires et théories*, Villeneuve d'Ascq, Presses universitaires du Septentrion, 51-65.
- Russell, B. (1919), *Introduction to mathematical philosophy*, Allen & Unwin, Londres.
- Sales, M.-P. (2008), *Influence du lexique et de la syntaxe sur la reprise pronominale : exemple de ça*, Thèse de 3<sup>e</sup> cycle, Université Paris Ouest La Défense.
- Sapir E. (1921), *Language. An introduction to the study of speech*, New York, Harcourt, Brace & World.
- Sadock, J. M. (1977), « Truth and Approximations », *Proceedings of the 3rd Annual Meeting of the Berkeley Linguistics Society*, 430-439.
- Sandfeld, Kr. (1970), *Syntaxe du français contemporain : Les pronoms*, Paris, Champion.
- Sanford A. J., Filik, R. (2006), « “They” as gender-unspecified singular pronoun : eye-tracking reveals a processing cost », *The Quarterly Journal of Experimental Psychology* 60(2), 171-178.
- Sanford A. J., Filik, R., Emmott, C. & Morrow, C. (2008), « They're digging up the road again: the processing cost of Institutional *They* », *The Quarterly Journal of Experimental Psychology* 61(3), 372-380.
- Sanford, A. J. & Garrod, S. C. (1981), *Understanding written language : explorations of comprehension beyond the sentence*, Chichester, John Wiley & Sons.
- Sanford, A. J. & Garrod, S. C., Lucas, A. & Henderson, R. (1983), « Pronouns without explicit antecedents ? », *Journal of Semantics* 2(3-4), 303-318.
- Sanford A. J. & Sturt, P. (2002), « Depth of processing in language comprehension : Not noticing the evidence », *Trends in Cognitive Science* 6(9), 382-386.
- Saussure, F. de (1916=2008), *Cours de linguistique générale* (éds. Ch. Bally & A. Sechehaye, avec la coll. d'A. Riedlinger, éd. critique T. de Mauro), Paris, Payot.
- Saussure, F. de (2002), *Écrits de linguistique générale* (éds. S. Bouquet & R. Engler), Paris, Gallimard.
- Schank, R. C. & Abelson, R. (1977), *Scripts, plans, goals, and understanding : An enquiry into human knowledge structures*, Hillsdale, N. J., Erlbaum.
- Schiffrin, D. (1994), « Making a List », *Discourse Processes* 17, 377-406.
- Schmid, H.-J. (1997), « Constant and ephemeral hypostatization : thing, problem, and other “shell nouns” », in B. Caron (éd.) *Proceedings of the 16th International Congress of Linguists (Paris, July 22-25, 1997)*, CD-ROM, Amsterdam, Elsevier.
- Schmid, H.-J. (2000), *English abstract nouns as conceptual shells : From corpus to cognition*, Berlin, Walter de Gruyter.

- Schneedecker, C. (1997), *Nom propre et chaînes de référence*, Paris, Klincksieck.
- Schneedecker, C. (2000), « L'un... sans l'autre : encore un "indéfini" marginal », *Scolia* 13, 195-213.
- Schneedecker, C. (2003), « Les uns... une pluralité singulière », *Langages* 151, 87-104.
- Schneedecker, C. (2008), « Tout nom et tout pronom : qu'est ce qui fait la différence ? », in J. Durand, B. Habert & B. Laks (éds.) *Actes du Congrès Mondial de Linguistique Française 2008*, CMLF08, EDP Sciences, 2159-2171, <http://dx.doi.org/10.1051/cmlf08190>
- Schneedecker, C. (2011), « Monsieur tout le monde, Maud Machin-Chouette, Denise Trucmuche, et les autres.... Inventaire et comportement des noms propres " indéfinis " du français », in D. AMiot, W. De Mulder, E. Moline & D. Stosic, *Ars Grammatica. Hommage à Nelly Flaux*, Berne, Peter lang, 37-54.
- Schneedecker, C. (2014), « Chaînes de référence et variations selon le genre », *Langages* 195(3), 23-42.
- Schneedecker, C. (à par.), « Les pronoms dits "indéfinis" référant à la personne humaine : des grammaires à l'usage et réciproquement », *Bulletin de la Société de Linguistique de Paris*.
- Schneedecker, C. & Charolles, M. (1993), « Les référents évolutifs : point de vue ontologique et phénoménologique », *Cahiers de linguistique française* 14, 197-227.
- Schneedecker, C. & Landragin, F. (2014), « Les chaînes de référence : présentation », *Langages* 195(3), 3-22.
- Schwarzschild, R. (2011), « Stubborn distributivity, multiparticipant nouns and the count/mass distinction », in S. Lima, K. Mullin & B. Smith (éds), *Proceedings of NELS 39*, Amherst (MA), GLSA, 661-678.
- Searle, J. R. (1969), *Speech acts : An essay in the philosophy of language*, Cambridge, Cambridge University Press.
- Searle, J. R. (1985), *L'intentionnalité : essai de philosophie des états mentaux*, Paris, Les Editions de Minuit.
- Secova, M. (2014), « 'Je sais et tout mais...' might the general extenders in European French be changing ? », *Journal of French Language Studies* 24(2), 281-304.
- Selting, M. (2007), « Lists as embedded structures and the prosody of list construction as an interactional resource », *Journal of Pragmatics* 39(3), 483-526.
- Sidner, C. L. (1981), « Focusing for the interpretation of pronouns », *Computational Linguistics* 7/4.
- Sidner, C. L. (1983), « Focusing in the comprehension of definite anaphora », in M. Brady and R. C. Berwick (éds.), *Computational models of discourse*, Cambridge MA, MIT Press, 267-330.
- Siepmann, D., Bürgel, Ch. & Diwersy, S. (2016), « Le Corpus de référence du français contemporain (CRFC), un corpus massif du français largement diversifié par genres », SHS Web of Conferences 27(11002), 5<sup>e</sup> Congrès Mondial de Linguistique Française, <http://dx.doi.org/10.1051/shsconf/20162711002>.
- Siewierska, A. (2004), *Person*, Cambridge, Cambridge University Press.
- Siewierska, A. (2008), « Ways of impersonalizing : pronominal vs. verbal strategies », in M. Gómez González et al. (éds.), *Current Trends in Contrastive Linguistics : Functional and cognitive perspectives*, Amsterdam, John Benjamins, 3-26.
- Siewierska, A. (2010), « From Third Plural to passive : incipient, emergent and established passives », *Diachronica* 27(1), 73-109.
- Siewierska, A. (2011), « Overlap and complementarity in reference impersonals : Man-constructions vs. third person plural-impersonals in the languages of Europe », in A. Malchukov et A. Siewierska *Impersonal constructions. A cross-linguistic perspective*, John Benjamins, Amsterdam/Philadelphia, 57-90.

- Siewierska, A. & Papastathi, M. (2011), « Towards a typology of third person plural impersonals », *Linguistics* 49(3), 575-610.
- Sinclair, J. (1996), « Preliminary recommendations on corpus typology », Technical report, Eagles (Expert Advisory Group on Language Engineering).
- Simone, R. (2014), « Une théorie simple de la saillance », in M. Avanzi et al., *Enseignement du français : les apports de la recherche en linguistique. Réflexions en l'honneur de Marie-José Béguelin*, Bruxelles, Peter Lang, 343-359.
- Smith, J. C. (1995), « L'évolution sémantique et pragmatique des adverbes déictiques *ici, là* et *là-bas* », *Langue française* 107, 43-57.
- Sperber, D. & Wilson, D. (1986), *Relevance: communication and cognition*, Oxford, Blackwell.
- Sperber, D. & Wilson, D. (1990), « Rhetoric and Relevance », in D. Wellbery & J. Bender (éds.), *The Ends of Rhetoric: History, Theory, Practice*, Stanford, Stanford University Press, 140-155.
- Stark, E., Ueberwasser, S. & Ruef, B. (2009-2014), *Swiss SMS Corpus*, University of Zurich, <https://sms.linguistik.uzh.ch>
- Strawson, P. (1959), *Individuals. An essay in descriptive metaphysics*, Londres, Methuen.
- Strawson, P. (1964), « Identifying reference and truth values », *Theoria* 30, 96-118.
- Strozer, J. (1976), *Clitics in Spanish*, PhD Dissertation, UCLA.
- Sturt, P., Sanford, A.J., Stewart, A. & Dawydiak, E. (2004), « Linguistic focus and good-enough representations : An application of the change-detection paradigm », *Psychonomic Bulletin and Review* 11, 882-888.
- Suñer, M. (1983), « *pro<sub>arb</sub>* », *Linguistic Inquiry* 14(1), 188-191.
- Tarski, A. (1935), « Der Wahrheitsbegriff in den formalisierten Sprachen », *Studia Philosophica* 1, 261-405.
- Tasmowski-De Ryck, L. & Verluyten, S. P. (1982), « Linguistic control of pronouns », *Journal of Semantics* 1(3-4), 323-346.
- Tasmowski-De Ryck, L. & Verluyten, S. P. (1985), « Control mechanism of anaphora », *Journal of Semantics* 4(4), 341-370.
- Tesnière, L. (1959), *Éléments de syntaxe structurale*, Paris, Klincksieck.
- Teston-Bonnard, S., Lefevre F. & Morel, M.-A. (2011), « Valeur prototypique de *quoi* à travers ses usages en français oral et contemporain », *Neophilologische Mitteilungen (Bulletin de la Société Néophilologique)*, 37-61.
- Theissen, A. (2008), « *Ceci* en emploi cataphorique », in O. Bertrand, S. Prévost, M. Charolles, J. François, C. Schnedecker (éds.), *Discours, diachronie, stylistique du français*, Bern, Peter Lang, 159-173.
- Townsend, D. J., & Bever, T. G. (2001), *Sentence comprehension : The integration of habits and rules*, Cambridge MA, MIT Press.
- Trésor de la langue française informatisé*, Atilf, Université de Lorraine, <http://atilf.atilf.fr/>.
- Vaguer, C. (2010), « Quand "facilement" rencontre un numéral, ça parle d'approximation », *L'information grammaticale* 125, 29-35.
- Valentin, P. (1985), « How does anapher work ? », in U. Pieper & G. Stickel (éds.), *Studia Linguistica Diachronica et Synchronica. Werner Winter sexagenario anno MCMLXXXIII*, Berlin/New York/Amsterdam, Mouton de Gruyter, 849-859.
- Van de Velde, D. (2006), *Grammaire des événements*, Villeneuve d'Ascq, Presses universitaires du Septentrion.

- Vendler, Z. (1967), *Linguistics in Philosophy*, Ithaca NY, Cornell University Press.
- Vendler, Z., (1970), « Say what you think », in J.L. Cowan, (éd.), *Studies in thought and language*, Tucson Arizona, The University of Arizona Press, p. 79-97.
- Vincent, D. (1993), *Les ponctuels de la langue et autres mots du discours*, Québec, Nuit Blanche.
- Vincent, D. & Demers, M. (1994), « Les problèmes d'arrimage entre les études discursives et prosodiques. Le cas du *là* ponctuel », *Langues et linguistique* 20, 201-212.
- Viollet, C. (1988), « Mais qui est ON? » *Linx* 18, 67-75.
- Vuillaume, M (1993), « Le repérage temporel dans les textes narratifs », *Langages* 112, 92-105.
- Von Wartburg, W. & Zumthor, P. (1973), *Précis de syntaxe du français contemporain*, 3<sup>e</sup> éd., Berne, A. Francke.
- Wagner, R.-L. & Pinchon, J. (1962), *Grammaire du français classique et moderne*, Paris, Hachette.
- Waltereit, R. (2007), « À propos de la genèse diachronique des combinaisons de marqueurs. L'exemple de *bon ben* et *enfin bref* », *Langue française* 154(2), 94-109.
- Ward, G. & Birner, B. (1993), « The semantics and pragmatics of *and everything* », *Journal of Pragmatics* 19, 205-214.
- Wierzbicka, A. (1988), *The semantics of grammar*, Amsterdam, John Benjamins Publishing.
- Wiese, B. (1983), « Anaphora by pronouns », *Linguistics* 21, 373-417.
- Willems, D. (1998), « *La mer, c'est beau*. Le sujet dans certaines structures non canoniques », in M. Forsgren, K. Jonasson & H. Kronning (éds.), *Prédication, assertion, information : actes du colloque d'Uppsala en linguistique française, 6-9 juin 1996*, Uppsala, Uppsala University Press, 595-604.
- Wilmet, M. (1986), *La détermination nominale*, Paris, PUF.
- Wilmet, M. (1996), « A la recherche du *nom abstrait* », in N. Flaux, M. Glatiny & D. Samain (éds.), *Les Noms abstraits. Histoire et Théories*, Paris, Presses Universitaires du Septentrion, 67-76.
- Wilmet, M. (2007), *Grammaire rénovée du français*, Bruxelles, De Boeck-Duculot
- Wilmet, M. (2010), *Grammaire critique du français* (5<sup>e</sup> éd.), Bruxelles, De Boeck-Duculot.
- Windisch, E. (1869), *Untersuchungen über den Ursprung des Relativpronomens in den indogermanischen Sprachen*, Doctoral dissertation, Leipzig.
- Winter, E. O. (1992), « The notion of unspecific versus specific as one way of analysing the information of a fund-raising letter », in W.C. Mann & S.A. Thomson (éds.), *Discourses descriptions. Diverse analyses of a fund-raising text*, Amsterdam-Philadelphia, Benjamins, 131-170.
- Wisniewski, E. J. (2009), « On using count nouns, mass nouns, and pluralia tantum : What counts? », in F. J. Pelletier (éd.), *Kinds, Things, and Stuff : Mass Terms and Generics*, Oxford, Oxford University Press, 166-190.
- Wittgenstein, L. (2011), *Philosophische Untersuchungen : auf der Grundlage der Kritisch-genetischen Edition*, J. Schulte (éd.), Frankfurt, Suhrkamp.
- Wolters, M. K. (2001), *Towards entity Status*, Doctoral dissertation, Rheinischen Friedrich-Wilhelms-Universität, Bonn.
- Yule, G. (1982), « Interpreting anaphora without identifying reference », *Journal of Semantics* 1(3-4), 315 -322.
- Zadeh, L. A. (1975), « Foreword » in A. Kaufmann (éd.), *Introduction to the Theory of Fuzzy Subsets*, New York, Academic Press.
- Zribi-Hertz, A. (1996), *L'anaphore et les pronoms*, Paris, Presses universitaires du Septentrion.

Zumwald Küster, G. (2014), « Quel statut pour les sujets *je, tu, il, ...* et les objets *me, la, lui, en... ?* », *Verbum* 36(1), 17-34.