

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

HÉTÉROGÉNÉITÉ DE LA CONNAISSANCE ET PRATIQUE SCIENTIFIQUE:
RÔLE EXPLICATIF ET IMPLICATIONS ÉPISTÉMOLOGIQUES DE LA
PSYCHOLOGIE ÉVOLUTIONNISTE ET DE LA NEUROPSYCHOLOGIE DE
LA VISION

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
À LA MAÎTRISE EN PHILOSOPHIE

PAR
FRÉDÉRIC-ISMAËL BANVILLE

Février 2013

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

AVANT-PROPOS

La forme initiale du projet était autrement plus radicale et ambitieuse qu'elle ne l'est devenue : le plan original était d'éliminer complètement le terme « connaissance » du vocabulaire philosophique. Force a été de reconnaître, toutefois, qu'il vaut mieux commencer par le commencement (et faire preuve de modération). Aussi, le présent mémoire doit être vu comme la première étape d'un long projet qui, je l'espère, sera un jour achevé (ou à tout le moins avancé).

Je tiens à remercier toutes les personnes que j'ai croisées dans mon parcours académique. Je remercie particulièrement : mon directeur de recherche Pierre Poirier, professeur de philosophie à l'UQÀM pour ses conseils, ses critiques et sa (grande) patience; mon co-directeur Jean-Pierre Marquis, professeur de philosophie à l'Université de Montréal, Luc Faucher et Serge Robert, membres du jury, professeurs de philosophie à l'UQÀM. Je remercie également la Fondation de l'UQÀM pour son soutien financier.

Il serait injuste d'omettre de remercier les quelques amis indispensables qui m'ont épaulé, d'une manière ou d'une autre, durant les deux dernières années (certains pendant plus longtemps encore) : Maxime Héroux-Legault, pour son omniprésence; Guillaume Beaulac, pour ses conseils ainsi que pour m'avoir fourni l'idée à la base de ce mémoire, Evelyn McDuff, pour sa patience infinie; Mirna Panic, qui sait remettre les choses en perspective; ainsi que Michaël Simard et Frédérick Armstrong, qui me rappellent qu'il y a un monde hors de l'université. Mes sincères remerciements à ma famille, dont le support a été aussi indispensable qu'inconditionnel : Danielle Salvail et Martin Banville, mes parents; Arnaud Banville, mon frère; Laurence Olivier, ma « sœur »; mes cousins Patrick et Sébastien Mouttet; mon oncle Clément Banville et ma tante Agnès ainsi que mes grands-parents, Christine et Bertrand Banville.

Je dédie ce mémoire à mon grand-père, professeur de physique à l'Université de Montréal de 1964 à 1996. Je ne suis certes pas devenu physicien, mais je dois ma passion pour la science et la connaissance à nos longues marches dans le parc d'Oka.

TABLE DES MATIÈRES

AVANT-PROPOS	I
TABLE DES MATIÈRES	II
LISTE DES FIGURES.....	V
LISTE DES TABLEAUX.....	VI
RÉSUMÉ	VII
INTRODUCTION	1
0.1. Des phrases aux neurones	2
0.2. Méthodologie et démarche.....	6
0.3. Structure	9
CHAPITRE I	
HÉTÉROGÉNÉITÉ: EXEMPLES ET PARAMÈTRES	12
1.1. Hétérogénéité : exemples	12
1.1.1. Griffiths : les espèces naturelles.....	13
1.1.2. Griffiths et les émotions : élimination du concept	17
1.1.3. Machery et les concepts	23
1.2. Hypothèses d'hétérogénéité	30
1.2.1. Espèces naturelles: une perspective sur la pratique scientifique.....	30
1.2.2. Hypothèses d'espèce naturelle	38
1.2.3. Hétérogénéité : conditions	40
1.3. Conclusion	42

CHAPITRE II

ÉNONCÉ DU PROBLÈME DE RECHERCHE ET ARGUMENT DE NÉCESSITÉ EXPLICATIVE.....	44
2.1. Histoire et caractérisation du paradigme phrastique	45
2.1.1. Émergence des conceptions phrastiques de la connaissance	45
2.1.2. Caractérisation du paradigme phrastique.....	48
2.2. Présentation et évaluation de l'éliminativisme	53
2.2.1. Problème 1 : le développement	56
2.2.2. Problème 2 : connaissance animale	57
2.2.3. Problèmes 3 et 4 : cognition adulte et vérité	59
2.2.4. Synthèse de la position naturaliste de Churchland : un naturalisme de remplacement .	63
2.3. Métathéorie : caractérisation générale du débat.....	64
2.4. Hétérogénéité : conditions contextuelles et argument de nécessité explicative.....	66
2.4.1. Comment le débat sur la nature de la connaissance remplit les conditions contextuelles	67
2.4.2. Condition méthodologique, partie 1 : raisons conceptuelles de proposer l'hétérogénéité; assises de l'argument de nécessité explicative	68
2.4.3. Condition méthodologique, partie 2 : argument de nécessité explicative.....	71
2.5. Conclusion	73

CHAPITRE III

IMPLICATIONS ÉPISTÉMOLOGIQUES DE LA PSYCHOLOGIE ÉVOLUTIONNISTE ET DE LA NEUROPSYCHOLOGIE DE LA VISION; CONCEPTION HÉTÉROGÈNE DE LA CONNAISSANCE	74
3.1. Considérations préliminaires.....	74
3.1.1. Conditions minimales	75

3.1.2. Justification; un premier axe de différenciation.....	78
3.2. Psychologie évolutionniste : implications pour le critère de vérité; second axe de différenciation.....	80
3.2.1. Argument des mécanismes d'attribution de valeur.....	81
3.2.2. Critique de l'argument.....	85
3.2.3. Vérité pragmatique.....	88
3.2.4. L'apport de la psychologie du raisonnement.....	93
3.2.5. Analyse des implications : comment établir la pertinence du second axe de différenciation.....	101
3.3. Implications de la neuropsychologie de la vision : pertinence des caractérisations phrastiques.....	103
3.3.1. Le cas D.F. : considérations neuroanatomiques et description.....	103
3.3.2. Analyse des implications : comment établir le troisième axe de différenciation.....	107
3.4. Synthèse.....	109
3.4.1. Argument de nécessité explicative, prise 2.....	109
3.4.2. Hétérogénéité de la connaissance : une première suggestion.....	112
3.5. Conclusion.....	114
3.5.1. Obstacles à surmonter.....	114
3.5.2. Résumé.....	115
CONCLUSION.....	117
BIBLIOGRAPHIE.....	121

LISTE DES FIGURES

Figure		Page
3.1	Effet <i>less-is-more</i> : performances comparatives de procédés heuristiques et d'algorithmes complexes (Gigerenzer et Brighton, 2009 p. 112)	97
3.2	Tâches comportementales sensibles aux lésions corticales visuelles chez les singes (Ungerleider et Macko, 1983, p. 120)	104

LISTE DES TABLEAUX

Tableau		Page
1.1.	Correspondances des niveaux explicatifs biologiques et psychologiques. Entre parenthèse sont les caractéristiques des programmes d'affect pertinentes à ce niveau d'explication. Tableau adapté de Griffiths 1997, p. 221, informations sur les programmes d'affect tirées du chapitre 9, section 2 (Griffiths 1997)	19
3.1.	Typologie des formes de connaissance (version provisoire)	112

RÉSUMÉ

Ce mémoire est consacré à une question centrale en épistémologie, soit celle du niveau descriptif approprié à l'étude de la connaissance (ou, simplement, de la nature de la connaissance). L'auteur remet en question la forme « traditionnelle » des solutions proposées à ce problème par les deux principaux courants de l'épistémologie naturalisée. En effet, les théories de la connaissance naturalistes optent généralement pour l'une de deux possibilités : soit la connaissance est adéquatement étudiée au niveau phrastique, soit au niveau du substrat physique (souvent, il s'agit du niveau neurologique). L'intention est de fournir une caractérisation de la connaissance sous forme d'« espèce naturelle ». Le présupposé implicite à ces propositions est qu'il n'existe qu'une seule « Connaissance », que celle-ci a une nature unique. Toutefois, les critiques adressées au « paradigme » phrastique suggèrent que celui-ci est insuffisant au plan explicatif parce que, 1) il ne s'applique directement qu'à une infime partie de l'ensemble des organismes pouvant avoir des états épistémiques et 2) parce qu'il repose sur des notions superficielles, issues de la psychologie du sens commun, qui doivent être éliminées (Churchland 1979, 1989, 2007). Ces approches se heurtent également à un problème explicatif lorsqu'il est question des formes de connaissance reposant sur la capacité pour le langage (telles que les définitions). Pour expliquer le phénomène de la connaissance il est nécessaire d'avoir recours à (au moins) ces deux approches. Par conséquent, la forme moniste de la solution au problème de la nature de la connaissance semble être inappropriée. L'objectif du mémoire est de défendre cette thèse. Il s'agit d'un travail métathéorique qui constitue la première étape d'un projet à plus grande échelle visant à développer une conception hétérogène de la connaissance. Le mémoire comporte trois chapitres :

- 1) Le premier chapitre vise à mettre en place les fondements de l'argument développé. Il détaille deux exemples d'approches hétérogènes tirées de la philosophie de la psychologie et en tire un plan général pour le développement d'un argument en faveur de l'hétérogénéité de la connaissance.
- 2) Le second chapitre procède à une analyse métathéorique des approches naturalistes à la connaissance. Ceci permet de formuler l'argument de nécessité explicative à l'endroit de la connaissance, central au mémoire.
- 3) Le troisième chapitre fournit un support empirique à la prémisse centrale de l'argument de nécessité explicative. L'argument est ensuite réitéré, et ses implications sont détaillées. Est ensuite proposée une première, et provisoire, ébauche d'une conception hétérogène de la connaissance.

Mots-clés : Épistémologie naturalisée, psychologie, neuropsychologie, connaissance

INTRODUCTION

De manière générale, il est fort louable de tenter de déterminer, dans le cadre d'un débat formel ou lors de discussions quotidiennes, sur quoi exactement porte le désaccord. L'étude de la connaissance n'y fait pas exception : il est intuitif de supposer que l'épistémologie, pour avancer de manière substantielle, doit répondre à la question du niveau descriptif approprié à l'étude de la connaissance ou, en d'autres termes, de la nature de celle-ci. Il s'agit d'une question difficile, mais cette difficulté est peut-être la conséquence d'une mauvaise manière d'aborder cette question, d'une conception erronée de la forme que doit prendre une telle réponse.

La forme appropriée de la réponse que doit recevoir la question de la nature de la connaissance est le sujet du présent mémoire. Nous argumentons contre l'idée implicite à la majeure partie de la recherche épistémologique qu'il existe quelque chose comme la Nature de la connaissance. Il faut plutôt considérer le terme « connaissance » comme référant à un ensemble d'entités hétérogènes, chacune pouvant (et devant) être considérée comme une instance de connaissance et dont l'étude est mieux abordée à partir d'une perspective particulière.

Notre travail est de nature métathéorique : il s'agit de montrer que la forme traditionnelle de la solution, qui est moniste, est inadéquate. Nous nous limitons à la démonstration de cette idée, et nous proposons une première ébauche, partielle, de ce à quoi pourrait ressembler une conception hétérogène de la connaissance. Il s'agit d'une suggestion, que nous comptons certes développer dans des travaux futurs, mais qui ne doit pas être vue comme une proposition substantive complète.

0.1. Des phrases aux neurones

L'épistémologie, au 20^e siècle, est caractérisée par l'adoption généralisée de l'idée que les phrases sont le véhicule privilégié de la signification (Quine 2004). Ceci explique la place centrale des croyances (véhiculées par les phrases) en épistémologie « traditionnelle », où la connaissance est (du moins avant Gettier) définie comme une « croyance vraie justifiée » (Steup 2008). Cette définition se heurte au fameux problème de Gettier (1963), soit que la condition de justification, censée empêcher les occurrences de « chance épistémique » n'est simplement pas suffisante pour remplir ce rôle. L'idée que la connaissance correspond à des phrases, toutefois, est demeurée sensiblement inchangée, le travail se portant sur l'enjeu de trouver une fondation solide et privilégiée à la connaissance. Cette entreprise est incarnée dans le fondationnalisme (des prémisses), selon lequel toute connaissance est justifiable sur la base d'un noyau de croyances fondamentales jouissant d'une sorte de privilège épistémique.

Quine (1969) rejette le projet fondationnaliste, le voyant comme une tentative de reconstruire de manière rationnelle le développement de la connaissance. Selon lui, il faut plutôt expliquer ce développement de manière causale. L'empirisme de Quine lui fait affirmer que puisque tout ce dont nous disposons pour aborder le monde sont nos impressions sensorielles, il faut s'intéresser aux processus par lesquels ces sensations deviennent des connaissances : comment de simples sensations nous passons à des énoncés observationnels et théoriques qui reçoivent l'assentiment de nos pairs. Quine veut donc expliquer *causalement* le développement de la connaissance, et pour ce faire il recommande un changement de projet et de méthode important. Le projet de l'épistémologie passe de trouver les fondements de la connaissance à l'étude du développement causal de celle-ci. Ceci impose un changement méthodologique important : il faut que l'épistémologie adopte les méthodes de la psychologie. Bien que Quine propose une sorte d'intégration d'une discipline à l'autre, sa position est souvent désignée par l'expression de « naturalisme de remplacement » : les méthodes de l'épistémologie traditionnelle sont remplacées par celles de la science empirique (Feldman 2001). Il existe d'autres versions du naturalisme. Le naturalisme coopératif, par exemple, affirme simplement que le travail épistémologique doit

tenir compte des avancées scientifiques pertinentes. Kornblith (2002) est un exemple de ce type d'approche.

À travers les changements proposés par Quine et les travaux subséquents au sein de l'épistémologie naturalisée, l'idée que la connaissance est adéquatement décrite par la terminologie phrastique est demeurée prépondérante (nous discutons en détail de cette prépondérance au chapitre II). Il existe toutefois une position radicale affirmant que nos concepts folkloriques concernant l'esprit humain sont inappropriés parce que superficiels au plan explicatif. Ceux-ci doivent être éliminés au profit de notions scientifiquement cohérentes, celles-ci étant souvent situées au niveau neurologique ou, plus généralement, au niveau du substrat physique. Cette position est entre autres développée par Paul Churchland (1979, 1988, 1989, 2007). Les implications épistémologiques de cette position sont évidentes : le paradigme phrastique doit être éliminé parce qu'il est superficiel au plan explicatif.

Les arguments de Churchland reposent sur deux questions distinctes (mais non indépendantes) : la question des véhicules de la connaissance et la question du contenu de ces véhicules. Il considère, d'une part, que répondre à la question des véhicules de la connaissance par la conception phrastique est une erreur parce qu'il existe 1) des organismes qui n'emploient pas de phrases mais sont tout de même des agents épistémiques (animaux non-humains, enfants en bas âge, aphasiques, voir Churchland 1979) et 2) des instances de connaissance qui même chez une humaine adulte ne sont pas phrastiques (les schémas moteurs requis pour jouer au baseball, voir Churchland 2007). D'autre part, en développant une théorie physicaliste et neurologique de la cognition humaine (1992) dans laquelle les véhicules informationnels pertinents sont des vecteurs d'activation neuronale, il rejette l'idée que la connaissance soit propositionnelle. En effet, de tels vecteurs d'activité n'ont pas de valeur de vérité, mais sont évalués sur la base d'un critère multivalent de réduction d'erreur. Donc, les réponses du paradigme phrastique aux questions des véhicules et du contenu de la connaissance (phrastique et propositionnel, respectivement) sont rejetées par Churchland, qui développe une position impliquant l'élimination de la terminologie phrastique.

Les arguments de Churchland sont convaincants, mais uniquement dans la mesure où ils démontrent l'insuffisance du paradigme phrastique. En effet, il existe des contextes dans lesquels une approche phrastique est plus puissante, au point de vue explicatif, qu'une approche réductionniste. En particulier, les définitions, qui sont des ensembles de conditions nécessaires et suffisantes communiquées de manière verbale (par le langage naturel) jouent un rôle central dans l'utilisation et la transmission de la connaissance chez les êtres humains. Poirier et Beaulac (2011) suggèrent qu'en plus des trois espèces naturelles qui constituent l'ensemble des concepts (exemplaires, prototypes, théories), les définitions doivent être considérées comme un quatrième type de concept, principalement employé par les processus de type 2 (par opposition aux processus de type 1, considérés comme automatiques et associatifs). Les définitions permettent d'acquérir le concept d'une chose avec laquelle un individu n'a jamais eu de contact. De plus, elles dépendent de manière importante de la faculté pour le langage. En ce sens, le phénomène de la connaissance, chez l'être humain, ne peut recevoir d'explication purement neurologique, il faut des outils pour analyser les particularités du langage et son influence sur des processus de plus bas niveau.

Le problème qui découle de cet état de la question sur le débat concernant le niveau descriptif approprié à l'étude de la connaissance est donc le suivant : les épistémologues tentent de donner une réponse moniste à la question de la nature de la connaissance. Une telle forme de réponse est inappropriée : le terme « connaissance » semble référer à des entités qui n'ont que très peu de propriétés communes, et qui en ce sens ne peuvent être étudiées comme une entité unique. La problématique du présent mémoire est la suivante : la solution au problème de la nature de la connaissance peut-elle être une solution moniste? Si non, comment pouvons-nous dépasser ce problème?

Nos objectifs sont doubles. Premièrement, nous argumentons contre l'idée que la solution au problème de la nature de la connaissance puisse être moniste. Deuxièmement, nous tentons de montrer qu'une manière prometteuse de dépasser, ou de résoudre le problème, est de considérer la possibilité que la réponse à la question de la nature de la connaissance puisse être une réponse qui repose sur une hypothèse d'hétérogénéité à l'endroit

de la connaissance. En d'autres termes, nous voulons montrer que pour expliquer adéquatement la connaissance, il est nécessaire d'abandonner le projet de trouver une nature unique de celle-ci et adopter une conception hétérogène qui reconnaisse 1) l'existence d'entités fondamentalement différentes qui peuvent et doivent être considérées comme des instances de connaissance en vertu de leur rôle explicatif et 2) l'apport important de différentes approches méthodologiques.

Il importe de bien comprendre que notre objectif n'est *pas* de défendre une version achevée d'une conception hétérogène de la connaissance. En effet, il est surtout question de poser les bases de cette conception, ce qui veut dire argumenter en faveur de l'idée que tant les énoncés observationnels ou théoriques de forme phrastique que les vecteurs d'activation neuronale, en passant par les états informationnels « primitifs » des animaux non-humains peuvent et doivent être considérés comme des connaissances. Ceci s'appuie sur le rôle explicatif de ces différentes entités, dont la postulation est une étape nécessaire dans la formulation d'explications comportementales, par exemple.

L'objectif, donc, est de démontrer qu'il y a une nécessité explicative qui motive et justifie l'idée de rejeter la forme moniste des réponses traditionnelles au problème de la nature de la connaissance. Nous développons également une première version, préliminaire et schématique, d'une conception hétérogène de la connaissance, considérée comme une solution alternative prometteuse à ce problème. En ce sens, notre problématique et nos objectifs sont de nature métathéorique. Rappelons également que si notre travail s'inscrit dans le contexte de l'épistémologie naturalisée, la classe des approches souscrivant au « paradigme phrastique » peut être étendue pour inclure les approches « traditionnelles » non-naturalistes. Notre argument fonctionne, *mutatis mutandis*, pour ces approches.

0.2. Méthodologie et démarche

Nous employons trois outils méthodologiques : l'analyse théorique, la recherche bibliographique et l'interprétation de résultats empiriques. L'analyse théorique est employée en conjonction avec la recherche bibliographique tout au long du mémoire. Dans le premier chapitre, ces méthodes servent à 1) identifier les conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer recevable et 2) la marche à suivre pour développer une hypothèse d'hétérogénéité. Au second chapitre, nous employons ces outils pour 1) construire une perspective métathéorique du débat sur la nature de la connaissance et 2) développer un argument de nécessité explicative reposant sur la validité d'une prémisse empirique. C'est ici qu'entre en jeu l'interprétation de données empiriques, qui est centrale au travail que nous accomplissons au chapitre III. Nous proposons, premièrement, l'idée qu'il existe des formes de connaissance qui reposent sur un critère de vérité fondamentalement différent du critère propre aux approches phrastiques. Ce critère est *pragmatique* et repose sur un argument évolutionniste énoncé par Tooby, Cosmides et Barrett (2005) et remanié à la lumière du travail de Goode et Griffiths (1995). Cette partie repose sur l'analyse théorique et la recherche bibliographique. Nous fournissons un support empirique à cet argument en interprétant les recherches de Gigerenzer (1999, 1999, 2009) en psychologie du raisonnement. Ensuite, nous discutons des travaux de Goodale et Milner sur la préservation des capacités de préhension de patients atteints d'agnosie aperceptive (1991, 1992, 1994) afin de montrer que la catégorie des connaissances non phrastiques n'est pas, elle non plus, homogène.

En ce qui concerne les détails de la démarche adoptée, nous devons d'abord faire quelques précisions. Par « perspective métathéorique », nous entendons ici une perspective portant sur les théories de la connaissance. L'aboutissement de notre projet est certes, à long terme, la formulation d'une théorie de la connaissance. Toutefois, le travail que nous accomplissons ici est un travail concernant les théories naturalistes de la connaissance et la forme de la réponse que celles-ci tentent de donner à la question de la nature de la

connaissance. En ce sens, plusieurs des termes et des notions que nous développons ici sont de nature métathéorique.

Les conditions contextuelles et méthodologiques dans lesquelles une hypothèse d'hétérogénéité peut s'avérer justifiée sont des notions métathéoriques que nous tirons de l'analyse de deux approches hétérogènes en philosophie de la psychologie, soit celle de Griffiths concernant les émotions (1997, 2004) et celle Machery concernant les concepts (2005, 2009).

La notion de « paradigme » phrastique est également une notion métathéorique. Elle désigne l'ensemble des approches naturalistes endossant l'idée que la connaissance est adéquatement décrite au niveau phrastique. De la même manière, l'idée du « débat sur la nature de la connaissance »¹ est une notion métathéorique qui réfère à une question générale, abordée par plusieurs théories. De plus, l'ensemble de notre travail se situe au niveau métathéorique, surtout lors du second chapitre, qui évalue la portée explicative de diverses théories.

Concernant le pendant empirique du présent mémoire, il s'agit d'une composante nécessaire en regard de l'engagement naturaliste que nous prenons. Cet engagement est simplement l'idée que l'épistémologie devrait s'intéresser au développement causal de la connaissance, tel qu'il se produit *dans les faits*. Churchland (1979, 1989, 2007) adopte cet engagement, tout comme le fait Quine (1969). Il s'agit d'un engagement très général, et le rôle qu'il joue ici est central. En regard des considérations conceptuelles et théoriques que nous développons au chapitre 2, cet engagement nous force à reconnaître que tant les approches phrastiques que l'approche réductionniste de Churchland font face à des problèmes d'insuffisance explicative. Notons également que malgré cet accent sur l'aspect empirique, nous ne tentons pas d'éliminer le pendant normatif de l'épistémologie. D'une part, la

¹ Cette expression doit être comprise comme désignant le débat sur le niveau descriptif approprié à l'étude de la connaissance.

conception pragmatique de la vérité (développée au chapitre III) fournit un critère normatif opératoire. D'autre part, la conception hétérogène de la connaissance que nous cherchons à développer, parce qu'elle reconnaît l'apport de diverses approches épistémologiques, reconnaît *de facto* l'importance du projet mélioratif de l'épistémologie. De plus, le pendant empirique est requis par la forme que prend l'argument qui sous-tend la proposition d'une conception hétérogène de la connaissance. Cet argument est l'argument de nécessité explicative énoncé au chapitre II. Pour bien apprécier son rôle dans le mémoire, il sera utile de le formuler dès maintenant :

Argument de nécessité explicative à l'endroit de la connaissance

1. Pour expliquer nos capacités épistémiques, nous devons postuler que nous possédons quelque chose comme la connaissance. (Prémisse 1)
2. Les constructions théoriques mises de l'avant par les théories phrastiques et non phrastiques² de la connaissance n'ont que très peu de propriétés communes. (Prémisse 2, venant de l'analyse métathéorique)
3. Afin d'expliquer nos capacités cognitives de manière satisfaisante, nous devons employer les deux sortes de constructions théoriques (Prémisse 3, venant de l'analyse métathéorique et de l'exposition des raisons conceptuelles en 2.3.2)
4. Donc l'explication de nos capacités épistémiques requiert d'avoir recours à des constructions théoriques n'ayant que très peu de propriétés communes. (Conclusion 1)
5. Donc, il n'existe pas de nature unique de la connaissance (ou : la connaissance n'est pas une espèce naturelle). (Conclusion 2)

Ce qui est important ici est de remarquer que la prémisse 3 est une prémisse empirique. Afin de déterminer si elle est valide ou non, il faut fournir non seulement des raisons conceptuelles

² Une autre version de l'argument pourrait inclure les approches non naturalistes dans les approches phrastiques.

mais également un support empirique solide. Pour comprendre comment nous arrivons à cet argument et avoir une meilleure idée de son rôle crucial dans notre entreprise, il convient de présenter la structure du présent mémoire.

0.3. Structure

Le premier chapitre présente deux exemples d'hypothèses d'hétérogénéité tirés de la philosophie de la psychologie. Le premier exemple est l'approche hétérogène de Griffiths concernant les émotions (Griffiths 1997, Griffiths et Scarantino 2009). Griffiths rejette le terme vernaculaire « émotion » (ainsi que les termes spécifiques comme « colère » sur la base du fait qu'il s'agit d'une instance de référence partielle : le terme « émotion » renvoie à des entités qui n'ont que peu de (ou aucune) propriétés communes. Comme le terme « jade » qui réfère autant à la néphrite qu'à la jadéite, il s'agit d'un terme qui n'est pas approprié à l'élaboration d'une psychologie scientifique des émotions. Le second exemple est l'approche de Machery concernant les concepts (Machery 2005, 2009). L'idée est sensiblement la même : « concept » renvoie à des entités n'ayant que peu de propriétés communes mais remplissant tout de même une fonction explicative importante. Ces deux approches reposent sur la théorie des espèces naturelles de Boyd (1991), que nous employons afin de rendre explicite le rôle des espèces naturelles dans la pratique scientifique, et surtout, le rôle des hypothèses d'espèce naturelle. Une hypothèse d'espèce naturelle est un procédé par lequel on fait l'hypothèse qu'un terme T du vocabulaire scientifique réfère à une espèce naturelle. Lorsqu'une telle hypothèse est erronée, nous voyons émerger certains signes (prolifération de débats sur la nature d'une entité fondamentale à la discipline, par exemple). En analysant les exemples d'hétérogénéité à la lumière de ces considérations sur le rôle et l'effet des hypothèses d'espèce naturelle, nous identifions deux conditions contextuelles et une condition méthodologique qui doivent être remplies pour qu'une hypothèse d'hétérogénéité puisse s'avérer justifiée. Ceci nous fournit une marche à suivre claire pour la suite du travail.

Le second chapitre commence par la définition des deux positions impliquées dans le débat sur la nature de la connaissance. La première, que nous appelons « paradigme phrastique » est abordée par le biais de son émergence et de sa prépondérance au sein de l'épistémologie naturalisée. Cette prépondérance est attribuable à plusieurs facteurs, que nous discutons en détail dans la première section du chapitre. Nous procédons ensuite à la présentation de la seconde position, soit le matérialisme éliminativiste de Churchland (1979, 1992, 2007)³. Nous présentons d'abord les grandes lignes et les engagements théoriques de Churchland. Il est ensuite question des arguments qu'il développe en faveur de l'élimination du paradigme phrastique. Nous construisons ensuite une perspective métathéorique du débat sur la nature de la connaissance, perspective qui met en évidence le fait que ces deux positions voient la solution au problème de la nature de la connaissance comme étant moniste (il n'existe qu'une seule nature de la connaissance). Nous montrons ensuite que ce débat remplit les conditions contextuelles applicables à l'élaboration d'une hypothèse d'hétérogénéité et que nous disposons de raisons conceptuelles de rejeter la solution moniste. En effet, après évaluation des arguments de Churchland, ceux-ci sont jugés convaincants *uniquement* dans la mesure où ils montrent que l'approche phrastique est insuffisante. La position de Churchland est elle aussi insuffisante parce qu'elle néglige l'apport du langage. L'explication de cet apport est une tâche toute désignée pour les approches phrastiques. Sur la base de ces considérations, nous formulons l'argument de nécessité explicative présenté plus haut.

Le troisième chapitre a plusieurs objectifs. Premièrement, il faut fournir un support empirique à la prémisse 3 de l'argument de nécessité explicative (pour expliquer nos capacités cognitives de manière satisfaisante, nous devons employer les approches phrastiques et non phrastiques). Nous commençons par établir une conception minimale de la connaissance, sous la forme de trois conditions relativement faibles : les structures

³ Nous désignons parfois sa position comme « réductionniste ». Lorsque cette expression est employée, elle doit être comprise comme « réductionnisme éliminativiste ».

informationnelles candidates à l'attribution du statut de « connaissance » doivent 1) entretenir un lien avec le monde, 2) jouer un rôle dans l'explication comportementale et 3) être produites de manière fiable. Le résultat est une conception de la connaissance qui correspond à la notion d'information en sciences cognitives : soit les signaux d'entrée et de sortie d'un système cognitif. Il s'agit en fait de l'un des apports intéressants de la perspective que nous développons : dans le cadre moniste que nous critiquons, il était fort difficile d'intégrer la notion d'information qui, pourtant, est employée comme une notion épistémique en sciences cognitives. Le modèle que nous proposons intègre facilement la notion. Il faut noter que cette conception n'est pas substantive, elle est heuristique : elle nous permet simplement d'identifier les formes les plus fondamentales de connaissance. À cela s'ajoutent des considérations sur la notion de justification : les connaissances de plus « bas niveau » doivent remplir des critères plus simples que des formes plus complexes de connaissance. Ceci nous fournit un premier axe d'hétérogénéité ou de différenciation; les critères justificatifs pertinents peuvent nous aider à différencier diverses formes de connaissance. Nous développons ensuite un argument évolutionniste supporté par des recherches en psychologie du raisonnement pour mettre en évidence un second axe de différenciation (critère de vérité). Finalement, nous présentons certains résultats de la neuropsychologie de la vision pour mettre en évidence un troisième axe de différenciation (forme phrastique ou non).

Nous faisons ensuite la synthèse de ces différentes considérations en rappelant l'argument de nécessité explicative et comment les considérations empiriques présentées fournissent un support empirique à la prémisse 3 de l'argument. Le résultat est que notre argument de nécessité explicative est validé. Il en découle que la solution moniste au problème de la nature de la connaissance doit être rejetée. Nous proposons ensuite une première ébauche d'une conception hétérogène de la connaissance, comme solution alternative au problème de la nature de la connaissance.

CHAPITRE I

HÉTÉROGÉNÉITÉ: EXEMPLES ET PARAMÈTRES

Ce premier chapitre présente les fondements des thèses avancées dans le mémoire. L'objectif est de clarifier l'inspiration principale du projet et les conditions dans lesquelles il peut s'avérer justifié de proposer une hypothèse d'hétérogénéité. Nous présentons deux approches hétérogènes tirées de la philosophie de la psychologie. Ces exemples sont ensuite analysés afin de rendre explicites les éléments suivants : 1) les notions d'espèce naturelle (*natural kind*) et d'hypothèse d'espèce naturelle, 2) leur rôle en science (ainsi que les effets associés à l'adoption d'une hypothèse d'espèce naturelle) et 3) les conditions dans lesquelles une hypothèse d'hétérogénéité (comprise comme remettant en question l'adoption d'une hypothèse d'espèce naturelle jugée erronée) peut s'avérer justifiée.

1.1. Hétérogénéité : exemples

Cette section présente deux exemples d'hypothèses d'hétérogénéité. L'objectif est de mettre en évidence les éléments qui nous permettront par la suite d'identifier les conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer justifiée. Conséquemment, nous n'allons pas discuter plus que nécessaire du détail des considérations empiriques mises de l'avant par les auteurs (nous en présentons évidemment les grandes lignes, ainsi que les détails cruciaux à la bonne compréhension de leurs argumentaires).

Les exemples discutés sont l'élimination des catégories vernaculaires « émotion » et d'émotions particulières (« colère », « joie » et « peur », par exemple) du vocabulaire scientifique (Griffiths 1997, 2004, Griffiths et Scarantino 2005) et l'éliminativisme scientifique de « concept » (Machery 2005, 2009). Nous commençons par présenter la

position de Griffiths et ce parce qu'il développe certaines considérations sur lesquelles s'appuie Machery.

1.1.1. Griffiths : les espèces naturelles

Afin de bien comprendre l'éliminativisme de Griffiths, il est nécessaire de présenter sa position sur les espèces naturelles et leur rôle dans la pratique scientifique. L'objectif de Griffiths (1997) est de porter un regard critique sur le concept vernaculaire d'émotion, ainsi que sur le concept d'émotion compris comme une espèce naturelle, soit une catégorie qui correspond à des distinctions naturelles permettant l'élaboration d'une théorie.

Son point de départ est que les émotions sont le référent des termes employés par les théories qui étudient les phénomènes émotifs. Pour expliquer adéquatement ces phénomènes il est nécessaire de réviser les concepts traditionnels d'émotion au sens large et les concepts désignant des émotions particulières. Cette révision vise à regrouper des éléments similaires et discriminer les éléments dissimilaires : il n'y a parfois aucune raison valable (du point de vue explicatif) de regrouper sous une même catégorie des éléments tombant sous un concept vernaculaire. Par exemple, la catégorie des objets supra lunaires n'offre aucun gain explicatif, parce qu'il s'agit d'une catégorie purement arbitraire, comme celle des objets plus lourds que la tasse de café de Jim (Griffiths 1997, p. 171). Le type de révision conceptuelle que propose Griffiths n'est pas inusité; au contraire, il est commun lorsque sont découvertes de nouvelles données de réviser certains concepts (les espèces préscientifiques « feu » et « chaleur » deviennent les phénomènes décrits par les sciences physiques, par exemple). Les concepts sont radicalement modifiés, voire abandonnés comme catégories scientifiques. Ceci est souvent la conséquence du fait que ce type de révision conceptuelle découle d'une découverte du fait que sous ces concepts se cachent des phénomènes fort différents. Il s'agit précisément de ce que nous tentons de mettre en lumière en ce qui a trait à la connaissance.

Ce type de révision conceptuelle est devenu une préoccupation centrale de la philosophie des sciences et du langage suite au rejet des théories descriptives de la signification des termes théoriques. Ces théories ne peuvent accommoder la révision conceptuelle suite à une

découverte empirique : la conceptualisation du référent d'un terme ne peut être erronée parce qu'elle définit ce qui peut être compté comme un référent de ce terme (*ibid*, p. 171)⁴. Ce type de position est en fait commun en philosophie des émotions (voir note ici-bas). Griffiths (1997) argumente plutôt que comprendre ce qu'est réellement la peur, par exemple, demande d'expliquer les états psychologiques auxquels le terme « peur » réfère, et non de s'appuyer sur une analyse des croyances concernant la peur (*ibid*, p. 172). Griffiths soutient donc que la vraie nature des émotions doit être révélée par les théories qui expliquent le mieux les phénomènes émotionnels, même si cela implique d'abandonner les conceptions traditionnelles.

Cette idée est cohérente avec les théories causales de la référence développées entre autres par Putnam (1975) et Kripke (1980). Les termes d'espèce (*kind terms*) du langage et les concepts employés par les gens ont une relation référentielle aux arbres, aux chiens ou aux cellules. La théorie causale de la référence vise à remplacer, pour certains concepts, la vision descriptive par une vision causale de la relation référentielle : notre concept d'arbre de est approprié à cause de la séquence causale appropriée au fait d'être « à propos des arbres ». Cette théorie permet de s'attaquer à la question de la révision conceptuelle suite à des découvertes empiriques. Le « stéréotype » d'un terme est l'ensemble des croyances concernant son référent. Ces croyances déterminent d'un ensemble d'instances qu'il est l'extension du terme. Les découvertes empiriques peuvent changer le stéréotype du terme, changeant ainsi son extension (Griffiths 1997, p. 172). Griffiths note que les premiers tenants de la théorie causale de la référence étaient fortement réalistes, ce qui est reflété dans leur traitement des espèces naturelles.

⁴ Davis (1988) soutient par exemple que pour découvrir la nature de la peur, il faut analyser la compréhension qu'ont les locuteurs ordinaires du terme « peur », parce qu'il considère que la signification de « peur » correspond au contenu descriptif du terme pour les locuteurs ordinaires. Conséquemment, seulement une chose correspondant à ce contenu descriptif peut être de la peur (Griffiths 1997, p. 171)

L'idée centrale de cette théorie des espèces naturelles est qu'il existe un mécanisme sous-jacent expliquant les corrélations entre propriétés qui nous permettent de classifier les choses en espèces (*kinds*) distinctes (*ibid*, p. 173). Pour prendre le mécanisme explicatif de prédilection des réalistes, l'on peut dire par exemple que parce que les choses dans une catégorie C ont la même essence (explication sous-jacente), elles ont des propriétés qui sont corrélées entre elles. Ceci implique également que lorsque sont découvertes de nouvelles données sur un membre de C, celles-ci peuvent être projetées sur les autres membres de manière fiable (*ibidem*). Ces catégories sont des espèces naturelles, et leur existence permet de résoudre le problème de l'induction⁵ : ces catégories peuvent supporter l'induction. Comme leurs membres ont la même essence, leurs propriétés sont corrélées les unes aux autres et nous sommes donc justifiés de projeter les données concernant un membre d'une espèce naturelle sur d'autres membres (*ibidem*). Par exemple, du fait qu'un rubis R a une dureté de 9 sur l'échelle de Mohs, nous pouvons conclure qu'un rubis A' aura la même dureté, parce qu'ils ont la même essence.

Le réalisme se heurte toutefois au problème de la dépendance théorique des observations : Putnam (1994) rejette le réalisme parce qu'il reconnaît que les catégories ne peuvent être totalement indépendantes des concepts théoriques qui les désignent (voir Griffiths 1997, p. 173 et Putnam 1975). L'abandon du réalisme implique que les espèces auparavant comprises comme indépendantes de la théorie ne sont plus disponibles pour ancrer les concepts théoriques suite à une révision conceptuelle (en l'absence d'essences indépendantes de la théorie, la relation référentielle n'est plus préservée).

Boyd (1991) argumente toutefois que la théorie des espèces naturelles peut être désengagée de la notion d'essences. Pour lui, la postulation d'espèces naturelles est nécessaire à la compréhension de certains aspects de la pratique scientifique (nous revenons sur ce point en 1.2.1). Les scientifiques ont de bonnes raisons de vouloir utiliser des concepts

⁵ Il s'agit, à tout le moins, d'une solution proposée au problème de l'induction.

qui supportent la généralisation (*projectables*) parce que seuls ceux-ci peuvent être utiles à l'explication et l'induction. Le potentiel de généralisation (*projectability*) d'un concept est jugé sur la base de théories d'arrière-plan et les concepts sont construits de manière à référer à des catégories jouant un rôle dans les meilleures théories du domaine. Donc, plutôt que dire que le progrès scientifique dépend de la mise en correspondance de nos concepts à la structure réelle du monde, Boyd adopte une perspective empiriste selon laquelle les théories empiriquement adéquates sont celles qui nous fournissent les données nécessaires à la construction de catégories utiles. Pour anticiper sur la suite, le réalisme de Boyd consiste à maintenir l'idée que les termes théoriques réfèrent à des entités réelles dans le monde tout en rejetant l'idée que cette relation référentielle est due à des « essences ».

Boyd (1989, 1991, voir également Keil 1989) emploie la notion d'homéostasie causale pour expliquer l'utilité d'une catégorie constituant une espèce naturelle. Une catégorie scientifiquement pertinente regroupe des objets dont les propriétés sont corrélées. La catégorie C est en homéostasie causale si et seulement si l'ensemble de corrélations qui la compose possède une explication sous-jacente qui fait en sorte que ces corrélations sont généralisables (*projectable*). Une catégorie scientifiquement pertinente, selon Keil, en est une qui capture un mécanisme d'homéostasie causale, soit une explication qui fait en sorte que les corrélations observées sont stables, au sens où nous pouvons, avec une certaine fiabilité, supposer qu'elles valent également dans le cas d'instances encore non observées (Griffiths 1997, p. 187)⁶.

Le concept d'homéostasie causale fournit donc un critère permettant d'identifier les espèces naturelles. Les espèces fondamentales⁷ de la chimie (éléments du tableau périodique)

⁶ Deux remarques sur le concept d'homéostasie causale : 1) il implique une définition large de l'« essence » d'une catégorie, parce qu'elle peut être n'importe quelle structure théorique qui explique le potentiel de généralisation d'une catégorie (*projectability*) et 2) nous discuterons plus en détail de ce concept lorsqu'il sera question de la position de Machery.

⁷ À ne pas confondre avec ce que Machery appelle les espèces fondamentales, qu'il distingue des espèces naturelles par le fait qu'elles sont des catégories regroupant des objets ou entités n'ayant qu'un

sont des catégories causalement homéostatiques parce que les corrélations de leurs propriétés sont expliquées par leur structure interne (qui ici est le mécanisme d'homéostasie causale). Les taxons biologiques sont des catégories causalement homéostatiques dont le mécanisme causal est la descendance commune (*ibid*, p. 188). À l'opposé, la catégorie « lys » n'est pas une espèce naturelle : elle dénote une variété de taxons ayant des propriétés variées et ne regroupe rien sinon une vague ressemblance visuelle (*ibidem*).

Nous disposons à présent des éléments nécessaires pour aborder directement la position de Griffiths sur les émotions. Sa thèse est que la catégorie vernaculaire « émotion » doit être éliminée du vocabulaire scientifique dans le contexte de l'explication des phénomènes émotionnels.

1.1.2. Griffiths et les émotions : élimination du concept

Griffiths part donc de l'idée que le développement d'une psychologie scientifique des émotions requiert que les concepts d'émotion soient révisés (ou simplement remplacés) de manière à ce que les catégories auxquelles réfèrent ces concepts aient une forte homéostasie causale, parce que de telles catégories sont généralisables (*projectable*) et donc utiles à l'explication et l'induction (Griffiths 1997, p. 227). Il examine deux manières de catégoriser les émotions qui pourraient répondre de manière satisfaisante (en identifiant les concepts d'émotion à des catégories causalement homéostatiques) à la question de ce que sont les émotions.

La première est de les identifier aux catégories dérivées de la théorie des programmes d'affect (*affect program theory*). Cette théorie est étroitement liée aux recherches de Darwin ([1872]1965) sur les expressions faciales des émotions chez les humains et, surtout, aux

nombre limité de propriétés en commun, par opposition aux espèces naturelles dont les membres partagent plusieurs propriétés (Machery 2005).

recherches plus contemporaines sur le sujet. Ekman (1971, voir également Ekman et Friesen 1971) montre que certaines expressions faciales associées à diverses émotions se retrouvent dans des cultures différentes. Ces similarités ne peuvent être d'origine culturelle, car elles se retrouvent même dans des cultures pré-littéraires (donc coupées d'une quelconque culture occidentale ou européenne, par exemple) et géographiquement distancées (n'appartenant pas au même ensemble culturel). Pour expliquer ces découvertes, Ekman développe la théorie des programmes d'affect, selon laquelle les réponses émotionnelles sont complexes, coordonnées et automatiques (Griffiths 1997, p. 77). Elles sont *complexes* parce qu'elles impliquent plusieurs éléments (changements faciaux expressifs, réponses locomotrices, changements vocaux expressifs, changement endocriniens, changements dans le système nerveux). Elles sont *coordonnées* car ces éléments se produisent ensemble, dans une séquence reconnaissable et elles sont *automatiques* parce qu'elles se produisent (de manière coordonnée) sans contrôle conscient. Le *programme d'affect* est cet ensemble coordonné d'événements physiologiques qui constitue la réponse émotionnelle. Ekman identifie six programmes d'affect particuliers (spécifiques aux humains) : colère, peur, surprise, joie, tristesse et dégoût (*ibid*, p. 78) et argumente en faveur d'une identification des émotions à ces entités théoriques : de la même manière que « eau » est *en fait* H₂O, « colère » est *en fait* le programme d'affect correspondant.

Les programmes d'affect semblent de bons candidats référentiels pour les termes d'émotion. Griffiths soulève que les émotions sont des phénomènes biologiques *et* psychologiques, et qu'il est bien connu qu'il existe trois niveaux d'explication en psychologie. Ces trois niveaux, identifiés par Marr (1982)⁸ sont: 1) le niveau computationnel (*task level* ou *ecological level*, qui décrit la tâche accomplie par le système psychologique), 2) le niveau algorithmique (qui décrit les processus de traitement de l'information qui

⁸ Griffiths (1997) présente ces trois niveaux comme étant tâche, computationnel et implémentation. La terminologie française correspondante est computationnel, algorithmique et implémentation. Les niveaux en italiques sont tels que désignés par Griffiths.

permettent au système d'accomplir la tâche) et 3) le niveau de l'implémentation (qui décrit la manière dont les processus computationnels sont implémentés dans le cerveau). Les programmes d'affect sont également des traits biologiques, qui peuvent être décrits aux trois niveaux explicatifs biologiques⁹ : il s'agit de traits évolués, servant probablement à faciliter la communication interspécifique entre humains (niveau écologique général), le programme d'affect « peur » peut être classifié comme un homologue de la peur chez d'autres primates et il est possible de classifier les programmes d'affect par leurs particularités physiques. Donc, les programmes d'affect permettent de former des catégories utiles aux trois niveaux d'explication biologique *et* psychologique.

Niveaux d'explication biologiques	Niveaux d'explication psychologiques
Écologique général : traits classifiés par problème adaptatif solutionné/rôle adaptatif (communication interspécifique)	Computationnel : description de la tâche accomplie (communication d'une réaction)
Histoire naturelle (<i>natural historical level</i>) : traits classifiés par homologie (programme d'affect « peur » est homologue à peur chez d'autres primates)	Algorithmique : Processus de traitement de l'information (réponses fonctionnelles évolutivement anciennes)
Anatomique : traits classifiés par caractéristiques physiques (ensemble des éléments physiques des programmes d'affect)	Implémentation : de quelle manière ces processus sont implémentés dans le cerveau (changements endocriniens, production d'hormones, effet neurologique)

Tableau .1.1 : Correspondances des niveaux explicatifs biologiques et psychologiques. Entre parenthèse sont les caractéristiques des programmes d'affect pertinentes à ce niveau d'explication. Tableau adapté de Griffiths 1997, p. 221, informations sur les programmes d'affect tirées du chapitre 9, section 2 (Griffiths 1997).

Le tableau 1.1, il faut le relever, illustre un parallèle relativement lâche entre les niveaux explicatifs de la biologie et ceux de la psychologie. Il montre toutefois que la théorie des programmes d'affect offre une explication assez complète de certains phénomènes

⁹ Voir Griffiths 1997, chapitre 8, section 2.

émotionnels. Puisque les programmes d'affect sont des réponses automatiques évolutivement anciennes, la théorie explique également les phobies et autres conflits entre les capacités cognitives de haut niveau (conscientes et impliquant des croyances explicites) et certaines émotions. Par exemple, elle explique pourquoi certaines réactions émotionnelles provoquent une dissonance avec nos croyances explicites¹⁰.

Le pouvoir explicatif de la théorie des programmes d'affects pourrait mener à la considérer comme une théorie des émotions de la colère, de la surprise, de la joie, etc. Mais cette identification des concepts vernaculaires aux concepts des programmes d'affect impliquerait une révision conceptuelle importante : identifier « colère » à un programme d'affect distinguerait les instances « réelles » de colère (les instances de programme d'affect « colère ») d'autres états n'impliquant pas les mêmes changements physiologiques mais tombant sous le concept vernaculaire de « colère ». Ceci implique d'exclure des réactions comme la désapprobation ou les résolutions à long terme associées à la vengeance ou à l'idée de « redresser un tort » de la classe des émotions. En d'autres termes, identifier les termes d'émotion aux programmes d'affect implique d'exclure une large classe de réactions émotionnelles de plus haut niveau¹¹, qui font indéniablement partie des phénomènes

¹⁰ Une idée semblable est développée par Gendler (2008a, 2008b) pour expliquer les émotions récalcitrantes comme la peur de tomber lorsque Jim marche sur un plancher de verre situé à haute altitude même s'il *sait* explicitement que la surface est sans danger. L'idée est qu'en plus de sa croyance occurrente que la surface est sans danger, il a un *alief*, soit une « croyance » implicite provoquée de manière automatique par un stimulus particulier (dans ce cas-ci, une quantité impressionnante de vide sous ses pieds).

¹¹ Griffiths n'est pas clair sur la nature exacte de ces phénomènes émotionnels de haut niveau. Il est toutefois relativement peu controversé de reconnaître que les réactions émotionnelles ne sont pas *uniquement* des réactions physiologiques rapides et automatiques, mais qu'il existe une classe de phénomènes émotionnels à plus long terme (voir Griffiths 1997, chapitre 5).

émotionnels qu'une psychologie scientifique devrait expliquer (Griffiths 1997, p. 215). Bref, la théorie des programmes d'affect n'explique qu'une partie des phénomènes émotionnels¹².

Il faut trouver une autre explication aux phénomènes émotionnels de haut niveau. Griffiths et Scarantino (2005) proposent une perspective situationniste des émotions¹³, qui pourrait combler ce manque explicatif. Cette perspective fournit une explication des phénomènes émotionnels laissés de côté par la théorie des programmes d'affect (sans toutefois s'appliquer uniquement aux réactions émotionnelles de haut niveau, ce qui ne réduit pas le pouvoir explicatif de la théorie des programmes d'affect parce que celle-ci offre une explication adéquate d'une classe particulière de réactions émotionnelles).

Rappelons que l'objectif de Griffiths est de répondre à la question « que sont les émotions? ». Pour expliquer convenablement l'émotion il faut parvenir à identifier les concepts d'émotion à des catégories causalement homéostatiques. La manière la plus intéressante de procéder est d'identifier ces concepts aux catégories des programmes d'affect, mais les programmes d'affect n'expliquent qu'une partie des phénomènes émotionnels regroupés sous les concepts vernaculaires. La partie qui n'est pas expliquée par cette théorie doit être expliquée autrement (soit à l'aide de la perspective situationniste développée par

¹² La seconde manière de catégoriser les émotions est de les identifier aux catégories dérivées d'une explication écologique. L'idée est simple : il s'agit de catégoriser les émotions sur la base de leur fonction adaptative. En adoptant cette base de catégorisation plus générale, il *serait* possible d'inclure les processus cognitifs de haut niveau, laissés de côté par la théorie des programmes d'affect. Griffiths rejette toutefois cette idée, et ce pour trois raisons : 1) cette approche ne peut expliquer les phénomènes émotionnels qu'au niveau de la description de la tâche et est impuissante aux niveaux computationnel et de l'implémentation, 2) les explications écologiques de l'émotion sont encore à formuler et ne le seront peut-être jamais et 3) même si de telles explications voient le jour, rien ne garantit que les catégories écologiques correspondraient mieux aux concepts vernaculaires (*ibid*, p. 228). Les programmes d'affect offrent une meilleure alternative, en dépit de leurs limites explicatives.

¹³ Au sens de « cognition située ». L'idée générale est que les émotions sont construites de manière à fonctionner dans un contexte social (une émotion est souvent une tentative de reconfigurer une relation sociale par le biais d'un signal social), sont une forme d'engagement avec le monde qui n'a pas besoin d'être médiatisée par une pensée conceptuelle, dépendent de l'environnement et y sont couplées par le fait qu'elles sont influencées et influencent l'environnement (Griffiths et Scarantino 2005).

Griffiths et Scarantino, 2005), et donc les concepts qu'on ne peut identifier aux catégories des programmes d'affect devront être identifiés à d'autres catégories, issues d'une autre théorie. La conséquence est la suivante : le terme « émotion » et les termes d'émotion spécifiques sont des cas de *référence partielle* (Griffiths 2004). L'exemple classique de référence partielle est le terme « jade ». Ce terme fait référence à l'un ou l'autre de la néphrite ou de la jadéite. Le terme fait partiellement référence à deux sortes de minéraux et donc, pour les fins de la géologie et de la chimie, « jade » ne peut être traité comme une seule chose. Il en va de même avec le terme « émotion » et les termes d'émotion spécifiques. « Colère » fait partiellement référence au programme d'affect « colère » et à des instances de « colère » différentes requérant une autre explication. Pour les fins d'une psychologie scientifique des émotions, « colère » (comme « émotion ») ne peut être traité comme référant à une seule sorte de chose. Tout comme il n'y a pas de telle chose que le jade (au sens du vocabulaire scientifique), il n'y a pas de telle chose que des émotions : il y a des programmes d'affects, des biais motivationnels, des attitudes socialement conditionnées, etc. (*ibid*, p. 902). Dans le cadre d'une psychologie scientifique des émotions, les concepts vernaculaires doivent être éliminés.

Ainsi, Griffiths énonce une hypothèse d'hétérogénéité à l'endroit des émotions en se basant sur une certaine définition des espèces naturelles (théorie de l'homéostasie causale), de leur rôle en science (fournir des catégories généralisables utiles à l'explication et l'induction) et sur une analyse des explications disponibles des phénomènes émotionnels (théorie des programmes d'affect, perspective située, etc.). Son constat est que le terme « émotion » fait référence à des catégories hétérogènes et que considérer « émotion » comme une seule chose est une erreur parce que le terme est un cas de référence partielle (ce qui implique qu'il s'agit d'une catégorie scientifiquement inutile). L'ensemble des états psychologiques et processus neurologiques regroupés sous « émotion » sont trop dissimilaires pour permettre une psychologie unifiée des émotions, il faut reconnaître que ce terme réfère à des catégories hétérogènes et développer des explications particulières de ces catégories. Nous pouvons à présent passer à la position de Machery sur les concepts.

1.1.3. Machery et les concepts

La position de Machery (2005 2009)¹⁴ est, structurellement, semblable à celle de Griffiths. Il s'agit de montrer que « concept » réfère en réalité à des catégories hétérogènes, en se basant sur la théorie de l'homéostasie causale des espèces naturelles pour spécifier les critères de définition d'une catégorie. Le point de départ de Machery est le suivant : 1) les concepts sont les structures informationnelles (*data structures*) qui sont emmagasinées dans la mémoire à long terme et employées par défaut par les processus cognitifs de haut niveau et 2) les psychologues des concepts supposent que ces structures, ou représentations, partagent plusieurs propriétés scientifiquement pertinentes et que la psychologie des concepts doit décrire ces propriétés. En bref, il affirme que les psychologues des concepts supposent que les concepts sont une espèce naturelle (Machery 2005, p. 1). Machery appelle cette présupposition « *natural kind assumption* (ou *NK assumption*) ». Nous adopterons l'expression « PEN » (présupposition d'espèce naturelle).

Il convient de détailler la PEN. Selon Machery, les sciences cognitives et la psychologie supposent que les concepts forment une classe homogène de représentations mentales : CHAT et CHIEN¹⁵ sont deux concepts différents mais il est supposé qu'au-delà de leurs différences, ils partagent plusieurs propriétés scientifiquement pertinentes. En accord avec cette idée, la tâche de la psychologie des concepts est de mettre en évidence ces propriétés. De plus, si il est le cas que les propriétés des concepts expliquent une partie des propriétés des processus cognitifs de haut niveau dans lesquels ils sont employés, alors une théorie adéquate des concepts est un prérequis à toute théorie cognitive. La PEN, donc, est l'idée que les concepts sont une espèce naturelle.

¹⁴ Puisque nous sommes intéressés à la thèse de l'hétérogénéité des concepts, nous allons surtout discuter de l'article de 2005, qui a l'avantage de présenter la position de Machery de manière plus schématique.

¹⁵ Nous suivons la convention de nommer les concepts en utilisant des majuscules.

Il faut à présent savoir en quel sens les concepts sont des espèces naturelles. Nous avons discuté en détail des propriétés des espèces naturelles dans la section précédente, et Machery adopte une définition analogue à celle de Griffiths : les espèces naturelles sont des catégories, ou des ensembles d'objets, dont les membres partagent, en vertu de mécanismes causaux, des propriétés corrélées, ce qui fait de ces catégories les éléments fondamentaux des généralisations scientifiques. C'est pourquoi, comme le relève Machery, les sciences empiriques visent à identifier les espèces naturelles de leur domaine : pour être en mesure de développer des théories empiriques adéquates, et la psychologie n'y fait pas exception (Machery 2005, p. 4, voir également Griffiths 1997 chapitre 7, section 2).

Il est ensuite question de déterminer quelle théorie des espèces naturelles est appropriée pour les fins de la discussion sur les concepts. Machery établit la condition suivante : la théorie adoptée doit être assez large pour accommoder les espèces psychologiques (*psychological kinds*), autrement sa thèse selon laquelle les concepts ne sont pas des espèces naturelles serait triviale. Machery rejette d'emblée la théorie essentialiste (Putnam 1975, voir aussi Putnam 1994 pour un désaveu de l'essentialisme) selon laquelle les membres d'une espèce naturelle sont caractérisés par la possession d'une essence commune. Cette théorie exclut *de facto* les espèces psychologiques, du moins dans un contexte où les théories fonctionnalistes ont encore une certaine légitimité (il est certes difficile de voir comment caractériser les espèces psychologiques autrement que par leur fonction, qui, comme nous l'avons exposé dans la section précédente, font partie des trois niveaux explicatifs en psychologie). La seconde théorie rejetée par Machery est la théorie nomologique des espèces naturelles : les termes d'espèce naturelle (*natural kind terms*) figurent dans des lois comprises comme des généralisations sans restrictions spatiales ou temporelles (*spatially and temporally unrestrcited*) et supportant des contrefactuels (Collier 1996). Machery affirme, suivant Fodor (1974), que les espèces psychologiques ne figurent pas dans des lois (mais seulement dans des généralisations *ceteris paribus*). La théorie nomologique est donc rejetée parce qu'elle ne peut accommoder les espèces psychologiques.

Machery adopte finalement (ce qu'il appelle) la notion causale des espèces naturelles, qui correspond à la théorie de l'homéostasie causale de Boyd (1991).

« A class C of entities is a natural kind if and only if there is a large set of scientifically relevant properties such that C is the maximal class whose members tend to share these properties because of some causal mechanism. » (Machery 2005, p.8)

En d'autres termes (et pour rendre explicite la parenté entre les notions d'espèce naturelle adoptées par Machery et Griffiths), une classe, ou catégorie, C d'entités est une espèce naturelle si et seulement si les (nombreuses) propriétés que ses membres partagent sont expliquées par un même mécanisme causal¹⁶. Cette définition permet de faire des espèces psychologiques des catégories sur la base desquelles il est possible de faire des généralisations inductives non accidentelles. Notons également que Machery affirme explicitement adopter cette théorie (*ibidem*).

Nous pouvons à présent revenir à la question de la PEN. Celle-ci implique, en bref, que les structures informationnelles employées par les processus cognitifs de haut niveau ont trois caractéristiques importantes : 1) il existe une large classe de propriétés que ces structures tendent à posséder, 2) elles possèdent ces propriétés à cause d'un mécanisme causal quelconque et 3) cet ensemble de propriétés est spécifique à cette catégorie de structures informationnelles (Machery 2005, p. 10). Évidemment, les psychologues ne sont pas engagés à dire qu'il n'existe pas de différences entre certains types de concepts, l'idée étant que malgré ces différences, ils partagent assez de propriétés pertinentes. La présupposition d'espèce naturelle a donné lieu au développement de ce que Machery appelle « la perspective standard » (*received view*) : l'idée que les concepts constituent une classe homogène à propos de laquelle il est possible de formuler des généralisations (Machery 2005, p. 10, voir

¹⁶ Machery mentionne l'objection selon laquelle cette définition devrait inclure des catégories dont les membres ne possèdent que quelques propriétés fondamentales en commun. Il soulève premièrement qu'il reste incertain que ces catégories existent réellement et, deuxièmement, que même s'il en existait, elles auraient des propriétés explicatives fort différentes des espèces naturelles paradigmatiques. En ce sens, il serait plus judicieux de les distinguer des espèces naturelles (Machery 2005 p. 9).

également Machery 2009 pp.53-56). Cette idée est adoptée entre autres par Barsalou (1999), qui argumente que les la majorité (sinon la totalité) des concepts sont des symboles perceptuels et Prinz (2002), qui affirme que la majorité (sinon la totalité) des concepts sont des proxotypes (forme particulière de symboles perceptuels).

Il n'est pas intéressant ici d'entrer dans les détails des théories particulières des concepts. Nous l'avons fait dans le cas de l'émotion (voir article 1.1.2) parce que nous avons besoin d'avoir une idée plus précise de la façon dont une hypothèse d'hétérogénéité est développée. Comme nous le verrons dans ce qui suit, Machery opère sensiblement de la même manière que Griffiths : en analysant le pouvoir explicatif de diverses théories de manière à montrer qu'aucune n'offre d'explication satisfaisante de l'ensemble des phénomènes à expliquer. L'important ici est de comprendre que la perspective standard a motivé la méthodologie de la psychologie des concepts de manière cruciale. Les psychologues, partant de l'idée que les concepts sont un ensemble homogène, ont tenté de trouver des données empiriques aisément expliquées par leur théorie des concepts particulière et, surtout posant des problèmes explicatifs aux théories concurrentes (Machery 2005, p. 11). Le résultat est une prolifération de théories partielles des concepts (pour un survol de ces théories, voir Machery 2009, chapitres 1 et 2 ainsi que chapitre 3, section 3)¹⁷.

L'argument de Machery est, à partir de ce moment, relativement simple. Compte tenu de la notion causale (théorie de l'homéostasie causale), la PEN ne peut être vraie parce qu'il n'est pas le cas que les structures informationnelles utilisées par défaut dans les processus cognitifs de haut niveau possèdent plusieurs propriétés communes expliquées par un même mécanisme causal. La classe des concepts est divisée en plusieurs espèces (*kinds*) de représentations qui ont peu en commun (hypothèse d'hétérogénéité des concepts). Nous avons plusieurs concepts de chien, chacun appartenant à une espèce naturelle particulière. CHIEN₁, CHIEN₂ et CHIEN₃ sont, par exemples, employés pour la catégorisation, le

¹⁷ Cette situation est analogue à celle qui se retrouve, selon nous, en épistémologie.

raisonnement inductif, et le raisonnement déductif, respectivement (*ibid*, p. 12). Chacun de ces concepts est employé par défaut dans divers processus cognitifs. Cette conception est évidemment en opposition totale à la perspective standard, parce qu'elle soutient que nos processus cognitifs ne sont *pas* dépendants d'un ensemble homogène de représentations mentales. L'idée de Machery est qu'au lieu de tenter (en vain) de caractériser les propriétés des concepts, les psychologues devraient s'atteler à caractériser les espèces (*kinds*) fondamentales de concepts (qui, selon Machery, correspondent, au moins, aux exemplaires, théories et prototypes). L'argument de Machery est présenté ci-dessous¹⁸.

Argument de la nécessité explicative à l'endroit des concepts

1. L'existence des concepts est supposée afin d'expliquer les propriétés des processus cognitifs de haut niveau. (Prémisse 1)
2. Les constructions théoriques mises de l'avant par les principales théories des concepts n'ont que peu de propriétés communes. (Prémisse 2)
3. Afin d'expliquer les propriétés de nos processus cognitifs de haut niveau, nous devons supposer que ces constructions théoriques sont utilisées par défaut dans nos processus cognitifs de haut niveau. (Prémisse 3)
4. Donc, nos processus cognitifs utilisent plusieurs espèces (*kinds*) de représentations n'ayant que peu de propriétés communes. (Conclusion 1)
5. Donc, les concepts ne sont pas des espèces naturelles. (Conclusion 2)

L'idée centrale est que les concepts sont des entités théoriques postulées pour permettre d'expliquer les propriétés de nos processus cognitifs de haut niveau. Pour produire une explication satisfaisante, toutefois, nous devons postuler différentes espèces de concepts qui, au final, n'ont que très peu de propriétés communes et donc ne sont pas des espèces naturelles (*ibid*, p. 12). Il s'agit également d'un argument empirique, parce que les conclusions ne tiennent que si les entités théoriques postulées par les diverses théories des

¹⁸ Argument adapté de Machery 2005, p. 12

concepts sont *dans les fait* dissimilaires. Machery prend la prémisse 1 pour acquis et présente des arguments empiriques en support des prémisses 2 et 3.

La prémisse 2 est supportée par le fait que les prototypes, exemplaires et théories sont généralement compris comme ayant 1) des propriétés représentationnelles différentes et 2) des propriétés fonctionnelles différentes. Compte tenu des propriétés pertinentes à la caractérisation des représentations mentales (représentationnelles et fonctionnelles, selon Machery), les principales théories psychologiques des concepts postulent effectivement trois espèces différentes de concepts ayant peu de propriétés communes aux niveaux pertinents pour l'explication psychologique, en plus d'être employés dans des processus différents (*ibid*, pp. 16-20). Ceci supporte la prémisse 2, en plus de permettre de formuler l'énoncé conditionnel suivant : s'il est démontré que les trois types de structures informationnelles sont employés par défaut dans les processus cognitifs de haut niveau, alors « concept » ne réfère pas à une espèce naturelle.

Ceci nous amène à considérer les données empiriques mobilisées par Machery en support de la prémisse 3, soit les recherches sur la combinaison de concepts (*concept combination*). La combinaison de concept est la capacité de créer des structures informationnelles à propos d'une classe d'entités pour laquelle aucune structure informationnelle n'est emmagasinée dans la mémoire (*ibid*, p. 21). La représentation GRAND-MÈRE ESPIONNE est formée des représentations GRAND-MÈRE et ESPIONNE, et est en fait un concept complexe. Un concept est simple si et seulement si il s'agit d'une structure informationnelle permanente entreposée dans la mémoire à long terme. À l'opposé, un concept complexe n'est pas permanent, mais composé ponctuellement pour les besoins d'une tâche cognitive donnée (*ibidem*).

Machery soutient que les recherches de Osherson et Smith (1981) sur la combinaison de concepts démontrent que cette opération requiert trois types d'information différente : 1) des informations statistiques sur les propriétés typiques des différentes catégories, 2) des connaissances modales et causales concernant ces catégories et 3) des connaissances concernant certains membres de ces catégories (Machery 2005, p.28). Il mentionne ensuite

que les trois espèces fondamentales de concepts (prototypes, exemplaires et théories) sont, selon les théories qui les postulent, supposés encoder respectivement ces types d'information. (*ibidem*). De plus, d'autres processus cognitifs reposent sur ces trois sortes de concepts. La prémisse 3 de l'argument de nécessité explicative est validée. Expliquer la combinaison de concepts requiert de postuler l'existence de trois sortes de structures informationnelles. Ces trois sortes de concepts sont dissimilaires *et* figurent dans l'explication d'autres capacités cognitives. Finalement, en vertu de l'énoncé conditionnel découlant de la prémisse 2 (s'il est démontré que les trois types de structures informationnelles sont employés par défaut dans les processus cognitifs de haut niveau, alors « concept » ne réfère pas à une espèce naturelle), les concepts ne sont pas une espèce naturelle.

La position de Machery débouche sur un éliminativisme scientifique de la notion de « concept ». De la même manière que Griffiths prône l'élimination du terme vernaculaire « émotion » (est des termes d'émotion spécifiques) du vocabulaire scientifique visant à expliquer les phénomènes émotionnels, Machery soutient en s'appuyant sur la théorie de l'homéostasie causale des espèces naturelles, que le terme « concept » renvoie à des catégories hétérogènes et qu'il ne s'agit donc pas d'un terme d'espèce naturelle (Machery 2009, chapitre 8, section 2).

Nous disposons à présent des éléments nécessaires à l'analyse plus serrée des exemples d'hypothèses d'hétérogénéité présentés dans la présente section. Notre objectif, rappelons-le, est de déterminer les conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer justifiée et de présenter clairement les fondements des thèses défendues dans le mémoire. La prochaine section discute plus en détails des notions d'espèce naturelle, d'hypothèse d'espèce naturelle, et du rôle de ces catégories dans la recherche scientifique, qui sont les éléments nécessaires à la défense de la thèse présentée au chapitre II.

1.2. Hypothèses d'hétérogénéité

Dans la section précédente, nous avons présenté de manière schématique la théorie des espèces naturelles telle qu'employée par Griffiths et Machery. Il importe à présent d'en donner une explication plus complète, qui nous permette de comprendre clairement le rôle que jouent les espèces naturelles (ainsi que les hypothèses d'espèce naturelle) dans la pratique scientifique. Ceci nous permet également d'identifier les conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer justifiée. Nous commençons par présenter la théorie des espèces naturelles telle que développée par Boyd (1989, 1991). Ensuite, nous définissons la notion d'hypothèse d'espèce naturelle. Nous concluons la section (et le chapitre) par une discussion des conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer justifiée.

1.2.1. Espèces naturelles: une perspective sur la pratique scientifique

La théorie des espèces naturelles employée dans les argumentaires respectifs de Griffiths et Machery est celle développée par Boyd (1989, 1991) dans le cadre d'une réflexion sur le regain d'intérêt concernant les espèces naturelles et une défense de l'extension de la théorie des espèces naturelles aux espèces de grappes de propriétés (*property cluster kinds*) et aux espèces sociales (*social kinds*) contre les critiques de Hacking (1991). Selon Hacking, les différentes approches aux espèces naturelles ont en commun de supposer qu'une espèce naturelle est une espèce qui : 1) est définie par un ensemble de propriétés nécessaires et suffisantes telles que 2) la possession de ces propriétés est, dans les faits (et non simplement de manière logique) indicative d'un grand nombre d'autres propriétés méthodologiquement pertinentes (il y a une corrélation factuelle entre propriétés), ce qui implique que 3) ces propriétés définitives sont naturelles et non sociales (Boyd 1991, pp. 127-128). Lorsque les propriétés qui définissent une espèce rencontrent la seconde condition, alors il s'agit d'un fait à propos de la *nature* que l'espèce en question existe, de même qu'il s'agit d'un fait naturel que la référence à cette espèce permette la formulation de généralisations causales liées à la

corrélation factuelle de propriétés. Cette conception est très restreinte et implique qu'il n'existe que peu d'espèces naturelles. À l'opposé, ce que Boyd appelle « l'enthousiasme à l'endroit des espèces naturelles », surtout présent en philosophie des sciences, reflète une conception beaucoup plus large des espèces naturelles, incluant les espèces de grappes de propriétés (*property cluster kind*) et les espèces sociales (*ibidem*). Nous nous concentrons ici sur le premier type (*property cluster*).

Pour Hacking, ces types d'espèces ne doivent pas être assimilés à des espèces naturelles parce qu'ils sont, selon lui, construits sur la base d'une vague ressemblance assimilable à un « air de famille » (*family resemblance*) et que les regroupements de ce type (vagues, lâches, etc.) ne sont pas d'origine naturelle, mais sociale. En conséquence, les ressemblances sur la base desquelles sont rassemblés les membres de ce type d'espèce incluent non seulement des propriétés naturelles, mais également des propriétés sociales. Les espèces de grappes de propriétés ne répondent donc pas aux conditions 1 et 3. En ce qui concerne les espèces sociales Hacking rejette leur assimilation à des espèces naturelles parce que les espèces sociales dépendent de nos théories, et ne peuvent donc être naturelles. Ce dernier point peut sembler étrange, si l'on accepte la prémisse que toute observation scientifique est informée par une théorie d'arrière-plan (prémisse qui n'est pas particulièrement controversée). En effet, l'argumentaire de Boyd en faveur de l'idée selon laquelle il existe de bonnes raisons d'étendre la théorie des espèces naturelles aux espèces de grappes de propriétés et sociales repose en partie sur une observation de la pratique scientifique telle qu'elle se déroule dans les faits, de l'importance des théories d'arrière-plan pour l'observation scientifique, et du rôle des espèces naturelles en science. Cet argument devient important lorsqu'il est question du rôle des espèces naturelles dans la pratique scientifique. Pour le moment, il convient de présenter la théorie des espèces naturelles de Boyd.

L'inspiration de la théorie des espèces naturelles développée par Boyd provient de théories sémantiques en philosophie du langage ordinaire, qui défendent la possibilité de

définitions ne fournissant pas de conditions nécessaires et suffisantes à l'appartenance d'un objet à une catégorie donnée (Boyd 1989, p. 15). Ces *property cluster theories*¹⁹ soutiennent que la définition de certains termes est fournie par un ensemble cohésif (grappe) de propriétés telles que la possession d'un nombre adéquats de ces propriétés est suffisante pour qu'un objet tombe sous l'extension d'un terme donné (*ibidem*). Évidemment, dans le contexte de la philosophie du langage ordinaire, la question de savoir quelles propriétés appartiennent à l'ensemble (et quelle combinaison de propriétés est suffisante pour tomber sous l'extension du terme) est une question conceptuelle (*a priori*). Dans la mesure où l'on accorde un poids plus important à certaines propriétés dans le jugement d'appartenance à l'extension d'un terme, la détermination de l'importance d'une propriété donnée est également une question conceptuelle. L'idée de Boyd est, simplement, d'appliquer une version modifiée de cette théorie à la question de la définition naturelle des termes théoriques. Par « définition naturelle », nous entendons simplement une définition d'un terme référant à une entité naturelle (par opposition à une entité sociale ou purement conceptuelle, par exemple).

La théorie des grappes de propriétés fournit une manière intéressante d'aborder les définitions naturelles. Boyd soutient qu'il existe un grand nombre d'espèces scientifiquement importantes dont la définition naturelle est très semblable aux définitions de grappes de propriétés. Ces espèces sont les espèces causalement homéostatiques dont nous avons discuté précédemment. Des définitions naturelles de ce type impliquent 1) un ensemble de propriétés associé à 2) une indétermination extensionnelle. La forme de *property cluster* de ces définitions et l'indétermination extensionnelle qui y est associée sont « dictées » par l'emploi, en sciences, de catégories correspondant, tant au point de vue explicatif qu'inductif, à des structures causales pertinentes. L'indétermination extensionnelle, surtout, est inévitable. Pour l'éliminer, il faudrait rendre ces définitions non-naturelles, au sens où elles seraient basées sur une délimitation *a priori* de l'extension des termes naturels, les rendant ainsi scientifiquement inutiles, parce qu'elles ne permettraient pas de former des raisonnements

¹⁹ Ou *criteria attribute theories* (critères d'attributs).

inductifs (*ibid*, p. 16). Il convient à présent de donner une caractérisation claire de la théorie d'agglomérats de propriétés (*property cluster kind*)²⁰.

Suivant la démarche de Boyd (1989), nous présentons ici les 11 étapes de son exposé²¹.

- 1) Il existe une famille F de propriétés qui sont, de manière contingente, regroupées dans la nature au sens où elles sont cooccurentes dans un grand nombre de cas.
- 2) Leur cooccurrence est typiquement le résultat d'une forme d'homéostasie. Soit la présence de certaines des propriétés comprises dans F tend à favoriser la présence d'autres propriétés, soit il existe des mécanismes causaux qui tendent à maintenir la présence de certaines propriétés, ou les deux.
- 3) L'assemblage homéostatique de certaines propriétés dans F est *causalement pertinent* : il existe des effets importants causés par la cooccurrence de (plusieurs) propriétés dans F et par (ou en conjonction avec) des mécanismes causaux.
- 4) Il existe un terme d'espèce T qui réfère aux choses dans lesquelles l'assemblage homéostatique de plusieurs (ou de la plupart) des propriétés dans F se présente.
- 5) T n'a pas de définition analytique (*a priori* ou conceptuelle). L'ensemble homéostatique (ou la majeure partie de cet ensemble) en conjonction avec certains ou la totalité des mécanismes causaux sous-tendant l'ensemble homéostatique fournissent la *définition naturelle* de T. La question de savoir quelles propriétés et mécanismes sont pertinents pour les jugements d'appartenance à F est une question *empirique (a posteriori)*.

²⁰ Nous emploierons ci-après les expressions « grappes de propriétés » et « agglomérats de propriétés ».

²¹ Ce qui suit est adapté de Boyd 1989, pp. 16-18.

6) Une homéostasie imparfaite est concevable et potentiellement réelle. Certaines choses exhibent peut-être seulement une partie des propriétés dans F²².

7) Dans ce type de cas, la question de l'importance relative des différentes propriétés dans F et des différents mécanismes causaux pour déterminer si un objet donné tombe sous l'extension de T est une question *théorique* et non conceptuelle.

8) Qui plus est, les instances d'indétermination extensionnelle sont nombreuses.

9) L'importance causale de l'ensemble homéostatique de propriétés F en conjonction avec les mécanismes causal sous-jacents pertinents est telle que le type de propriété dénoté par T est une espèce naturelle. En d'autres termes, le fait que ces éléments jouent un rôle causal important est ce qui fait en sorte le terme qui les dénote est un terme qui dénote une espèce naturelle, soit une catégorie permettant la généralisation, l'induction et l'explication.

10) Il n'existe pas de raffinement de l'usage remplaçant T par un terme dont l'extension est moins vague qui en même temps préserve l'aspect naturel de l'espèce à laquelle T réfère.

11) L'ensemble homéostatique de propriétés qui fournit la définition de T n'est pas individué de manière extensionnelle, au sens où les propriétés qui définissent les conditions à remplir pour tomber sous T peuvent varier à travers le temps (suite à des observations empiriques) sans que T ne change de définition. Ceci fait en sorte que la théorie de Boyd accommode aisément la révision conceptuelle (voir section 1.1.1).

²² Ceci est cohérent avec le fait que, selon Machery, une espèce naturelle doit regrouper des objets ou entités partageant *un grand nombre* de propriétés scientifiquement pertinentes. Il n'est pas nécessaire que *toutes* les propriétés comprises dans F se retrouvent dans un objet donné pour que celui-ci fasse partie de l'extension du terme T.

Ces éléments nous permettent d'étoffer la définition des espèces naturelles que nous avons employée dans les sections précédentes. Une espèce naturelle est une famille, un ensemble, ou une « grappe » (*cluster*) de propriétés en homéostasie causale. Cette homéostasie causale découle de mécanismes causaux sous-jacents. Ainsi, les taxons biologiques, les éléments chimiques, les molécules, etc. sont des espèces naturelles. À l'opposé, des catégories qui sont, de manière évidente, arbitraires, comme « les objets plus gros que la tasse de Jim », « les fourchettes à dessert », « les objets supra lunaires » et une variété d'autres catégories construites sur la base de conventions sociales ou encore dont la définition dépend de l'importance relative de certaines propriétés ou mécanismes déterminés de manière *conceptuelle (a priori)* ne sont *pas* des espèces naturelles. Dans le cas de « les objets plus grands que la tasse de Jim », c'est parce qu'il ne s'agit que d'une propriété commune à un ensemble vraisemblablement hétérogène. De plus, comme nous l'avons mentionné plus haut, les définitions naturelles reposant sur des agglomérats de propriétés homéostatiques ne peuvent qu'être empiriquement déterminées (en conjonction, évidemment, avec une théorie d'arrière-plan), ce qui exclut les catégories dont le mécanisme homéostatique ne serait que conceptuel.

Il reste à présent à clarifier le rôle des espèces naturelles dans la pratique scientifique. Si nous en avons déjà une idée vague, les arguments de Boyd (1991) pour l'extension de la théorie des espèces naturelles aux espèces sociales et d'agglomérats de propriétés (*property cluster kinds*) permettent de mieux cerner ce rôle. Le raisonnement présenté ici se divise en deux étapes. La première consiste à expliciter le rôle des espèces naturelles dans la pratique scientifique. La seconde consiste à montrer que les espèces d'agglomérats de propriétés (*property cluster kinds*) sont, en vertu des considérations développées dans la première étape, assimilables à des espèces naturelles.

Pour Boyd (1991), comprendre la portée de la théorie des espèces naturelles requiert de comprendre, avant tout, la contingence radicale de l'épistémologie (*radical contingency*). Cette notion reflète un état de fait qui caractérise les raisonnements inductifs, à savoir que ceux-ci reposent toujours sur des jugements de potentiel de généralisation (*projectability*) a

posteriori (Boyd 1991, pp 136-139)²³. Afin de faire des généralisations inductives (dire par exemple que toutes les instances de cuivre sont des instances d'un métal qui fond à 1083° C), il est nécessaire d'employer des catégories définies *a posteriori*, de manière à ce qu'elles reflètent la structure causale de la nature. Ainsi, les espèces naturelles, qui sont précisément ce type de catégories, reflètent une stratégie qui consiste à faire déférence à la nature lorsque nous devons faire des jugements de potentiel de généralisations. En d'autres termes, on peut penser cette stratégie comme une heuristique : les espèces naturelles sont définies *a posteriori*, de manière à refléter la structure causale du monde, précisément parce que nous ne sommes pas en mesure d'identifier ou formuler des généralisations appropriées sans avoir recours à cette stratégie (Boyd 1991).

Les espèces naturelles sont donc des catégories dont la définition se fait suite à des observations empiriques informées par une théorie d'arrière-plan donnée. L'induction et l'explication requièrent d'employer des catégories qui, le plus possible, reflètent la structure causale de la nature. Le rôle des espèces naturelles dans la pratique scientifique est donc le suivant : permettre d'élaborer des théories sur la base de catégories dont la définition (et la structure sémantique) est telle qu'elle permet de refléter le plus possible la structure causale de la nature. Pour Boyd, ce rôle est adéquatement rempli par, notamment, les espèces d'agglomérats de propriétés (*property cluster kinds*).

Partons de l'idée de Hacking que les espèces naturelles sont définies par des conditions nécessaires et suffisantes. Toutes les espèces naturelles sont *causalement* associées à un

²³ Cette idée est présentée par Boyd comme découlant de l'échec du fondationnalisme. D'une part, le fondationnalisme des prémisses, selon lequel toute connaissance est justifiable sur la base d'un noyau de croyances fondamentales jouissant d'une sorte de privilège épistémique est écarté par (entre autres) les travaux de Quine (1969). D'autre part, le fondationnalisme des inférences (*inference foundationalism*) selon lequel il existe des principes justifiables d'inférence inductive réductibles à des principes inférentiels justifiables *a priori* est clairement faux si l'on accepte (comme nous le faisons ici) que le raisonnement inductif repose sur des jugements de potentiel de généralisation *a posteriori*. Nous revenons sur ce point au prochain chapitre. Pour un exposé de l'argument réaliste en faveur de cette idée, voir Boyd 1991, pp.136-138.

certain nombre de propriétés méthodologiquement pertinentes²⁴. D'abord, le caractère *naturel* des espèces définies par des conditions nécessaires et suffisantes reflète le fait qu'elles sont associées à un ensemble plus large de propriétés que les seules propriétés « privilégiées » incluses dans les conditions nécessaires et suffisantes (toutes les instances de H₂O ont certaines propriétés, mais seul un sous-ensemble de celles-ci sont comprises dans les conditions nécessaires et suffisantes). Boyd argumente qu'il n'est pas inconcevable que dans certains cas (qui en fait constituent une large part de l'activité scientifique), nous employons des catégories, ou des espèces, définies par cette large famille de propriétés plutôt que par le sous-ensemble privilégié fournissant des conditions nécessaires et suffisantes (Boyd 1991). En vertu du fait que les espèces naturelles nous fournissent une stratégie nous permettant de formuler des explications et des généralisations inductives dont l'utilité et la valeur sont jugées *a posteriori* (catégories qui, au demeurant, sont aisément révisables suite à de nouvelles informations), il ne semble pas y avoir de bonne raison de rejeter les espèces d'agglomérats de propriétés de la théorie des espèces naturelles. Cette interprétation est également cohérente (dans ses grandes lignes) avec l'approche mécaniste en philosophie des sciences, surtout dans la version développée par Bechtel (2008, voir également Bechtel et Richardson 2010). Bechtel met l'accent sur l'importance de développer une philosophie des sciences informée par l'observation de la pratique scientifique telle qu'elle se déroule dans les faits.

Les espèces naturelles sont donc des catégories définies par un ensemble de propriétés partagées (parfois de manière partielle ou imparfaite) par ses membres. Ces catégories jouent un rôle important en science, soit celui de permettre de formuler des explications et des raisonnements inductifs généralisables. Il reste à voir de quelle manière ce rôle est implémenté.

²⁴ Cet aspect est central à l'argument de Machery (2005). Voir aussi section 1.1.3 ci-haut.

1.2.2. Hypothèses d'espèce naturelle

L'on peut concevoir le rôle des espèces naturelles dans la pratique scientifique comme implémenté par les hypothèses d'espèce naturelle. Faire une telle hypothèse revient à affirmer qu'un terme T réfère à une espèce naturelle. Face à ces hypothèses, deux choses peuvent arriver. Dans le premier cas, l'hypothèse fait consensus et il est généralement accepté que le terme T réfère à une classe C d'entités, classe (ou catégorie) à propos de laquelle nous pouvons formuler des généralisations sur la base desquelles développer des explications et des théories. Dans le second cas, l'hypothèse ne fait pas consensus et il s'agit dès lors de réfuter ou de corroborer l'hypothèse. Dans les deux cas, les débats au sein de la discipline sont influencés par la réception de l'hypothèse d'espèce naturelle : soit les questions « importantes » de la discipline concernent la valeur des explications et théories proposées (hypothèse acceptée), soit les questions se situent au niveau des fondements de la discipline (hypothèse rejetée)²⁵. Une fois qu'il est accepté que, par exemple, les termes du vocabulaire de la chimie renvoient à des espèces naturelles, la recherche dans ce domaine concerne ces éléments et composés chimiques. Pour prendre un exemple familier, on peut dire que l'adoption d'une hypothèse d'espèce naturelle à propos du terme « concept » en psychologie fait en sorte que les chercheurs supposent que les concepts forment une classe d'entités qui partagent suffisamment de propriétés scientifiquement pertinentes pour permettre des généralisations à leur endroit.

Mais une hypothèse d'espèce naturelle, même largement acceptée par la communauté, peut être erronée. Les découvertes sur la mémoire sont un excellent exemple. Alors que la mémoire à long terme a longtemps été considérée comme une seule chose (Bechtel 2008, p.54), les travaux de Scoville et Milner (1957) sur le patient H.M. ont révélé qu'il en était

²⁵ Il existe aussi le cas où les philosophes des sciences remettent en question des hypothèses d'espèce naturelle généralement acceptées par la communauté scientifique (Machery 2005 et Griffiths 1997, par exemple), mais il est ici question du rôle de ces hypothèses au sein d'un champ scientifique donné.

autrement. Suite à l'ablation chirurgicale d'une large part du gyrus temporal moyen de H.M. (dans le cadre d'un traitement pour l'épilepsie), le patient était incapable d'assimiler de nouvelles informations (amnésie antérograde). Toutefois, il était en mesure d'apprendre de nouvelles habiletés sans pour autant se souvenir de les avoir apprises (Bechtel 2008, p. 54). Ces découvertes ont motivé les chercheurs à remettre en question l'idée que la mémoire à long terme est « unique », et les a amené à la considérer comme (au moins) deux systèmes de traitement d'information partiellement superposés. Les modèles de la mémoire sont aujourd'hui plus complexes, mais l'important est de comprendre qu'une hypothèse d'espèce naturelle peut être erronée, comme l'était celle à propos de la mémoire à long terme. Cet exemple illustre également le fait que si l'erreur n'est pas relevée, on peut voir émerger des débats contre-productifs sur la nature de la mémoire, alors qu'elle n'est pas une espèce naturelle.

Donc, l'adoption d'une hypothèse d'espèce naturelle a un effet semblable à l'adoption d'un paradigme kuhnien: elle informe et dirige les débats et recherches dans la discipline. Toutefois, lorsque le travail fondé sur une hypothèse d'espèce naturelle erronée commence à se heurter à des insuffisances explicatives, la recherche risque de retourner à un niveau fondamental, soit celui de la définition de la nature de l'objet d'étude. C'est précisément la situation à laquelle Machery (2005) et Griffiths (1997) tentent de répondre. Dans le premier cas, certains psychologues des concepts supposent qu'il n'existe qu'un seul type de concept et développent des théories sur la base de cette idée. Dans le second, l'idée que les catégories scientifiques dénotant les phénomènes émotionnels ne correspondent pas aux catégories vernaculaires associées aux émotions est un fait qu'il faut reconnaître pour être en mesure de comprendre ces phénomènes. Finalement, le constat qu'une hypothèse d'espèce naturelle est erronée appelle à corriger la situation, et l'une des manières de procéder est d'énoncer une hypothèse d'hétérogénéité.

1.2.3. Hétérogénéité : conditions

Une hypothèse d'hétérogénéité est, simplement, la négation d'une hypothèse d'espèce naturelle. Nous allons à présent revenir sur les exemples discutés (émotions et concepts) pour en tirer les conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer justifiée.

La première condition est la suivante : le terme, le concept ou la catégorie qui fait l'objet de l'hypothèse d'hétérogénéité doit également faire l'objet d'une hypothèse d'espèce naturelle *ou* d'un procédé analogue, dont les modalités et conséquences sont proches de celles d'une hypothèse d'espèce naturelle (nous revenons à cette condition plus loin, parce qu'elle est importante pour l'argumentaire général développé dans le chapitre II). Machery (2005), par exemple, part de l'idée que les psychologues des concepts prennent pour acquis que ces derniers forment une espèce naturelle. En l'absence d'une telle supposition de la part des chercheurs, il n'y aurait aucune raison d'avancer l'idée que les concepts sont hétérogènes. Dans le cas de Griffiths (1997), il est plutôt question d'une présupposition de correspondance entre termes vernaculaires et catégories scientifiques. Son argumentaire repose sur la démonstration du fait que le terme vernaculaires « émotion » (ainsi que les termes d'émotions spécifiques) sont des cas de référence partielle, référant d'une part à des programmes d'affect et d'autre part à des stratégies sociales (Griffiths et Scarantino 2005). En d'autres termes, une hypothèse d'hétérogénéité s'applique lorsqu'il s'agit de remettre en question l'homogénéité d'une catégorie.

Il ne s'agit pas d'une condition à proprement parler, mais dans certains cas, l'hypothèse d'hétérogénéité peut sembler particulièrement controversée si elle nie une hypothèse d'espèce naturelle faisant consensus au sein d'une discipline donnée. Ceci pourrait laisser penser que la valeur d'une hypothèse d'hétérogénéité dépende de telles considérations, mais il n'en est rien. Nier ou défendre une hypothèse d'espèce naturelle marginale peut être fort utile en philosophie des sciences. Nous n'avons pas à traiter plus en détails de cette complication, puisque dans les exemples dont nous avons discutés, les hypothèses d'espèce naturelles sont relativement bien établies dans les disciplines concernées. La position de Machery est une réponse directe à un présupposé théorique qui caractérise une large part de la recherche en

psychologie des concepts. Celle de Griffiths est dirigée contre l'idée que les catégories vernaculaires sont appropriées à l'explication scientifique des phénomènes émotionnels *et* contre l'idée que ces phénomènes ne correspondent qu'à des programmes d'affect. En effet, si la théorie des programmes d'affect peut expliquer une partie des phénomènes émotionnels, il faut une autre théorie (Griffiths et Scarantino proposent la perspective située sur les émotions) pour expliquer ce qui est hors de la portée de la théorie des programmes d'affect. En ce sens, cette seconde condition n'est pas nécessaire, mais permet de mesurer l'impact d'une hypothèse d'hétérogénéité.

La troisième condition stipule que nous devons disposer de raisons empiriques *et* théoriques pour remettre en question l'hypothèse d'espèce naturelle (ou la présupposition d'homogénéité). Les raisons théoriques sont explicites dans le travail de Griffiths (1997) : son point de départ est l'idée que pour construire une psychologie scientifique des émotions, il faut opérer une révision conceptuelle des concepts d'émotion et d'émotions particulières de manière à ce que ces concepts réfèrent à des catégories causalement homéostatiques, de manière à permettre l'élaboration d'explications et de raisonnement inductifs généralisables. L'hypothèse d'hétérogénéité de Griffiths est donc motivée par des considérations théoriques issues de la théorie des espèces naturelles et de la philosophie des sciences. Sa thèse est que les catégories vernaculaires d'émotion et d'émotions particulières sont des exemples de référence partielle, qui sont par conséquent inutiles à la construction d'une psychologie scientifique des émotions. Il s'agit d'une thèse empirique : Griffiths doit montrer que les termes vernaculaires réfèrent à des choses qui sont, *dans les faits* différentes les unes des autres. Il existe (au moins) deux types d'entités très différentes auxquelles font référence les termes vernaculaires d'émotion : des programmes d'affect *et* des stratégies d'interaction sociale (voir section 1.1.2). De la même manière, Machery développe l'argument de la nécessité explicative sur la base de considérations théoriques, ramenant la pertinence de son argument à une question empirique. En effet, les conclusions ne tiennent que si les entités théoriques postulées par les diverses théories des concepts *sont dans les fait* dissimilaires, ce qui demande d'apporter un support empirique aux prémisses 2 et 3 de son argument (voir section 1.1.3). La forme générale des hypothèses d'hétérogénéité est donc la suivante : des

considérations théoriques sont mobilisées afin de formuler un argument empirique. Si celui-ci s'avère juste (en regard de diverses considérations empiriques et d'inférences tirées de l'analyse de résultats empiriques), alors l'hypothèse d'hétérogénéité peut être considérée.

De ces trois conditions, deux (1 et 2) sont contextuelles. La troisième condition est méthodologique. Les conditions 1 et 2 sont suffisantes pour *motiver* la formulation d'une hypothèse d'hétérogénéité. La troisième condition concerne la forme que doit prendre une telle hypothèse : il faut 1) des raisons théoriques de remettre en question une hypothèse d'espèce naturelle (ou une supposition d'homogénéité) et 2) une composante empirique qui fournit un certain support à l'hypothèse d'hétérogénéité. Cette composante empirique est également rendue nécessaire, du moins en regard des exemples présentés ici, par la structure des arguments déployés par Griffiths et Machery : il s'agit d'arguments reposant sur des questions empiriques (par exemple : « les concepts sont-ils vraiment dissimilaires? »). Il faut également noter que ces conditions ne sont pas définitives, mais simplement le résultat de l'analyse des exemples présentés plus haut. Pour anticiper sur le prochain chapitre, nous pouvons relever qu'elles laissent une certaine marge de manœuvre quant au contexte dans lequel il peut s'avérer pertinent de formuler une hypothèse d'hétérogénéité.

1.3. Conclusion

L'objectif du présent chapitre est, rappelons-le, de clarifier l'inspiration principale de notre projet et de placer les éléments nécessaires à la défense des thèses mises de l'avant dans les chapitres suivants. Nous avons donc présenté deux exemples d'approches hétérogènes tirées de la philosophie de la psychologie. Ces exemples sont toutefois plus qu'une simple source d'inspiration pour le développement d'une thèse d'hétérogénéité de la connaissance. En effet, suite à une discussion plus détaillée des espèces naturelles et de leur rôle en sciences, nous avons identifié la notion d'hypothèse d'espèce naturelle. Il s'agit d'un procédé par lequel on fait l'hypothèse qu'un terme T du vocabulaire scientifique réfère à une espèce naturelle. Ce type de procédé est central à la recherche scientifique, et lorsqu'une hypothèse

d'espèce naturelle potentiellement erronée est formulée, cela a des effets importants et, surtout, relativement faciles à observer. Une solution possible, lorsqu'aux prises avec les problèmes qui découlent d'une hypothèse d'espèce naturelle erronée, est de formuler une hypothèse d'hétérogénéité à l'endroit du terme faisant l'objet de l'hypothèse d'espèce naturelle. C'est ce que font Machery et Griffiths. En analysant leurs démarches respectives, nous avons pu formuler trois conditions qui doivent encadrer le développement d'une espèce naturelle. Ainsi, les exemples de Machery et Griffiths fournissent une sorte de marche à suivre pour l'élaboration de notre propre thèse d'hétérogénéité. Il s'agit, en fait, de l'apport principal de ce chapitre.

Dans le prochain chapitre, nous allons, premièrement, construire une perspective métathéorique du débat sur le niveau descriptif approprié à l'étude de la connaissance. Ceci nous permet d'avoir une vue d'ensemble du débat afin de confirmer que celui-ci présente effectivement des similarités avec, par exemple, la situation dans la recherche sur les concepts. Ensuite, nous développons un argument de nécessité explicative, dont la composante empirique est traitée en détail dans le chapitre III. En d'autres termes, nous tâchons de montrer que le débat sur la nature de la connaissance remplit les conditions contextuelles, pour ensuite construire un argument remplissant la condition méthodologique.

CHAPITRE II

ÉNONCÉ DU PROBLÈME DE RECHERCHE ET ARGUMENT DE NÉCESSITÉ EXPLICATIVE

Nous avons, dans le chapitre précédent, mis en place les éléments préalables nécessaires à la défense des thèses du mémoire. Ces éléments sont les conditions dans lesquelles une hypothèse d'hétérogénéité peut s'avérer recevable et la forme générale que doit suivre le développement de ce type d'hypothèse.

Nous abordons dans le présent chapitre l'étape cruciale de notre discussion. Nous commençons par analyser le débat concernant le niveau de description approprié à l'étude de la connaissance²⁶ (sections 2.1 et 2.2) de manière à rendre explicite le problème que nous cherchons à soulever et, à terme, résoudre (section 2.3). Par la suite, nous élaborons un argument de nécessité explicative inspiré de celui de Machery (2005, 2009) débouchant sur la conclusion que la forme traditionnelle de la solution au problème de la nature de la connaissance est inappropriée (section 2.4). Cet argument repose sur le fait de démontrer qu'il existe des structures informationnelles qui 1) peuvent être considérées comme des connaissances et 2) sont irréductibles aux formes de connaissances postulées par l'épistémologie traditionnelle. La démonstration est menée dans le chapitre III. Les objectifs du présent chapitre sont donc les suivants : 1) montrer que le débat sur la nature de la connaissance remplit les conditions contextuelles identifiées dans le chapitre I et 2) développer un argument de nécessité explicative, en accord avec les paramètres identifiés dans le chapitre I.

²⁶ L'expression « débat sur la nature de la connaissance » sera employée à certains endroits pour alléger le texte.

2.1. Histoire et caractérisation du paradigme phrastique

En accord avec la méthodologie établie au début du mémoire, nous commençons par développer une perspective métathéorique du débat sur la nature de la connaissance. La première étape est de décrire ce que nous appelons le « paradigme phrastique » : l'idée que le niveau phrastique est le niveau descriptif adéquat pour l'étude de la connaissance. En d'autres termes, la connaissance est véhiculée par des phrases. Celles-ci ont des conditions de vérité (et de justification ou de fiabilité, dépendamment de la perspective) et c'est en vertu de ces conditions qu'elles acquièrent leur pertinence épistémique. Le paradigme phrastique émerge au début du 20^e siècle et joue un rôle important dans l'élaboration de l'épistémologie « traditionnelle ». Il est également central à une large part du naturalisme épistémologique et c'est à ce niveau qu'il est le plus intéressant pour nos fins.

2.1.1. Émergence des conceptions phrastiques de la connaissance

Quine (2004) trace l'apparition de la conception phrastique de la connaissance à deux changements de perspective importants dans l'empirisme britannique²⁷. Le premier est le passage de l'intérêt pour les idées à l'intérêt pour les expressions linguistiques. L'empirisme est caractérisé par l'idée que seules les sensations sont significatives. C'est-à-dire qu'une idée n'est recevable, ou pertinente, que si elle est basée sur des impressions sensorielles. Quine considère que John Horne Tooke, en étendant cette idée aux expressions linguistiques (les mots n'ont de sens que s'ils peuvent être définis en termes sensoriels), est responsable du fait que de s'intéresser aux idées, l'empirisme s'est tourné vers l'étude des expressions linguistiques (Quine 2004, p. 301). Cette idée se heurte toutefois au problème des particules

²⁷ Ces deux tournants font partie de cinq changements qui ont mené l'empirisme vers le naturalisme épistémologique. Nous discutons plus en détail du naturalisme plus loin, et les deux changements que nous décrivons sont les plus intéressants pour nos fins.

grammaticales, comme les prépositions ou les articles. Celles-ci n'ont pas de référent sensoriel, et ne peuvent être définies en termes sensoriels.

La solution à ce problème, pour Quine, vient de Jeremy Bentham. Bentham suggère que les particules grammaticales acquièrent un sens *en contexte*, et non en isolation. Bentham étend la définition contextuelle à certains termes qui, malgré leur utilité peuvent parfois s'avérer ontologiquement difficiles. Ce type de terme, comme les particules grammaticales, n'a de signification que comme partie d'un tout. Le véhicule primaire de la signification est la *phrase*, et non le *terme*. Il s'agit du second tournant, soit un changement de perspective concernant les véhicules de la signification.

La conséquence de ce changement de perspective qui est la plus intéressante pour nous est la suivante : l'épistémologie du 20^e siècle s'est articulée autour de l'idée que les phrases sont le principal vecteur de la signification. Ceci explique la place prépondérante des croyances (puisque celles-ci sont ce que véhiculent les phrases) en épistémologie « traditionnelle », où la connaissance est (du moins avant l'arrivée du problème de Gettier) définie comme une « croyance vraie justifiée » (Steup 2008).

Gettier (1963) relève un problème profond de cette définition de la connaissance. Le critère de justification dans « croyance vraie justifiée » remplit la fonction suivante : il fait en sorte qu'il ne peut exister de connaissances « accidentelles ». En effet, la simple possession d'une croyance vraie peut être le résultat de la « chance épistémique » (*epistemic luck*). Gettier met en évidence le fait que cette condition est insuffisante.

Imaginons le cas suivant²⁸ : Jim a de bonnes raisons de croire la proposition (fausse) suivante :

1. John possède une Ford

²⁸ Exemple adapté de Steup (2008) et de Gettier (1963).

Supposons, de plus, que Jim infère de (1):

2. John possède une Ford ou est à Petawawa.

Puisque de (1) découle (2), Jim est justifié de croire (2). Supposons que John, par pure coïncidence, se trouve effectivement à Petawawa. Compte tenu de ce que nous avons dit, lorsque nous disons de Jim qu'il croit (2), nous pouvons dire qu'il a une croyance vraie justifiée : Jim croit (1), qui implique (2) et (2) est vraie parce que John est effectivement à Petawawa. Jim ne dispose d'aucune information (*evidence*) concernant la localisation de John. Il ne croit une proposition vraie que par pure chance. Donc Jim ne *sait* pas que John est à Petawawa. La condition de justification est insuffisante. Il faut donc soit la renforcer ou trouver une ou plusieurs autres conditions.

Une large part du travail épistémologique s'est orienté vers le projet de développer une analyse de la connaissance qui soit immunisée contre les cas Gettier. Les propositions bien connues sont de renforcer le critère de justification (Chisholm 1977) ou trouver une condition supplémentaire. Armstrong (1973) propose par exemple d'ajouter la condition suivante : « la croyance de S que P ne peut être inférée d'une proposition fausse », ce qui ne règle pas le problème, puisque les cas Gettier ne requièrent pas qu'il y ait inférence de la part du sujet. Une alternative est le fiabilisme, développé entre autre par Goldman (1976, 1986), selon laquelle certaines « représentations » cognitives comptent comme des instances de connaissance à cause de la fiabilité des processus et mécanismes cognitifs qui ont dans les faits produit cette représentation (cette approche est discutée plus en détail au chapitre III).

L'épistémologie « traditionnelle » a pris plusieurs formes, l'une des plus importantes étant celle développée par le Cercle de Vienne. Dans *Epistemology Naturalized* (1969), Quine formule une critique de l'épistémologie du Cercle de Vienne qui débouche sur le rejet du fondationnalisme des prémisses, selon lequel toute connaissance est justifiable sur la base d'un noyau de croyances fondamentales jouissant d'une sorte de privilège épistémique. L'empirisme logique a tenté de trouver cette fondation dans la logique, la théorie des ensembles, ou les *énoncés* empiriques avec des résultats mitigés, si l'on en croit Quine. Ce

dernier rejette l'idée d'une fondation privilégiée de la connaissance parce que tout ce dont nous disposons pour aborder le monde est nos impressions sensorielles²⁹. Il faut donc, pour expliquer et comprendre la connaissance, l'aborder par le biais de la psychologie, qui seule peut nous fournir une explication causale de comment, à partir d'impressions sensorielles, nous arrivons à la connaissance (et, surtout, à la connaissance scientifique). Il s'agit donc d'un changement de projet et de méthode.

Il est important de relever deux choses. Premièrement, Quine ne rejette pas le paradigme phrastique. Son objectif est de se débarrasser du projet fondationnaliste. L'ensemble des préceptes de l'épistémologie « traditionnelle » est préservé chez Quine. Pour lui, il y a des phrases observationnelles (en langage naturel, ex. : la phrase : « l'accéléromètre est passé de .8 m/s² à .9 m/s² »), des phrases théoriques (ex. : « la force gravitationnelle diminue au carré de la distance ») et des comportements (béhaviorisme) d'assentiment à ces phrases. L'épistémologie doit expliquer ce qui régit ces comportements. Deuxièmement, la position de Quine sur l'explication de la connaissance peut être reformulée pour mettre en évidence une préoccupation qui est centrale à l'ensemble des approches naturalistes, incluant les approches éliminativistes telle que celle de Churchland, que nous présentons dans la section 2.2. Cette préoccupation correspond à l'idée qu'une explication de la connaissance requiert d'étudier les processus causaux par lesquels celle-ci émerge (pour Quine, il s'agit d'expliquer quels stimuli contrôlent, dans le sens du conditionnement opérant, nos comportements d'assentiments à des phrases).

2.1.2. Caractérisation du paradigme phrastique

Nous avons donc une certaine base pour l'exposition de notre conception métathéorique du paradigme phrastique. Il nous faut maintenant en clarifier les caractéristiques. La première

²⁹ Il s'agit certes d'une « fondation » mais pas au même sens que les énoncés empiriques des empiristes logiques, qui avaient pris le tournant linguistique. Ils cherchaient donc à fonder la connaissance sur les énoncés empiriques plutôt que sur les sensations.

chose à noter est que notre notion de paradigme phrastique inclut l'ensemble des approches épistémologiques « traditionnelles » *ainsi que plusieurs approches naturalistes*. En fait, les approches « traditionnelles », ou « intellectualistes » ne nous intéressent pas principalement, et nous traiterons surtout des approches naturalistes phrastiques. Il importe de mentionner que la conception que nous proposons à la fin du chapitre III est fortement naturaliste. En ce sens, nous sommes opposés aux approches « traditionnelles » comme étant les *seules* approches pertinentes pour étudier la connaissance (elles demeurent utiles, comme nous le verrons plus loin). L'important est de comprendre que le débat auquel nous nous attaquons est celui entre les conceptions phrastiques et non phrastiques de la connaissance (plus précisément les approches prônant l'élimination de la terminologie phrastique). La raison de ce choix est que dans un cadre naturaliste, l'accent est (souvent) mis sur l'identification d'une espèce naturelle « connaissance », ce qui nous rapproche des conditions contextuelles « idéales » pour la proposition d'une hypothèse d'hétérogénéité de la connaissance³⁰.

Le paradigme phrastique se définit ainsi : perspective selon laquelle les états épistémiques et les signaux d'entrée sont véhiculés par des phrases liées entre elles par des règles logiques (ou causales, chez Quine) et exprimant des propositions. Ces règles logiques régissent également les processus de raisonnement. Il faut décompresser cette définition.

« États épistémiques » réfère aux états d'un organisme qui possède et emploie des informations concernant son environnement, les conséquences possibles de ces comportements, etc. Ces états sont représentés ou réalisés par des véhicules informationnels. En d'autres termes, un état épistémique consiste à connaître, à posséder de l'information. Les

³⁰ Nous pensons toutefois que l'argument que nous formulons à la fin du présent chapitre s'applique, *mutatis mutandis*, si nous incluons les approches non naturalistes. Nous présentons quelques raisons en support de cette idée, mais il faudrait probablement la défendre de manière systématique. Nous n'avons pas la place de le faire ici.

organismes qui ont des états épistémiques sont des agents épistémiques³¹. Dans le cadre du paradigme phrastique, les véhicules informationnels pertinents sont des phrases (comme le suppose Quine, par exemple).

Pour qu'une phrase ait une valeur épistémique, celle-ci doit être *vraie*. La vérité est une notion vague dans le cadre d'une épistémologie empiriquement fondée (comme celle que Quine). La vérité d'une phrase doit dépendre, au moins en partie, de si elle correspond ou non aux faits. Ce type de critère de vérité est commun aux approches « traditionnelles » et aux approches « intellectualistes » contemporaines. Williamson (2000) soutient par exemple que la connaissance est un état factif : Jim sait que A si et seulement si il est le cas que A (cet état factif est distinct d'une croyance mais est tout de même véhiculé de manière phrastique). Par contre, il existe des cas où la vérité d'un énoncé dépend plutôt de sa relation logique avec d'autres énoncés, et c'est ici qu'entrent en jeu les règles logiques : c'est par la cohérence avec ces règles que les énoncés qui ne sont pas directement observationnels (énoncés théoriques) obtiennent leur valeur épistémique. Une telle phrase est vraie si elle est logiquement liée à d'autres phrases dont la vérité est empirique. Comme le relève Livingston (1996), ces règles s'appliquent également au raisonnement : l'être humain doit raisonner en accord avec les principes de la logique, qui garantissent la rationalité du raisonnement, en plus de certains facteurs plus spécifiques (comme le fait, par exemple, de considérer chaque possibilité avant de prendre une décision). L'importance du critère de vérité est également fonction du fait que si la connaissance est exprimée par des phrases, alors le contenu des phrases exprimant des connaissances est propositionnel. Une proposition est soit vraie ou fausse. Seules les propositions vraies, et donc les phrases vraies, sont des connaissances³².

³¹ Nous empruntons cette terminologie à Churchland (1979).

³² Les considérations concernant la notion de vérité développées dans ce paragraphe sont très générales et mériteraient d'être approfondies. Toutefois, nous pensons que pour les fins de notre discussion, ce dont nous disposons ici est suffisant.

Nous avons mentionné à quelques reprises que ce paradigme est prépondérant à l'intérieur du naturalisme épistémologique. Il nous faut justifier cette affirmation. Le naturalisme épistémologique se divise, grossièrement, en trois types (Feldman 2001). Le premier type est le naturalisme de remplacement : l'épistémologie doit être remplacée par une discipline scientifique. Quine (1969) fait partie de cette catégorie : il vise le remplacement de l'épistémologie telle qu'elle était pratiquée par la psychologie. L'approche éliminativiste que nous discutons dans la prochaine section tombe sous cette catégorie.

Le second type est ce que Feldman (2001) appelle le naturalisme coopératif. Celui-ci soutient que l'épistémologie devrait tenir compte des avancées scientifiques en psychologie et sciences connexes pour régler des problèmes philosophiques liés à la connaissance. Kornblith (2002) propose une conception naturaliste de la connaissance qui repose sur l'identification d'une espèce naturelle « connaissance ». Comme nous l'avons vu au chapitre I, les espèces naturelles jouent un rôle crucial en science en rendant possible la formulation de généralisations inductives. Il est utile ici de refaire rapidement l'argument de Kornblith. Il commence par justifier l'attribution d'états intentionnels (incluant des représentations internes des besoins de l'animal) aux animaux. Cette pratique est justifiée en vertu du travail explicatif qu'elle accomplit : une explication physicaliste nous empêcherait de voir plusieurs comportements comme des instances d'un même type de comportement (*ibid*, p.34). Ceci est également vrai des états informationnels internes : une fourmi employant la navigation à l'estime pour retourner à la colonie après avoir trouvé de la nourriture doit posséder une forme ou une autre de représentation de son environnement, de la distance parcourue et ainsi de suite (*ibid*, p.37). Donc, les animaux ont des représentations internes de leurs désirs, et des états informationnels internes sur leur environnement (ou du moins, nous sommes justifiés de leur attribuer ces éléments). Ainsi, le comportement animal résulte de l'interaction entre les besoins (désirs et intentions) des animaux et l'information dont ils disposent (croyances).

Il faut ensuite montrer comment ces croyances acquièrent le statut de connaissances. Au niveau individuel, les croyances sont suffisantes pour expliquer le comportement. Par contre, dans un cadre évolutionniste tel que celui qu'adopte Kornblith, cela n'est pas suffisant : il est

peu probable qu'une espèce ait pu évoluer parce que, par chance, les croyances de la plupart de ses membres se sont avérées vraies. Ces croyances doivent être produites de manière fiable (*ibid*, p.53). L'argumentaire de Kornblith mène donc à définir l'espèce naturelle « connaissance » comme « une croyance vraie produite de manière fiable » (Kornblith 2002, p. 58). Ici, la fiabilité renvoie au fait que la simple possession de croyances vraies n'est pas suffisante pour expliquer le fait qu'une espèce entière a pu évoluer : la production de ces croyances doit être fiable chez les membres de l'espèce. Kornblith, en un sens, naturalise l'épistémologie de manière « *top-down* », en argumentant en faveur de l'attribution de connaissances à des animaux non-humains tout en conservant la terminologie phrastique.

Le troisième type de naturalisme épistémologique discuté par Feldman (2001) est le naturalisme substantif. Celui-ci concerne la nature de « faits épistémologiques ». Par exemple, Goldman (1979) propose une liste de termes propres aux énoncés épistémologiques paradigmatiques (« a de bonnes raisons », « justifié », etc.). Ces termes sont intrinsèquement évaluatifs, et supposent l'existence de faits épistémologiques objectifs. Le naturalisme substantif propose de caractériser la nature de ces faits de manière scientifiquement cohérente, mais ne remet pas nécessairement leur utilité en question. La différence entre ce type de naturalisme et le naturalisme coopératif est, simplement, une question d'objet d'étude. Le naturalisme substantif s'interroge sur le contenu des termes épistémologiques. Le naturalisme coopératif est une position plus générale.

Ainsi, le paradigme phrastique est prépondérant au sein du naturalisme épistémologique. Dans certains cas, cette prépondérance est due au fait que l'adoption du paradigme phrastique est associée à un ensemble d'engagements et de positions théoriques qui sont utiles aux épistémologues adoptant un cadre naturaliste. Kornblith est un bon exemple. Il faut noter que la présentation des théories que nous venons de faire est très schématique et nous reconnaissons qu'il y a beaucoup plus à dire sur la place du paradigme phrastique en épistémologie. Toutefois, l'idée est simplement de mettre en évidence le fait que l'approche phrastique n'est pas simplement l'apanage de l'épistémologie « traditionnelle » et qu'elle a une place importante au sein du naturalisme.

Dans ce qui suit, nous présentons la position éliminativiste de Churchland (1979, 1989, 2001). Par la suite, nous serons en mesure d'évaluer celle-ci et d'aborder la formulation de notre argument de nécessité explicative.

2.2. Présentation et évaluation de l'éliminativisme

L'éliminativisme matérialiste (auquel nous référons par « éliminativisme ») est une position radicale affirmant que nos concepts et notions folkloriques ou de sens commun concernant l'esprit humain doivent être remplacés par notions scientifiquement cohérentes, celles-ci étant souvent situées au niveau neurologique ou, plus généralement, au niveau du substrat physique. La position dont nous traitons ici est celle de Paul Churchland (1979, 1988, 1989, 2007). Les travaux de Churchland dépassent les frontières de l'épistémologie, mais nous nous limiterons à discuter sa critique de l'approche phrastique, et la position éliminativiste qui en découle.

La présente section aborde quatre problèmes distincts. Ces problèmes reposent sur un engagement théorique à expliquer causalement l'émergence de la connaissance (qui est l'engagement de Quine). Il convient de présenter rapidement les grandes lignes de la position de Churchland, surtout en ce qui concerne l'engagement théorique susmentionné. Il faut premièrement comprendre que Churchland s'attaque à deux choses : (1) la conception phrastique de la connaissance et (2) la conception propositionnelle de la connaissance.

Ou :

(1) La connaissance est véhiculée de manière phrastique

(2) La connaissance est propositionnelle (constituée de propositions ayant des conditions de vérité)

Ces deux éléments ne sont pas indépendants : (1) implique (2), mais (2) n'implique pas (1). Cela devient plus clair lorsque nous identifions les questions auxquelles répondent ces éléments.

(1) répond à la question des véhicules de la connaissance. La connaissance est nécessairement véhiculée d'une manière ou d'une autre. La question se pose donc de savoir *comment*. La réponse que donnent Quine et une large part de l'épistémologie « traditionnelle » est que la connaissance est véhiculée par des phrases. Par contre, si on répond à la question des véhicules en postulant un véhicule qui ne peut exprimer de propositions (un vecteur d'activation neurale, par exemple), cela contraint les réponses possibles à la question du contenu des véhicules.

(2) répond à la question de la nature du *contenu* des véhicules. Dans le cas qui nous occupe, le contenu est une (ou des) proposition(s). Mais une proposition peut être exprimée autrement que de manière phrastique (par un symbole ou en langage de la pensée, par exemple). De plus, les propositions ont des conditions de vérité qui, généralement, sont bivalentes.

Toutefois, et c'est là le point important, une phrase véhiculant une connaissance ne peut exprimer autre chose qu'une proposition, même si une proposition peut être exprimée autrement. Le problème de Churchland avec (1) est qu'il s'agit d'une perspective superficielle : il existe une kyrielle d'exemples où il y a utilisation de connaissance n'impliquant pas de phrases. Il y a premièrement les nombreuses instances de comportements indiquant la présence de connaissances pratiques observées chez des enfants en bas âge (ramper) et même chez des adultes (marcher, attraper une balle au vol). Il s'agit d'instances de connaissance non phrastique (Churchland 2007). Deuxièmement, nombre d'animaux non-humains ont des comportements assez complexes pour justifier qu'on leur attribue des connaissances (*ibidem*, voir également Kornblith 2002 ainsi que section 2.1.2 du présent mémoire). La conception phrastique ne semble donc pas appropriée à l'explication causale de la connaissance. Il existe d'autres arguments, que nous présentons plus en détails dans les sections suivantes. L'important est de comprendre que pour Churchland l'approche phrastique est insuffisante au point de vue explicatif. Donc, Churchland rejette (1) et propose une réponse alternative à la question des véhicules. Cette réponse est que la connaissance est véhiculée par des vecteurs d'activation neurale (Churchland 1989, 1992).

Le rejet de la perspective propositionnelle par Churchland repose sur deux choses : 1) les propositions dépendent d'une condition de vérité bivalente et 2) sa réponse à la question des véhicules, soit que ceux-ci sont des vecteurs d'activation neurale réalisées dans un réseau de neurones (Churchland 1992), élimine la possibilité d'un contenu propositionnel. Ces réseaux apprennent à faire des discriminations de stimuli complexes et le taux de réussite augmente proportionnellement à l'apprentissage. Mais le critère de validité épistémique n'est plus un critère de vérité bivalent : il s'agit d'un critère de réduction d'erreur multivalent³³.

Nous allons élaborer sur ces points dans ce qui suit. Pour le moment, il faut simplement comprendre que la position de Churchland découle d'un engagement théorique à expliquer causalement l'émergence de la connaissance. Au point de vue de la question des véhicules, ceci implique de s'intéresser aux véhicules qui entretiennent les relations causales pertinentes. En l'occurrence, les phrases ne sont pas appropriées. Concernant la question de la nature propositionnelle de la connaissance, ce que ces véhicules contiennent ne peut être propositionnel, à cause du problème de la bivalence du critère de la vérité³⁴.

Dans ce qui suit, nous présentons quatre critiques adressées au paradigme phrastique par Churchland. Celles-ci nous permettent de mieux comprendre sa position.

³³ Nous devons ici prendre une certaine distance vis-à-vis des arguments de Churchland. En effet, de nombreux développements contemporains en logique, applicables à l'épistémologie du phrastique, intègrent des critères de vérité multivalents. Comme notre objectif ici est de mettre en lumière le développement du débat entre les tenants du paradigme phrastique et les critiques de ce paradigme, nous n'allons pas discuter de cette faiblesse de l'argument de Churchland plus en détail, mais il est important de la relever.

³⁴ Nous verrons au chapitre III qu'il est possible de suggérer un critère de vérité qui peut être vu comme bivalent, mais qui est cohérent avec l'engagement théorique que nous décrivons ici.

2.2.1. Problème 1 : le développement

Le premier problème est celui des stades précoces du développement humain. Churchland formule un argument clair et relativement simple pour montrer que les approches phrastiques ne sont pas adéquates pour expliquer le développement des capacités épistémiques humaines :

- a) « Le développement intellectuel humain n'est *pas* essentiellement ou uniquement phrastique, et il ne peut être représenté ainsi, parce que les états phrastiques requièrent le langage³⁵ ».
- b) Les mêmes processus sont impliqués dans le développement intellectuel des enfants en bas âge que dans l'acquisition de connaissance chez l'adulte, il n'y a pas de différence fondamentale.
- c) Compte tenu de i) et ii), « les états phrastiques ne peuvent figurer parmi les paramètres primitifs inclus dans une théorie adéquate du développement intellectuel rationnel, et la pertinence des états phrastiques est superficielle, ou au mieux dérivée. » (Churchland 1979, p. 128, notre traduction).

Il faut développer quelque peu cet argument. S'il est valide, il pose un problème descriptif *et* normatif aux approches traditionnelles. Nous savons que bien avant d'acquérir des capacités langagières, l'enfant est en mesure de catégoriser des objets, de comprendre des relations causales de base et de résoudre des problèmes (Livingston 1996). En d'autres termes, les jeunes enfants humains possèdent une compréhension du monde, sans avoir de capacités

³⁵ Fodor (1975) répondrait qu'il existe, même chez nouveau-nés, un langage de la pensée, et donc qu'il est toujours possible de représenter le développement et les opérations intellectuelles humaines par des états phrastiques. Il pourrait également affirmer que les phrases sont en *mentalese*, de sorte que l'enfant peut employer des phrases sans avoir de connaissance du langage naturel. Nous ne discuterons pas de cette possibilité ici, et ce pour deux raisons. Premièrement, nous nous intéressons aux théories épistémologiques, et non aux théories représentationnelles de l'esprit. Deuxièmement, la théorie de Fodor est aux prises avec ses propres problèmes. Susan Schneider (2009) relève, par exemple, que la théorie, qui repose sur l'existence de symboles mentaux manipulés par les processus cognitifs, doit encore fournir une conception plausible de ces symboles.

langagières. Il n'est pas clair *comment* l'enfant arrive à manipuler des phrases, si ce n'est par l'opération de facultés plus fondamentales, qui ne requièrent pas le langage. Churchland ne nie pas que *certaines* instances de connaissance puissent, à la limite, être représentées de manière métaphorique, ou dérivée, par des phrases, mais ceci requiert la possession de capacités langagières (Churchland 1979). Le paradigme phrastique est donc aux prises avec le problème suivant : il ne peut expliquer 1) les capacités épistémiques des jeunes enfants, ni 2) comment les humains acquièrent les capacités langagières (à moins qu'il ne soit possible d'expliquer l'acquisition d'une capacité en faisant intervenir des entités et des propriétés qui dépendent de la possession de cette capacité). Ceci a trois implications pour les approches phrastiques :

- 1) Implication descriptive A : les approches phrastiques ne peuvent fournir d'explication des capacités des enfants en bas âge.
- 2) Implication descriptive B : Les approches phrastiques excluent les enfants en bas âge de la catégorie des agents épistémiques : si les états et les signaux d'entrée des agents épistémiques sont des phrases, alors tout organisme ne pouvant utiliser de phrases ne peut avoir d'états épistémiques.
- 3) Implication normative : si la terminologie phrastique est inadéquate pour décrire la connaissance et les processus par lesquels nous l'acquérons, elle a peu de chances de fournir des normes épistémologiques particulièrement intéressantes.

2.2.2. Problème 2 : connaissance animale

Le second problème est celui de la connaissance animale. Nous prenons comme donné que les seuls animaux qui emploient des phrases sont les humains³⁶. Pourtant, suivant

³⁶ Encore une fois, l'hypothèse du langage de la pensée de Fodor (1975) pourrait secourir l'approche phrastique. Le langage de la pensée de Fodor permettrait de dire que l'épistémologie phrastique fonctionne tout à fait bien si on adopte l'hypothèse du langage de la pensée, qui permettrait d'exprimer des phrases en un langage interne, propre aux opérations computationnelles d'un agent épistémique. Par contre, l'hypothèse du langage de la pensée et la théorie représentationnelle de l'esprit qui l'accompagne font face, depuis longtemps, au sérieux problème de formuler une conception plausible

l'argumentaire de Kornblith (2002) présenté en 2.1.2, les animaux peuvent et doivent être considérés comme des agents épistémiques. Toutefois, en vertu de l'engagement de Churchland à expliquer *causalement* la connaissance, l'argument de Kornblith est irrecevable (même si Churchland partage l'idée que les animaux non-humains sont des agents épistémiques, cf. Livingston 1996), parce qu'il s'agit d'un argument basé sur l'utilité explicative de la terminologie phrastique et intentionnelle. Churchland rejette cette idée parce que l'utilité de l'approche phrastique est au mieux dérivée, et très limitée (Churchland 1979).

Il faut discuter un peu plus des raisons pour lesquelles l'argument de l'utilité explicative est irrecevable. Prenons comme point de départ l'idée que les animaux peuvent être des agents épistémiques. Cette idée est mise à profit de manière intéressante par Patricia Churchland (1980). Elle prend au sérieux l'idée que « *what makes intelligent beings intelligent may only be attainable through the physics of the substrata* » (Churchland 1980, p.163). Churchland introduit ainsi l'idée que l'étude de la connaissance doit, comme le suggère Kornblith (2002), se concentrer sur l'identification d'une espèce naturelle. Toutefois, contrairement à Kornblith, elle pense que l'espèce naturelle pertinente n'est pas la connaissance, mais l'agent épistémique qui possède ce que l'on nomme « connaissance »³⁷. Ceci dirige naturellement l'attention vers les ressemblances *physiques* entre les humains et les autres animaux. Les humains ne seraient qu'une instance parmi d'autres de systèmes tombant sous l'extension de la catégorie « agents épistémiques ». Prenant pour acquis que le cerveau humain est le produit de l'évolution et que celui-ci présente des similarités

des représentations mentales nécessaires au fonctionnement d'une théorie représentationnelle (voir Schneider 2009). Comme cette question n'est toujours pas réglée, nous renvoyons le fardeau de la preuve du côté des tenants de l'hypothèse du langage de la pensée. Ceci est motivé par le fait que nous disposons maintenant de modèles de l'esprit n'ayant aucun besoin de postuler des représentations au sens où l'entend Fodor, et que la théorie représentationnelle de l'esprit n'est plus nécessaire.

³⁷ Nous ne développons pas en détail cette idée. Il y a cependant de bonnes raisons de croire que considérer tout système capable d'états épistémiques comme une seule espèce naturelle est erronée. Toutefois, l'intuition fondamentale derrière cette idée, soit que pour comprendre les capacités cognitives humaines il est nécessaire de s'intéresser à ce que nous avons en commun avec d'autres organismes, est importante.

importantes avec celui d'autres animaux, il semble que nous ayons de bonnes raisons de croire que l'explication des capacités épistémiques humaines doit passer, d'une manière ou d'une autre, par l'étude du substrat physique de ces capacités. Cette étude doit, aussi, passer par l'examen des similarités entre humains et animaux non-humains. Par exemple, nous savons que certains oiseaux chanteurs possèdent des expressions du gène *FoxP2* semblables à celles des humains (Webb et Zhang 2005). De plus, les études neurologiques sur des singes ont été très importantes dans les recherches sur le traitement de l'information visuelle (Pohl 1973). Il semble donc qu'il soit important, pour comprendre les capacités humaines, de s'intéresser aux fondements neurologiques de ces capacités.

Si une partie de la recherche sur la connaissance se fait au niveau neurobiologique, alors le paradigme phrastique est insuffisant. Notons également que, sous cette interprétation, l'étude de la connaissance doit se situer à ce niveau. Il s'agit d'une idée centrale à l'éliminativisme de Paul Churchland (1979, 1988, 1989, 2007).

Le défenseur du paradigme phrastique pourrait, à ce stade, affirmer que ce qui devrait nous intéresser est le fonctionnement de la connaissance chez une humaine adulte, qui possède le langage, parce que le langage est ce qui distingue l'être humain d'autres organismes. Notons que ce type d'argument ne pourrait venir de Kornblith, et que nous ne lui attribuons aucunement une telle idée. Il reste que, même si le langage est effectivement une capacité cruciale et une bouée de sauvetage pour l'approche phrastique (comme nous le verrons en 2.4), il y a, là aussi, un problème.

2.2.3. Problèmes 3 et 4 : cognition adulte et vérité

Un défenseur du paradigme phrastique pourrait objecter que ce qui est vraiment intéressant pour l'épistémologie est l'aspect mélioratif. En ce sens, il serait possible d'écarter les problèmes dont nous avons discutés jusqu'à présent en affirmant que le paradigme phrastique demeure pertinent lorsqu'il s'agit de fournir des normes régissant l'acquisition et l'utilisation de connaissance par des êtres humains adultes, rationnels et en possession de capacités langagières normales. Cette version de l'approche phrastique se heurte à deux

problèmes, le premier étant celui de la cognition adulte. Le second concerne la notion de vérité.

Rappelons que le paradigme auquel s'attaque Churchland est caractérisé par l'idée que le raisonnement suit des règles logiques³⁸ (Livingston 1996, voir également l'introduction de la présente section). Suivant l'idée que les phrases sont traitées de manière explicite par les agents épistémiques, le modèle du raisonnement qui accompagne le paradigme phrastique en est un qui impose des principes de rationalité classiques à ces agents. Ce type de conception de la rationalité est, toutefois, de moins en moins plausible.

La psychologie expérimentale (Kahneman, Slovic et Tversky 1982, par exemple) suggère que les humains emploient souvent des heuristiques, des raccourcis cognitifs minimisant l'effort requis pour résoudre un problème de décision plutôt que des raisonnements explicites fondés sur des règles. Gigerenzer soutient que ces procédés reflètent le fonctionnement normal d'un système fonctionnant selon des principes de rationalité écologique (Gigerenzer, Todd et ABC Research Group. 1999).

La rationalité écologique correspond à une version modifiée de la rationalité limitée de Simon (1956, 1990), développée en réaction aux conceptions classiques de la rationalité humaine qui supposaient que l'esprit humain avait à sa disposition des ressources computationnelles et temporelles illimitées. Simon rejette cette idée et développe une conception de la rationalité tenant compte des limites de l'esprit humain et qui, plutôt que de reposer sur un critère d'optimalité, repose sur un critère de satisfaction : dans le cadre de la résolution de problème, on établit un seuil de satisfaction et on cesse la recherche de solution une fois qu'une solution qui excède ce seuil est trouvée (Simon 1956, 1990). Gigerenzer modifie cette définition de deux manières. Premièrement, afin d'expliquer pourquoi certaines heuristiques fonctionnent bien dans certaines situations, il donne plus d'importance à la

³⁸ La position de Kornblith échappe probablement à ceci.

structure de l'environnement (une notion présente chez Simon mais selon Gigerenzer trop peu explicitée). Deuxièmement, il distingue deux types de rationalité limitée : celle reposant sur la satisfaction et celle qu'il appelle « rapide et frugale », reposant sur des règles d'arrêt de recherche et de décision simples. La littérature sur les heuristiques remet donc en question l'idée que le raisonnement humain repose sur les règles de la logique traditionnelle. La majeure partie de notre raisonnement semble plutôt reposer sur des procédés heuristiques. De plus, comme nous le verrons au chapitre III, il existe des situations dans lesquelles un système fortement biaisé, reposant sur plusieurs heuristiques spécialisées, produit des performances prédictives largement supérieures qu'un système général moins biaisé (Gigerenzer et Brighton 2009). Ceci implique que la défense habituelle du modèle classique, soit qu'il ne s'agit que d'un idéal normatif, n'a pas beaucoup de force, puisque même ce modèle idéal est moins efficace (dans certains contextes) qu'un système reposant sur des heuristiques. En tant qu'organismes devant naviguer dans un environnement complexe et changeant, avoir de bonnes performances prédictives est, nous présumons, important.

Une charge supplémentaire contre le modèle classique du raisonnement vient des théories à processus duaux. Ces théories suggèrent que l'esprit humain est double : il existe soit deux systèmes cognitifs ou deux types de processus cognitifs. Le système 1 (ou les processus de type 1) emploie des heuristiques et est associatif et rapide, alors que le système 2 (ou les processus de type 2) est impliqué dans des raisonnements complexes employant des règles logiques et demandant plus d'effort (voir le chapitre introductif de Evans et Frankish 2009 pour un survol de ces théories). En fonction de l'importance accordée au système 1, l'on se retrouve avec une vision de la cognition humaine qui donne une place plus ou moins importante aux raisonnements explicites. Considérant que la littérature psychologique sur les biais et les heuristiques (dont il sera question au chapitre III) suggère qu'une large part de nos décisions sont en un certain sens « prises » par le système 1 (ou des processus de type 1), le paradigme phrastique (en tant que perspective reposant sur une certaine conception du raisonnement) essuie un revers supplémentaire. Il essuie un revers supplémentaire parce que, même s'il est raisonnable de croire que les processus de type 2 jouent un rôle important dans la cognition humaine, il demeure qu'une part plus ou moins importante de nos tâches

cognitives sont traitées par des processus de type 1. Ceux-ci ne sont *pas* régis par des règles logiques explicites³⁹.

Ceci nous amène à considérer, très rapidement, le problème de la vérité. Rappelons que Churchland rejette, de par sa réponse à la question des véhicules, l'idée qu'un critère de vérité bivalent soit approprié (Churchland 1992, p.421). Il existe une autre raison de douter de la validité explicative du critère traditionnel de vérité.

Succinctement, la notion correspondantiste de la vérité n'est pas un critère approprié pour l'évaluation de certaines formes de connaissance. Comme nous fournissons un argumentaire détaillé en faveur de cette idée au chapitre III, nous nous limiterons ici à une présentation schématique du problème. L'important est simplement de relever que les procédés heuristiques visant à faire des prédictions ne visent pas à trouver une réponse qui corresponde à la réalité, mais plutôt à diminuer le nombre d'erreurs dont le coût est le plus élevé. Lorsqu'un lapin doit évaluer la probabilité de tomber sur un prédateur, celui-ci peut se tromper de deux manières. Soit il fait une erreur de type 1 (faux positif) ou une erreur de type 2 (faux négatif). Un faux négatif a un coût plus élevé qu'un faux positif. Le faux positif a comme conséquence, par exemple, un détour inutile. Un faux négatif a comme conséquence une rencontre potentiellement désastreuse avec un prédateur. Si une large part de nos raisonnements repose sur des procédés heuristiques, et que l'intuition fondamentale des théories à processus duaux est juste, alors la notion correspondantiste de vérité joue un rôle beaucoup moins central que ce que supposent certains tenants du paradigme phrastique. La psychologie évolutionniste remet aussi en question le primat de la notion correspondantiste de vérité. Selon Cosmides, Tooby et Barrett (2005), par exemple, notre système cognitif est composé de modules spécialisés pistant des solutions précises à des problèmes adaptatifs

³⁹ Notons que nous employons ici la distinction entre processus de type 1 et 2 de manière heuristique : nous ne souscrivons à aucune version précise des théories à processus duaux. De plus, il faut noter que certaines versions de ces théories n'excluent pas que certains processus de type 1 emploient quelque chose comme la connaissance phrastique. Il demeure, toutefois, que les processus de type 1 reposent sur des

précis. Sous cette interprétation, il n'existe pas de critère général pour le succès comportemental, et la notion de vérité ne serait pas centrale à une large part de nos processus cognitifs⁴⁰.

Nous avons donc une bonne idée des critiques du paradigme phrastique par Churchland.

2.2.4. Synthèse de la position naturaliste de Churchland : un naturalisme de remplacement

Churchland développe donc quatre critiques visant à justifier l'élimination du paradigme phrastique. Parce que celui-ci ne peut expliquer les capacités épistémiques des enfants en bas âge, ne peut rendre compte de manière satisfaisante des capacités épistémiques des animaux non-humains, repose sur un modèle irréaliste du raisonnement humain, et finalement présuppose d'un critère de vérité dont l'importance réelle est moindre, il suit que le paradigme phrastique est, tout simplement, superficiel.

Pour reprendre la division de l'épistémologie naturalisée proposée par Feldman (2001), Churchland se place dans la catégorie du naturalisme de remplacement. Sa position est toutefois fondamentalement différente de celle que Quine. En effet, Churchland argumente en faveur du remplacement de l'épistémologie « phrastique » par les disciplines neurologiques, qui seules peuvent expliquer correctement le phénomène de la connaissance. Il faut noter qu'il s'agit là du pendant épistémologique d'une position plus générale : l'éliminativisme matérialiste, qui affirme que nos conceptions folkloriques, et les méthodes et terminologies qui en découlent, doivent être éliminées. Ceci implique d'éliminer les notions de croyance, de vérité, d'attitudes propositionnelles, et ainsi de suite. Dans le cadre de l'épistémologie à proprement parler, Churchland reste cohérent avec l'engagement théorique de Quine, soit que

⁴⁰ L'argument évolutionniste présenté (très schématiquement) ici est développé et profondément remanié au chapitre III.

l'étude de la connaissance doit se concentrer sur l'explication causale de l'émergence de celle-ci. La différence, cruciale, est que Churchland rejette l'approche phrastique sur la base de sa faiblesse explicative. Il rejette par la même occasion le naturalisme coopératif de Kornblith, parce que celui-ci repose sur l'attribution aux animaux d'états mentaux qui, selon sa perspective, doivent être éliminés. Finalement, il rejette le naturalisme substantif parce que chercher la définition naturelle des référents de termes évaluatifs propres à l'épistémologie, tels que « justifiée » et « a de bonnes raisons » est une entreprise futile : ces termes seront éliminés.

Il faut mentionner que nous sommes, en général, sympathiques aux critiques de Churchland, mais *uniquement* dans la mesure où celles-ci démontrent que le paradigme phrastique n'est pas *suffisant* pour expliquer la connaissance. Nous présentons plus loin un argument *contre* l'éliminativisme de Churchland.

2.3. Métathéorie : caractérisation générale du débat

Dans cette section, nous clarifions notre perspective métathéorique. Ceci nous permet d'arriver à une formulation claire du débat que nous voulons dépasser. Notons que nous lorsque nous employons l'expression « éliminativisme de Churchland », nous parlons de l'éliminativisme à l'endroit de la terminologie phrastique.

Rappelons, avant tout, que nous nous intéressons surtout à la prépondérance du paradigme phrastique au sein du naturalisme épistémologique. Ce choix est motivé par le fait que les approches naturalistes tendent à privilégier l'idée de proposer une espèce naturelle « connaissance ». Nous sommes donc plus proches des conditions contextuelles dans lesquelles doit être proposée une hypothèse d'hétérogénéité. Notre argument s'applique *mutatis mutandis* si nous incorporons les approches « traditionnelles ».

La dichotomie que nous voulons mettre en évidence ici est celle qui découle de la position éliminativiste de Churchland à l'endroit de la terminologie phrastique.

L'épistémologie « traditionnelle », ainsi qu'une large part de l'épistémologie naturalisée, adopte ce que nous avons appelé le « paradigme phrastique ». Churchland (1979, 1989, 1992, 2007) soulève des objections contre les trois composantes de ce paradigme (présentées en 2.2).

Il s'agit d'une dichotomie qui s'exprime dans le cadre du débat sur la nature de la connaissance. Pour étudier la connaissance, l'idée généralement admise est qu'il faut commencer par fixer le niveau descriptif approprié à l'étude de celle-ci. Le choix d'un niveau descriptif est un choix qui a des implications importantes : il détermine les méthodes pertinentes et, surtout, les unités de base pertinentes à l'étude d'un phénomène donné. Dans le cas de la connaissance, il y a un « camp » qui maintient que le niveau descriptif approprié est le niveau phrastique. Un second camp maintient que le niveau approprié est le niveau neurologique, en affirmant que le niveau phrastique est inutile et que la terminologie qui lui est propre doit en conséquence être éliminée. Le débat qui nous intéresse ici est donc le débat sur le niveau descriptif approprié à l'étude de la connaissance.

Ce débat est caractérisé par l'idée que la solution au problème de la nature de la connaissance est une solution *moniste*, soit qu'il y a quelque chose comme La Nature de la Connaissance. Pour Churchland, cette nature unique de la connaissance correspond à des vecteurs d'activation neurale, et s'il y a une espèce naturelle de la connaissance, il s'agit d'un certain type de vecteurs d'activation (ayant certaines propriétés), ou de l'information véhiculée par ceux-ci — qui n'est définitivement *pas* phrastique, ni propositionnelle.

Pour d'autres, comme Kornblith, l'espèce naturelle « connaissance » correspond à des « croyances vraies produites de manière fiable », idée rejetée *de facto* par la position de Churchland. Une autre position, défendue par Murray Clarke (2004), affirme que la connaissance correspond à plusieurs espèces naturelles, mais celles-ci restent au niveau phrastique et donc seraient compatibles avec une approche comme celle de Kornblith, mais pas avec une approche éliminativiste à la Churchland. De la même manière, certaines approches « traditionnelles » ou « intellectualistes » proposent des définitions considérées comme exhaustives. Williamson (2000) propose par exemple que la connaissance correspond

à l'état factif le plus général, et que toute forme de connaissance (incluant, crucialement, le savoir-faire) se réduit à cette définition (Stanley et Williamson 2001).

Le problème est qu'il existe de bonnes raisons de penser que pour expliquer le phénomène de la connaissance, il est nécessaire d'employer non seulement des approches situées au niveau neurologique, mais également des approches situées au niveau psychologique (comme celle de Kornblith). Il est même nécessaire d'employer, dans certains contextes, des approches se situant hors du naturalisme épistémologique. Il en résulte donc que l'idée selon laquelle la question de la nature de la connaissance doit recevoir une réponse moniste⁴¹ est erronée : elle donne lieu à un débat stérile parce qu'il n'existe pas de nature unique de la connaissance. Le terme renvoie à des entités fondamentalement différentes, appartenant à des niveaux explicatifs différents. Pour en faire un slogan : le terme « connaissance » est une instance de référence partielle.

2.4. Hétérogénéité : conditions contextuelles et argument de nécessité explicative

Cette section reprend la marche à suivre pour l'élaboration d'une hypothèse d'hétérogénéité, telle que définie dans le chapitre I. Nous commençons par montrer que les conditions contextuelles sont remplies. Ensuite, nous remplissons la première partie de la condition méthodologique en fournissant des raisons conceptuelles de remettre en question la présupposition d'unicité à l'endroit de la connaissance. Finalement, nous remplissons la seconde partie de la condition méthodologique en énonçant un argument de nécessité explicative.

⁴¹ Pour ce qui est des propositions naturalistes, cette idée correspond à une hypothèse d'espèce naturelle. Si nous incluons les approches « traditionnelles », il s'agira plutôt d'une « présupposition d'unicité », soit un procédé analogue à une hypothèse d'espèce naturelle.

2.4.1. Comment le débat sur la nature de la connaissance remplit les conditions contextuelles

Commençons par rappeler les deux conditions contextuelles à remplir lors de l'élaboration d'une hypothèse d'hétérogénéité (voir chapitre I).

1. Condition 1 : le terme, le concept ou la catégorie qui fait l'objet de l'hypothèse d'hétérogénéité doit également faire l'objet d'une hypothèse d'espèce naturelle ou d'un procédé analogue (comme une présupposition d'unicité), dont les modalités et conséquences sont proches de celles d'une hypothèse d'espèce naturelle.
2. Condition 2 : l'hypothèse d'espèce naturelle (ou la présupposition d'homogénéité) doit faire l'objet d'un consensus dans la discipline concernée.

Le terme « connaissance », dans le contexte du débat que nous avons décrit plus haut, remplit ces deux conditions. Chaque camp considère que la connaissance ne correspond qu'à un seul type de véhicule informationnel. Cette présupposition a un effet analogue à celui de l'adoption d'une hypothèse d'espèce naturelle. En fait, si nous restreignons notre argumentaire au naturalisme épistémologique, nous sommes en présence d'une hypothèse d'espèce naturelle, bien qu'elle soit moins claire que celle concernant les concepts, parce que les différentes candidates se situent à des niveaux descriptifs différents. Si nous incluons les approches non naturalistes, il s'agit d'une présupposition d'unicité de la connaissance, qui a un effet analogue : elle cause la recherche d'une nature unique de la connaissance et donc l'émergence de débats sur la nature de celle-ci.

La seconde condition est remplie par le fait que tant les approches éliminativistes que les approches phrastiques argumentent en faveur de l'adoption de leur espèce naturelle de prédilection. Cette situation se manifeste parce que l'idée que « connaissance » doit référer à une espèce naturelle fait consensus (sauf dans le cas de Clarke, qui suppose qu'il y a plusieurs espèces naturelles). Rappelons qu'une hypothèse d'espèce naturelle affirme qu'un terme T du vocabulaire scientifique réfère à une catégorie naturelle. Le fait que personne ne s'entende sur la question de savoir de quelle catégorie il s'agit est plutôt un symptôme d'une hypothèse d'espèce naturelle erronée.

2.4.2. Condition méthodologique, partie 1 : raisons conceptuelles de proposer l'hétérogénéité; assises de l'argument de nécessité explicative

D'emblée, nous acceptons l'idée que pour expliquer la connaissance, il faut s'intéresser aux véhicules qui entretiennent, dans les faits, les relations causales recherchées. Cette idée, comme nous l'avons vu, fait beaucoup de travail pour les approches naturalistes en général, et surtout pour l'éliminativisme de Churchland. Toutefois, elle se retourne contre lui. En effet, si l'on pousse plus loin l'idée qu'il faut se concentrer sur les véhicules informationnels entretenant des relations causales pertinentes, on se rend compte que les arguments de Churchland, s'ils sont très convaincants dans la mesure où ils démontrent que l'approche phrastique n'est pas suffisante, ne sont pas assez forts pour justifier l'élimination de cette approche. La raison est simple : il existe des véhicules informationnels qui *sont* phrastiques, qui jouent un rôle causal *en vertu* de leur forme phrastique et de leur contenu propositionnel, et dont l'apport ne peut être négligé. Ces véhicules existent parce que l'être humain dispose du langage.

Dans une revue critique de *Doing Without Concepts* de Machery (2009), Poirier et Beaulac (2011) discutent d'une conséquence de l'adoption de l'hypothèse d'hétérogénéité : les définitions font un retour comme quatrième terme remplaçant « concept ». En adoptant une version minimale des théories à processus duaux, ils argumentent que les exemplaires, prototypes et théories (les trois types de concepts identifiés par Machery) sont des concepts propres aux processus de type 1 (inconscients, associatifs, automatiques, etc.), mais qu'il existe aussi des concepts employés par des processus de type 2 (explicites, contrôlés, etc.) : les définitions. Ces concepts, comme les concepts de type 1, regroupent des membres d'une même catégorie, et ils sont construits par des processus de type 2. À l'instar de Machery, Poirier et Beaulac construisent un argument de nécessité explicative pour justifier l'introduction des définitions. Nous savons que certains concepts sont acquis automatiquement à travers nos interactions avec le monde. Ceux-ci sont des concepts de type 1, soit des prototypes, des exemplaires et des théories. Mais il existe des cas où nous devons faire des distinctions fines entre des objets abstraits, ou des cas où nous acquérons le concept

d'un objet sans avoir de contact direct avec celui-ci, et les concepts de type 1 ne peuvent rendre compte de ces cas. De plus, les processus de type 1 et de type 2 ne sont pas isolés les uns des autres. Les processus de type 1 peuvent influencer les processus de type 2, et vice-versa. C'est la raison pour laquelle il faut introduire les définitions, soit des corps d'information (véhicules informationnels) construits et acquis de manière explicite et réflexive. Ce type de corps d'information prend, souvent, la forme d'ensembles de conditions nécessaires et suffisantes construits et transmis verbalement. De plus, ceux-ci sont appliqués par la remémoration de l'énoncé verbal de la définition (Poirier et Beaulac, 2011).

Les définitions sont d'abord des connaissances. Elles servent aussi à introduire certains concepts dans le cadre de la formation scientifique (on les retrouvera, par exemple, dans des manuels). Nous pouvons également employer des définitions pour acquérir le concept d'une catégorie avec laquelle nous n'avons eu aucun contact. Par exemple, Dwight n'a jamais eu de contact avec des takoyakis, mais grâce à Wikipedia, qui lui en fournit la définition, il sait qu'il s'agit d'un mets japonais, spécialité de la région d'Osaka, fabriqué à partir d'une pâte sucrée, de forme ronde, et contenant du poulpe grillé. Une large part des concepts acquis après l'enfance (surtout dans le milieu scolaire et plus tard professionnel) sont acquis par l'entremise de définitions énoncées verbalement. Les définitions jouent également un rôle normatif important en sciences, et permettent le processus de révision conceptuelle d'un terme du vocabulaire scientifique suite à de nouvelles découvertes (la redéfinition de « planète » par l'UAI, en 2006, est un bon exemple).

Les définitions ainsi conçues sont, dans les faits, des véhicules informationnels phrastiques. Leur existence est due au langage, et il y a fort à parier que nos capacités épistémiques, bien qu'elles soient liées à celles d'autres animaux au point de vue anatomique et fonctionnel, subissent l'influence du langage.

Un autre exemple de la pertinence de l'approche phrastique est le travail scientifique, dans lequel, même si nous employons des heuristiques de recherche, les règles de la logique doivent être observées, parce que l'objectif est souvent de fournir une explication complète d'un phénomène et dans de tels cas, il faut plus que la simple réduction d'erreur : nous

devons formuler des explications et des théories qui cadrent avec la réalité. De plus, l'entreprise scientifique est éminemment sociale, et repose donc, en partie, sur les avantages que nous confère le langage au point de vue de l'échange rapide d'information. Les disciplines académiques sont également un contexte où l'approche phrastique peut avoir certaines applications, parce que les définitions jouent un rôle tout aussi important en philosophie qu'en sciences ou qu'en histoire de l'art ou en architecture. Finalement, lorsqu'il est question d'expliquer, ou de décrire, nos pratiques quotidiennes reposant sur la notion de connaissance, il est fort probable que les approches phrastiques soient plus appropriées qu'une approche neuronale. À noter que même dans ces cas, il y aura probablement des interactions entre ces connaissances phrastiques et les connaissances de « de bas niveau », du même type que les interactions entre processus de type 1 et 2 présumées par Poirier et Beaulac (2011)⁴².

Toutes ces considérations, de concert avec notre engagement à expliquer causalement la connaissance, nous forcent à rejeter l'élimination de la terminologie phrastique. Cette approche se heurte à un problème d'insuffisance explicative : elle élimine des notions qui dans un certain contexte explicatif, sont nécessaires.

Une objection que pourrait faire un défenseur de l'éliminativisme est que le paradigme phrastique reste superficiel et que les capacités de plus haut niveau qu'il couvre sont réductibles aux processus de plus bas niveau. Cette objection est vite traitée. Les définitions, et autres véhicules phrastiques, sont parfois nécessaires au développement de capacités que nous pourrions considérer comme de plus bas niveau. Par exemple, apprendre le calcul propositionnel requiert, avant tout, d'apprendre des définitions. Mais plus quelqu'un pratique le calcul propositionnel, plus cela devient automatique, et les définitions ne sont plus, comme

⁴² Ceci suggère qu'une conception hétérogène de la connaissance adéquate devrait pouvoir rendre compte de ces interactions. Puisque notre objectif est de suggérer une telle conception, nous ne proposons aucunement un simple pluralisme méthodologique où chaque discipline travaille en isolation.

au début, remémorées chaque fois qu'il faut résoudre un problème de logique. Cela devient un automatisme. Le développement de nos capacités n'est pas à sens unique, mais se fait autant du bas vers le haut que du haut vers le bas. Le processus est évidemment beaucoup plus complexe que ce que nous avons décrit ici. L'apprentissage des définitions contrôle probablement (ou influence) les pratiques (comme faire des problèmes de logique) qui elles vont mener à la formation d'exemplaires et de prototypes.

2.4.3. Condition méthodologique, partie 2 : argument de nécessité explicative

Nous pouvons à présent formuler notre argument de nécessité explicative. Pour ce faire, nous allons suivre une démarche analogue à celle de Machery (2005). La première étape est de poser que l'épistémologie, pour expliquer nos capacités épistémiques, doit postuler l'existence de connaissances, soit d'information véhiculée d'une certaine manière, étant d'un certain type et manipulée par des agents épistémiques. Deuxièmement, nous affirmons que la définition de la connaissance, soit la définition de la nature et des propriétés des véhicules informationnels pertinents, est fonction du niveau descriptif adopté, et que le choix de ce niveau descriptif est ce sur quoi porte le débat. Troisièmement, nous savons que nous avons de bonnes raisons de croire que pour expliquer l'ensemble du phénomène de la connaissance, nous devons avoir recours *et* aux approches phrastiques, *et* aux approches non phrastiques (ainsi, nous supposons, qu'à des approches intermédiaires). En conséquence, nos capacités épistémiques doivent être expliquées en ayant recours autant aux véhicules phrastiques et aux contenus propositionnels qu'aux véhicules non phrastiques et aux contenus non propositionnels.

S'il n'existe pas de nature unique de la connaissance, alors le débat sur la nature de la connaissance, dans sa forme actuelle, est un débat stérile. De plus, la dichotomie phrastique-non phrastique doit être brisée. Nous proposons, donc, que l'adoption d'une conception hétérogène de la connaissance permettrait non seulement de mettre fin à ce débat stérile, mais pourrait également fournir une manière de voir les différents types de connaissance comme s'imbriquant dans une sorte de continuum, ou d'organisation générale en fonction des

véhicules informationnels impliqués, des critères justificatifs et du type de « vérité » pertinent. Mais nous anticipons sur le prochain chapitre. Pour le moment, nous allons simplement formaliser notre argument, tel que nous l'avions fait dans l'introduction :

Argument de nécessité explicative à l'endroit de la connaissance

1. Pour expliquer nos capacités épistémiques, nous devons postuler la connaissance. (Prémisse 1)
2. Les constructions théoriques mises de l'avant par les théories phrastiques et non phrastiques⁴³ de la connaissance n'ont que très peu⁴⁴ de propriétés communes. (Prémisse 2, venant de l'analyse métathéorique)
3. Afin d'expliquer nos capacités cognitives de manière satisfaisante, nous devons employer les deux sortes de constructions théoriques (Prémisse 3, venant de l'analyse métathéorique et de l'exposition des raisons conceptuelles en 2.3.2)
4. Donc l'explication de nos capacités épistémiques requiert d'avoir recours à des constructions théoriques n'ayant que très peu de propriétés communes. (Conclusion 1)
5. Donc, il n'existe pas de nature unique de la connaissance (ou : la connaissance n'est pas une espèce naturelle). (Conclusion 2)

S'il n'existe pas de nature unique de la connaissance, le débat sur la nature de la connaissance est totalement stérile et doit être dépassé. Pour clore le chapitre, il nous faut identifier ce qu'il reste à faire pour que nous puissions affirmer qu'une manière intéressante de dépasser ce débat est d'adopter une conception hétérogène de la connaissance.

⁴³ Une autre version de l'argument pourrait inclure les approches non naturalistes dans les approches phrastiques.

⁴⁴ Elles n'ont en fait aucune propriété commune, puisque l'approche non phrastique de Churchland implique l'élimination de l'approche phrastique et de son appareil conceptuel.

2.5. Conclusion

Les objectifs du présent chapitre étaient les suivants : 1) montrer que le débat sur la nature de la connaissance remplit les conditions contextuelles identifiées dans le chapitre I et 2) développer un argument de nécessité explicative, en accord avec les paramètres identifiés dans le chapitre I. Ce qu'il nous reste à faire se divise en deux étapes. Il faut, premièrement, apporter un soutien empirique à ce que nous avons développé jusqu'à maintenant. Cette étape est complétée en trois phases. Il faut premièrement montrer que, *dans les faits*, nous sommes justifiés de considérer certains véhicules informationnels comme des instances de connaissance malgré le fait qu'ils ne sont pas phrastiques, et ne dépendent pas des critères traditionnels. Deuxièmement, nous devons montrer qu'il existe des types de connaissance qui requièrent d'être étudiés au niveau neurologique, alors que d'autres sont mieux compris au niveau psychologique et d'autres encore sont mieux compris au niveau phrastique. Finalement, nous devons montrer que ces différents types de connaissance ne peuvent être réduits les uns aux autres.

La seconde étape est de suggérer une version partielle, préliminaire, provisoire, et temporaire de la conception hétérogène de la connaissance. Celle-ci est, nous l'admettons, schématique et incomplète, mais nous en présentons assez pour qu'il soit possible d'entrevoir son utilité, ainsi que son potentiel de fournir une étude descriptive efficace de la connaissance sans évacuer la dimension méliorative de l'épistémologie.

CHAPITRE III

IMPLICATIONS ÉPISTÉMOLOGIQUES DE LA PSYCHOLOGIE ÉVOLUTIONNISTE ET DE LA NEUROPSYCHOLOGIE DE LA VISION; CONCEPTION HÉTÉROGÈNE DE LA CONNAISSANCE

Dans le chapitre précédent, nous avons formulé un argument de nécessité explicative qui débouche sur la conclusion qu'il n'existe pas de nature unique de la connaissance. Ceci implique que la forme traditionnelle de la solution au problème de la nature de la connaissance doit être repensée et que le débat, dans sa forme actuelle, est futile. Il s'agit de la thèse négative du mémoire.

La prémisse empirique de l'argument reste toutefois à recevoir un support adéquat. Fournir ce support est le premier objectif du présent chapitre. Le second objectif est de soutenir la thèse positive du mémoire, soit qu'une conception hétérogène pourrait permettre de dépasser le débat. Nous commençons par présenter quelques considérations théoriques nécessaires à la suite de l'argumentation (section 3.1). Il est ensuite question (section 3.2) des implications épistémologiques concernant la question de la vérité d'un argument évolutionniste bien connu (Tooby, Cosmides et Barrett 2005), que nous retravaillons et augmentons d'un volet empirique tiré de la psychologie du raisonnement (Gigerenzer et Brighton 2009). Nous abordons ensuite les recherches de Milner et Goodale (2008), qui suggèrent que certains véhicules informationnels pouvant être considérés comme des formes de connaissance ne sont pas phrastiques (section 3.3). Nous concluons le chapitre en réitérant notre argument de nécessité explicative et en suggérant une première ébauche d'une conception hétérogène de la connaissance (section 3.4).

3.1. Considérations préliminaires

Cette section a pour objectif de poser quelques conditions minimales servant à distinguer les instances d'attribution de connaissance pertinentes des cas « accidentels » (succès

comportemental aléatoire) ou simplement inintéressantes (reposant sur des explications farfelues ou impertinentes).

3.1.1. Conditions minimales

Nous identifions trois conditions minimales à l'attribution de connaissance à un organisme. La première condition est évidente et triviale : la structure informationnelle candidate à l'attribution de connaissance doit entretenir un lien avec le monde. Plus précisément, elle doit représenter, pour l'organisme, un état du monde, un résultat possible d'un comportement, etc. Par exemple, la représentation (peu importe sa forme exacte) que possède une fourmi engagée dans une opération de navigation à l'estime (*dead reckoning*) serait, en fonction de ce critère, une candidate appropriée au statut de connaissance, parce qu'elle représente (d'une manière pertinente à la fourmi) l'environnement dans lequel celle-ci évolue, la distance qu'elle a parcouru, etc. (Kornblith 2002 p.37).

La seconde condition est moins évidente quoiqu'aussi simple que la première : la structure informationnelle candidate doit, potentiellement, jouer un rôle dans l'explication du comportement de l'organisme. C'est-à-dire que lorsqu'il est question de formuler une explication du comportement de l'organisme (et surtout d'un comportement récurrent remplissant une fonction quelconque), il y a un gain explicatif clair à attribuer à cette structure informationnelle le statut de connaissance, compris comme une structure informationnelle employée par l'organisme pour guider son comportement. Nous empruntons l'idée générale de notre argument ici à Kornblith (2002). Pour reprendre le cas de la navigation à l'estime, lorsque nous tentons d'expliquer le comportement de la fourmi, il est plus fécond de formuler une explication supposant la possession d'une certaine forme de connaissance plutôt que de formuler une explication strictement physique de son comportement. Il s'agit d'une explication plus féconde parce que se limiter à une explication physique risque d'occulter certaines similarités entre des comportements qui du point de vue physique semblent hétérogènes alors qu'ils sont des instanciations particulières d'un même

type de comportement (Kornblith 2002, p. 34)⁴⁵. En d'autres termes, certaines structures informationnelles acquièrent le statut de connaissance à cause de leur rôle explicatif.

La troisième condition est mise en place pour anticiper l'objection selon laquelle ces critères définissent en fait des croyances et non des connaissances. Une croyance peut être produite de manière accidentelle ou ponctuelle. Le type de structure informationnelle dont il est ici question a ceci de particulier qu'il s'agit de structures produites de manière régulière et fiable en présence d'une tâche ou d'une situation particulière. Cette idée s'inspire de Kornblith (2002). Une croyance n'est pas définie par son rôle dans la régulation du comportement. Il y a évidemment des croyances qui jouent un tel rôle, et dans ce cas, celles-ci (si elles sont vraies) sont des instances de connaissance. La troisième condition est donc la suivante : ces structures informationnelles sont employées de manière fiable dans la régulation comportementale *et* nous devons leur accorder un statut épistémique pour expliquer le comportement d'un organisme donné.

Les formes de « connaissances » remplissant ces conditions sont en fait assez semblables à la notion d'information en sciences cognitives. Définie comme l'ensemble des entrées et des sorties d'un système cognitif (Decety et Houdé 2003), celle-ci correspond en gros à ce que nous cherchons à caractériser ici. Nous parlons, spécifiquement, de structures informationnelles qui sont *traitées* par l'organisme et *utilisées* par celui-ci dans la régulation

⁴⁵ Cette seconde condition ressemble aux défenses de la pertinence de l'explication intentionnelle (voir entre autres Dennett 1987). Rappelons que nous parlons, toutefois, d'attribuer des *connaissances*, soit des états informationnels remplissant certaines conditions. Nous avons recours, ici, à un argument ressemblant à celui pour l'attribution d'états intentionnels parce que ce type d'argument est utilisé pour justifier l'attribution d'états d'un certain type à un système donné. Nous voulons justifier l'importance d'attribuer des états épistémiques à des organismes desquels on pourrait dire qu'ils n'ont que des états informationnels simples. Nous affirmons, toutefois, qu'il est plus fécond de leur attribuer des états épistémiques, compris comme des états informationnels remplissant certaines conditions. Une manière de penser cet argument est de remplacer « état intentionnel » par « état informationnel remplissant certaines conditions ». En d'autres termes, nous postulons l'existence de structures informationnelles ayant une dimension épistémique (parce qu'elles remplissent certains critères). Nous demeurons agnostiques (et flexibles) sur leur forme.

de son comportement⁴⁶. De plus, elle doit être fiable, en accord avec le critère fiabiliste dont nous discutons plus bas. En un certain sens, nous tentons ici de rendre explicite une dimension épistémique de la notion d'information. Nous supposons que les sciences cognitives emploient souvent la notion d'information de cette manière, mais la dimension proprement épistémique de celle-ci est rarement explicite, surtout parce qu'il est difficile d'intégrer la notion d'information dans un cadre épistémologique « traditionnel ». Par contre, la perspective que nous développons ici accommode aisément cette notion.

On peut aisément imaginer que les types plus complexes de connaissance doivent rencontrer plus de critères⁴⁷ (celles-ci remplissent *de facto* les conditions minimales). L'idée, que nous développerons en 3.4, est que la connaissance correspond, *au moins*, à des structures informationnelles remplissant ces conditions. L'ajout de conditions supplémentaires distingue les connaissances de « bas niveau » des connaissances de « haut niveau ».

⁴⁶ La notion d'information que nous invoquons ici est elle-même minimale. Le concept d'information a fait l'objet de travaux importants. Notamment, Fred Dretske (1981) propose une théorie informationnelle assez influente, dont l'idée centrale du processus de « digitalisation » de l'information sensorielle brute (analogue) pourrait être appliquée, dans une certaine mesure, à l'explication du processus par lequel une information (ou une « connaissance de bas niveau ») se complexifie. Ce point est spéculatif, mais suggère une sorte de continuum, ou de spectre sur lequel s'alignent diverses formes de connaissances. Tout organisme ayant un système cognitif emploie de l'information. Certains organismes emploient cette information de manière fiable et ont un comportement qui suggère que l'information ainsi employée possède certaines propriétés qui reflètent des propriétés épistémiques pertinentes comme d'être (en un sens ou un autre) fiable, vraie, et ainsi de suite. De l'information, nous passons donc à des connaissances simples et ainsi de suite jusqu'aux formes très complexes que nous employons en tant qu'humains. Nous revenons sur cette idée dans la prochaine section, ainsi qu'en 3.4.

⁴⁷ Cette idée de l'augmentation du nombre et de la complexité des critères en fonction de la complexité et de la forme des structures informationnelles auxquelles ils s'appliquent suggère la possibilité de développer une échelle, ou une métrique, des différents types de connaissance. Les connaissances simples (comme celles de la fourmi) n'ont qu'à satisfaire des critères faibles et peu nombreux, alors que la connaissance exprimée par une spécialiste de la pensée de Kant doit satisfaire un ensemble beaucoup plus complexe et étendu de critères.

3.1.2. Justification; un premier axe de différenciation

La justification est, à l'image de la connaissance, hétérogène. Il n'existe pas de critère justificatif universel. Nous pouvons pour le moment en présenter deux. Le premier correspond à l'interprétation que fait Churchland (2007) du fiabilisme de Goldman : certaines « représentations » cognitives comptent comme des instances de connaissance à cause de la fiabilité des processus et mécanismes cognitifs qui ont dans les faits produit cette représentation (Goldman 1976, 1986). L'adoption de ce type de fiabilisme fournit un critère de pertinence épistémique qui dirige naturellement l'attention vers les mécanismes responsables de nos représentations (au sens large), et surtout vers les véhicules que ces mécanismes manipulent, produisent et modifient (Churchland 2007). Ce critère s'applique aux formes non-phrastiques de connaissance et est donc surtout pertinent dans le cadre naturaliste.

Le second type concerne les connaissances plus complexes vraisemblablement véhiculées de manière phrastique. Dans ce cas, il est plutôt question de critères de justification multiples, suivant le cas. Jim est justifié de croire que la force gravitationnelle de la Terre est ce qui cause l'évènement « mes clés tombent par terre si je les lâche » parce que la théorie de la gravitation est suffisamment bien corroborée, qu'il s'agit de la meilleure explication disponible, etc. Jim *sait* donc que ses clés tombent à cause de la gravité. Un autre cas est celui des théories scientifiques, où les sources de justification sont nombreuses (résultats expérimentaux, force explicative, indices bibliométriques, etc.). Un troisième cas est celui d'un énoncé portant, par exemple, sur l'interprétation correcte d'un passage dans un texte philosophique obscur : ce type de travail exégétique impose des exigences différentes selon le champ philosophique et même l'auteur étudié. De la même manière, les énoncés du sens commun sont évalués à la lumière d'autres critères. Par exemple, Clarke (2004) soutient que les critères traditionnels de vérité et justification sont nécessaires à l'explication de nos tendances à vouloir supporter nos affirmations et rechercher la vérité d'une affirmation. Finalement, comme nous l'avons mentionné au chapitre II, les définitions sont des ensembles de conditions nécessaires et suffisantes véhiculées par le langage naturel. Ces structures informationnelles sont largement utilisées dans la formation scientifique et dans la vie de tous

les jours (Poirier et Beaulac 2011). La validité d'une définition dépend de critères justificatifs différents en fonction du contexte.

Il importe de préciser que la justification fiabiliste s'applique fort probablement aussi à certaines connaissances de « haut niveau ». L'important est que celles-ci doivent remplir, en plus du critère fiabiliste, des critères supplémentaires. Nous disposons également d'un premier axe de différenciation : celui de la justification. Certaines connaissances doivent remplir des critères justificatifs différents et plus complexes que d'autres. Notre conception minimale n'est donc pas incompatible avec des formes de connaissances comme la connaissance scientifique.

Concernant ce premier axe de différenciation, il faut relever qu'il ne s'agit pas d'un critère empirique, mais conceptuel, ou théorique. En ce sens, il serait possible de développer un argument conceptuel en faveur de l'idée que la justification est hétérogène, sans avoir à recourir à un quelconque présupposé naturaliste. Par exemple, Pryor (2000) développe une position dogmatiste, qui correspond en gros à une version plus élaborée de l'argument anti-scepticisme de Moore (1939), selon lequel il est possible de savoir quelque chose que l'on ne peut prouver. Pryor argumente que nos expériences nous fournissent une forme de justification immédiate pour certaines croyances perceptuelles de base. Il appuie sa position sur une distinction entre ce type de justification et les formes plus complexes associées à des croyances qui ne sont pas des croyances perceptuelles de base. Comme ce type de justification est immédiat, le sceptique n'a pas l'occasion d'introduire une hypothèse comme celle du malin génie, parce que ce type d'hypothèse fonctionne en exploitant le fait que la justification toute croyance reposant sur l'expérience tire sa justification d'une croyance préalable selon laquelle notre expérience est réelle. Quant à savoir si ce type de justification immédiate est fiabiliste, la question reste ouverte (Pryor rejette cette option, mais ne fournit pas de raison claire pour ce rejet)⁴⁸. En tous les cas, notre premier axe de différenciation

⁴⁸ En fait, une version du fiabilisme de Goldman restreinte aux processus perceptuels fournirait exactement le type de justification dont parle Pryor. Nous n'avons pas la place de défendre cette idée

suggère une manière d'introduire l'idée de l'hétérogénéité de la connaissance qui ne repose pas sur l'adoption préalable d'une perspective naturaliste.

3.2. Psychologie évolutionniste : implications pour le critère de vérité; second axe de différenciation

La réception philosophique des travaux en psychologie évolutionniste est souvent peu enthousiaste, parce que l'adoption d'un cadre théorique darwinien est souvent accompagnée de révisions considérables des idées classiques d'un champ de recherche. Joyce (2006) développe, par exemple, un scepticisme moral basé sur une explication évolutionniste de l'existence de concepts moraux. En adoptant une position nativiste qui suppose l'innéisme de la tendance humaine à poser des jugements moraux (et par extension à avoir des croyances morales portant sur l'existence de faits et principes moraux orientant les jugements), il offre une explication généalogique des jugements moraux qui n'implique aucunement la véracité des croyances morales⁴⁹.

Les implications épistémologiques d'une branche particulière de la psychologie évolutionniste représentée entre autres par Tooby, Cosmides, Pinker et Buss (Faucher 2007) sont de la même espèce. Ceux-ci suggèrent en effet que les critères de l'épistémologie classique ne sont pas de bons candidats pour la description et l'explication de la cognition humaine.

ici, mais il s'agit d'une avenue intéressante à explorer. Pryor ne fournit pas d'explication claire de *pourquoi* nos expériences nous fournissent des croyances justifiées de manière immédiate, sinon un appel à l'intuition (Pryor 2000, p. 536) qui peut aisément s'avérer insatisfaisant. Une manière de remplir cette lacune est d'adopter la forme minimale du fiabilisme mentionnée ci-haut.

⁴⁹ La théorie de Joyce n'est ici citée que comme exemple de l'application d'un cadre darwinien à des questions philosophiques.

Rappelons la prémisse 3 de l'argument formulé à la fin du chapitre II : « afin d'expliquer nos capacités cognitives de manière satisfaisante, nous devons employer tant les approches phrastiques que les approches non phrastiques ». Cette prémisse repose sur la démonstration 1) du fait que nous sommes justifiés de considérer certaines structures informationnelles « de bas niveau » comme des connaissances (ce que nous avons fait dans la section 3.1) et 2) du fait qu'il existe bel et bien des instances de connaissance qui sont fondamentalement différentes et irréductibles les unes aux autres. Nous présentons ici un argument évolutionniste développé afin de soutenir les thèses de l'innéisme et de la modularité massive, dont l'implication épistémologique est que la « vérité »⁵⁰ n'est pas un critère pertinent pour l'analyse de certains types de connaissance. Cet argument souffre toutefois d'une erreur méthodologique importante, que nous devons corriger. Malgré tout, il demeure qu'il nous donne des raisons conceptuelles valables de remettre en question la pertinence du critère de vérité classique et de remplacer ce dernier par un critère pragmatique de vérité, compris comme le fait de réduire un certain type d'erreur selon le contexte. Nous présenterons ensuite les résultats obtenus par Gigerenzer (1999, 2009), qui tendent à supporter l'idée que la cognition piste non pas la vérité classique, mais la vérité pragmatique. Ceci nous fournit un second axe de différenciation et vient corroborer la prémisse 3.

3.2.1. Argument des mécanismes d'attribution de valeur

Tooby, Cosmides et Barrett (2005) tentent de surmonter la réticence à l'adoption d'un cadre darwinien en sciences cognitives et disciplines connexes, qu'ils voient comme un obstacle à l'étude de la cognition humaine. Ils choisissent de procéder par la démonstration du fait que la vision « table rase » (*blank slate*) de l'esprit humain supposant une architecture cognitive générale (*general-purpose*) et indépendante du contenu (*content-independent*), dont les capacités ne sont que le produit de l'apprentissage et de l'expérience et fonctionnant sur la

⁵⁰ Comprise comme la vérité correspondantiste.

base de règles logiques, est théoriquement improbable en plus d'être incompatible avec ce que nous savons de l'esprit humain (*ibid*, p.309)⁵¹. L'idée est que s'il est établi que l'esprit humain dispose de contenu inné (dans le cas présent, des mécanismes spécialisés [*domain-specific*]) issus de la sélection naturelle, il est nécessaire d'ajuster nos théories en conséquence et d'abandonner la vision « table rase ».

S'il peut être démontré qu'un organisme développe (ou a besoin de développer) des compétences reposant sur des éléments inaccessibles aux mécanismes sensoriels et absents de leur environnement, alors il est impossible pour une architecture générale, non spécialisée et indépendante du contenu, de les acquérir. Selon les auteurs, les compétences relatives à l'attribution de valeur correspondent à un tel cas parce que ces compétences assignent des valeurs à des états du monde, des éléments de l'environnement (stimuli) ou à des *outcomes* et la valeur elle-même n'est pas une propriété externe, mais relationnelle (*ibid*, p.314). Aucun stimulus n'impose de réponse identique chez tous les organismes qui interagissent avec : les propriétés du stimulus ne peuvent expliquer par elles-mêmes les variations comportementales des divers organismes. Il est nécessaire de construire une explication qui mobilise tant les propriétés du stimulus que les particularités (besoins, capacités, etc.) des organismes desquels nous cherchons à expliquer le comportement (*ibid*, p. 315). L'idée est donc que les mécanismes neurologiques responsables de l'attribution de valeur sont issus de la sélection naturelle et font partie de l'architecture neurocomputationnelle des animaux (humains et non-humains). Puisque la valeur est absente de l'environnement, elle doit être incluse dans cette architecture (*ibid*, p. 316). Conséquemment, les mécanismes employés dans l'attribution de valeur ont été sélectionnés parce qu'ils augmentent la valeur adaptative (*fitness*). Il s'agit d'un argument de pauvreté du stimulus standard.

Ces mécanismes ont deux propriétés importantes : 1) ils sont sensibles au contenu des représentations qu'ils traitent : ils doivent pouvoir différencier des situations en fonction de

⁵¹ Voir également Pinker (2002) pour une critique détaillée de la vision « table rase ».

leur contenu; une situation où l'animal est confronté à un serpent n'a pas le même contenu que lorsqu'il cherche de la nourriture et 2) ils sont générateurs de contenu : en attribuant une valeur à une situation, ils l'étiquettent comme étant dangereuse, neutre, etc. (*ibid*, p. 317). Les représentations ainsi générées sont employées par l'organisme pour guider son comportement comme une forme de connaissance (ce qui est cohérent avec le premier critère présenté en 3.2.1). Il est important de relever que ce type de connaissance peut évidemment être acquis, mais uniquement lorsqu'il y a interaction entre des processus d'apprentissage reposant sur l'expérience et les mécanismes d'attribution de valeur qui sont innés (*ibid*, p. 318). Les indices employés pour la discrimination entre membres d'un groupe social et non-membres peuvent changer par apprentissage et influence sociale (Kurzban, Tooby et Cosmides 2001)⁵².

Puisque les mécanismes d'attribution de valeur sont employés pour guider le comportement de manière à promouvoir une forme de succès biologique, et compte tenu du fait que la définition de ce qu'est un résultat favorable diffère d'un domaine à un autre, il s'ensuit qu'il ne peut y avoir de définition générale de « succès » s'appliquant à tous les domaines. Par exemple, employer la même définition de « succès » pour choisir un partenaire sexuel et pour choisir de la nourriture produirait des habitudes alimentaires pour le moins étranges. Dans cet exemple, « nourriture » et « reproduction » sont deux « domaines motivationnels » différents. Définis comme « *the set of represented inputs (...) that a functionally specialized set of evaluative procedures was designed by evolution to act over* » (Tooby, Cosmides et Barrett 2005, p.319), ces domaines sont individués sur la base du critère d'incommensurabilité motivationnelle : si les critères d'attribution de valeur pour deux domaines ne peuvent être dérivés du même ensemble fondamental (*core set*), alors il s'agit de deux domaines motivationnels distincts (*ibid*, p.318). Cet argument en faveur de la spécialisation achève, selon les auteurs, de réfuter la vision « table rase », qui suppose que l'architecture cognitive est en somme générique et flexible. En effet, un mécanisme

⁵² Mais le système qui emploie ces indices reste un système spécialisé. Il ne fait qu'« apprendre » à utiliser des indices différents, sa tâche reste la même.

générique est beaucoup moins efficace qu'un ensemble de procédures spécialisées pour résoudre des problèmes adaptatifs variés et incommensurables (thèse appuyée par certains résultats de la psychologie du raisonnement, voir section 3.2.4). Notons que les psychologues évolutionnistes affirment explicitement que l'esprit n'est pas uniquement modulaire, mais bien qu'il est *surtout* modulaire. Ils ne sont pas fermés à l'existence d'un processeur générique (Cosmides et Tooby 1994).

Puisque nous disposons maintenant d'une idée claire de l'argument des mécanismes d'attribution de valeur, nous pouvons nous tourner vers sa principale implication épistémologique, soit qu'il faut repenser la manière dont nous pensons l'acquisition des connaissances. Si l'attribution de valeur a pour fonction de guider le comportement et que les « représentations » (concepts ou véhicules informationnels) dont nous disposons sont acquis par le biais de mécanismes innés d'attribution de valeur ayant pour fonction de guider le comportement, alors la fonction de ces représentations (et des processus qui les emploient) est la régulation du comportement (*ibid*, p. 327). En conséquence, la fonction propre de la cognition n'est pas la fixation de « croyances vraies » (Fodor 2000), mais bien de favoriser la production d'actions biologiquement avantageuses (Tooby, Cosmides et Barrett 2005, p.327).

En effet, sous l'interprétation de Tooby, Cosmides et Barrett, le système nerveux est une consolidation de mécanismes de régulation du comportement dont la fonction est de générer des comportements adaptatifs (*ibid*, p.328). Puisque les procédures de régulation comportementale et d'acquisition d'information (connaissance) sont réalisées par le même type de mécanisme, il faut réviser notre conception de l'acquisition de connaissance. Il n'est plus question d'acquérir une connaissance par une inférence conforme aux règles de la logique (respectant, par exemple, les critères de vérité ou les propriétés de transitivité, etc.), mais bien par une inférence adaptative, qui a comme critère d'évaluation non pas la vérité mais le succès comportemental, ou la maximisation de la valeur adaptative. L'action correcte ne dépend pas du fait qu'elle se base sur une représentation véridique, mais bien sur une représentation qui promeut l'action biologiquement correcte.

Les psychologues évolutionnistes sont donc des pragmatistes. La cognition ne piste (*track*) pas la vérité, mais plutôt un ensemble d'indices (favorisant la survie) qui varient en fonction du contexte et de l'individu. Faucher (2007) partage cette interprétation :

1) ce que la cognition piste (*track*) peut varier en fonction de l'âge, du sexe, de la position hiérarchique, de l'état émotionnel et motivationnel de l'individu, etc. (les systèmes régulateurs d'action dont parlent Cosmides et Tooby tiennent compte de ces variables dans leur évaluation des états du monde); donc: 2) la cognition ne piste (*track*) pas la vérité. (Faucher 2007, p.377)

À noter qu'il manque une prémisse dans l'argument cité : chaque problème adaptatif demande une solution spécifique, ce qui explique pourquoi un critère général comme la vérité n'est pas approprié. De plus, il faut insister sur le fait que ce que les psychologues évolutionnistes soutiennent est qu'il n'existe pas de vérité métaphysique : il existe plusieurs vérités, une grande partie étant égocentrique et contextuelle⁵³.

3.2.2. Critique de l'argument

La conclusion à laquelle arrivent les psychologues évolutionnistes est donc la suivante : la cognition (nos mécanismes cognitifs) n'a pas été sélectionnée parce qu'elle piste (*track*) la vérité mais parce qu'elle maximise la valeur adaptative (*fitness*). Cette conclusion est cependant insatisfaisante. Le problème n'est pas qu'elle est inexacte ou inappropriée, mais plutôt qu'elle est inintéressante. Afin de mieux comprendre ce que nous voulons dire, il faut aborder l'argument du point de vue des arguments de « cible de sélection » (*'target of selection' argument*), qui font appel à la distinction sélection pour/sélection de développée par Sober (1984a, 1984b, voir également Goode et Griffiths 1995). Ce type d'argument est employé pour justifier une explication causale de la sélection d'un trait particulier. Par exemple, Shapiro (1992) l'emploie pour soutenir que le détecteur de « petites formes noires ressemblant à une mouche » de certaines grenouilles est bel et bien un détecteur de nourriture

⁵³ Cette idée rejoint la conception biosémantique de la représentation chez Millikan (1989).

et non de mouches : il y a eu sélection *pour* la détection de nourriture et seulement sélection *de* la détection de « petites formes noires ressemblant à une mouche ». En d'autres termes, le fait que ce mécanisme détecte des choses qui ressemblent à des mouches est un effet secondaire « parasitaire » d'un mécanisme de détection de nourriture.

Pour en revenir aux psychologues évolutionnistes, rappelons qu'ils ne sont pas fermés à l'existence d'un processeur générique. Celui-ci, vraisemblablement, fonctionnerait sur la base de règles logiques et dépendrait d'un critère de vérité fort (pensons par exemple au raisonnement abductif). L'argument évolutionniste peut donc être reformulé ainsi : il y a eu sélection *pour* la maximisation de la valeur adaptative (*fitness*) et sélection *du* pistage (*tracking*) de la vérité. Le fait que nous soyons capables de pister la vérité est un effet secondaire de la sélection de nos mécanismes cognitifs, sélectionnés pour le fait qu'ils augmentent la valeur adaptative (*fitness*). Il s'agit cependant d'une mauvaise utilisation de la distinction.

Goode et Griffiths (1995) argumentent que la distinction sélection pour/sélection de est bien utilisée lorsqu'elle est appliquée à des propriétés situées à un même niveau explicatif. Par exemple, la fourrure de l'ours polaire est sélectionnée pour ses propriétés thermiques, mais une fourrure chaude est également lourde. Il y a donc sélection de la lourdeur en même temps que des propriétés thermiques. Le critère suggéré par Sober permet de distinguer les propriétés qui sont les cibles de la sélection (*targets of selection*) de celles dont la sélection est une coïncidence par le fait que seule la propriété ciblée par la sélection joue un rôle causal dans le processus sélectif. Le rôle causal d'une propriété est déterminé par la vérité d'énoncés contrefactuels. Si la fourrure avait été chaude mais non lourde, elle aurait tout de même été sélectionnée. Si elle avait été lourde et non chaude, elle ne l'aurait pas été. Les propriétés thermiques sont donc les propriétés ciblées (Goode et Griffiths 1995, pp. 8-9).

La distinction peut également être mal utilisée, comme dans cette autre version de l'exemple de l'ours polaire. La fourrure de l'ours polaire est à la fois chaude et favorise la valeur adaptative (*fitness-enhancing*). Si la fourrure n'avait été que chaude sans favoriser la valeur adaptative, elle n'aurait pas été sélectionnée. Si elle avait favorisé la valeur adaptative

sans être chaude, elle aurait été sélectionnée. Il y a donc eu sélection pour une fourrure qui favorise la valeur adaptative, mais seulement sélection concomitante de ses propriétés thermiques (*ibid*, p. 9). Cette conclusion semble farfelue, et avec raison. Le problème ici est qu'une explication de plus haut niveau est présentée comme une alternative à une explication de plus bas niveau.

Il est évident que la propriété « augmente la valeur adaptative » est la cible du processus de sélection, mais l'opposer à une propriété de plus bas niveau (la chaleur) revient à opposer une propriété X au rôle qu'elle réalise. Considérons l'exemple suivant: « l'hémoglobine a-t-elle été sélectionnée pour le fait qu'elle retient les molécules d'oxygène ou parce qu'elle contient du fer? »⁵⁴. Dans ce cas, le fait que l'hémoglobine contienne du fer est ce qui fait en sorte qu'elle retient les molécules d'oxygène : la question posée tente de faire d'un rôle et de son réalisateur des compétiteurs causaux dans l'explication de la sélection de l'hémoglobine. De plus, dans le cas de l'ours polaire, le fait qu'une fourrure maximisant la valeur adaptative (par opposition à une fourrure chaude) aurait été sélectionnée implique qu'une autre propriété de plus bas niveau prendrait la place des propriétés thermiques de la fourrure (parce qu'il y a nécessairement une raison pour laquelle la fourrure augmente la valeur adaptative) (*ibidem*).

L'implication pour l'argument évolutionniste qui nous concerne est claire : le pistage (*tracking*) de la vérité est une propriété de plus bas niveau que la maximisation de la valeur adaptative. Conséquemment, l'argument devrait reposer sur le fait qu'une autre propriété, de même niveau que le pistage de la vérité, a été la cible de la sélection. Que pourrait être cette propriété? Le succès comportemental n'est pas un candidat satisfaisant : il pourrait être réalisé par le pistage de la vérité. Il est en fait difficile d'imaginer une alternative au pistage de la vérité. Il semble donc que l'hypothèse selon laquelle la cognition piste la vérité ne peut être écartée⁵⁵. Il ne s'agit toutefois pas d'un problème, si nous comprenons bien qu'un

⁵⁴ Cet exemple est emprunté à Griffiths (2011).

⁵⁵ On pourrait objecter que le terme « cognition » est trop général. Si la modularité massive est vraie, il n'y a pas eu sélection de « la cognition » mais d'un ensemble de modules spécialisés, qui ne pistent

système cognitif est un système biologique et qu'en conséquence l'optimisation de toute capacité de ce système est *contrainte*.

Griffiths propose une conception pragmatique de la vérité qui permet d'expliquer pourquoi une solution donnée est adaptative. Cette conception de la vérité est tout à fait compatible avec la modularité, parce qu'elle a besoin de la spécialisation des modules pour fournir une explication complète de leur sélection. Nous exposons cette réflexion dans ce qui suit.

3.2.3. Vérité pragmatique

Il nous sera utile de définir les deux formes de vérité qui seront à l'œuvre ici :

Une croyance B est pragmatiquement vraie ssi, dans un contexte C donné, elle réduit l'erreur comportementale en fonction de la valeur attribuée à ce type d'erreur.

Une croyance B est fortement vraie ssi il est le cas que B.

L'argument de Griffiths mène à l'élaboration de la vérité pragmatique, définie ci-haut.

Il est difficile de trouver une alternative intéressante au pistage de la vérité. Griffiths et Wilkins (2009) mentionnent, entre autre, la suggestion de Plantinga (2002) selon laquelle la plupart de nos croyances pourraient être fausses, mais que les organismes ont certains désirs qui, en tandem avec ces croyances fausses, produisent des comportements adaptatifs. Par exemple, Jim peut aimer l'idée de se faire manger, mais lorsqu'il rencontre un ours polaire, il s'éloigne en quête d'un meilleur prospect (soit un animal qui, lui, voudra le manger) parce qu'il pense qu'il est improbable que l'ours polaire le remarque et tente de le manger. Ceci a

pas la vérité mais les indices permettant un comportement adaptatif. La réponse à cette objection est simple : la vérité que piste la cognition (ou celle que pistent nos modules) n'est pas, en elle-même, suffisante pour déterminer l'action adaptative. L'argument de Griffiths vise simplement à fournir une manière d'expliquer *comment* nos mécanismes cognitifs spécialisés font en sorte de maximiser notre valeur adaptative.

pour effet d'éloigner son corps du danger (comportement adaptatif) sans d'aucune manière avoir recours à des croyances vraies⁵⁶.

Cette idée est pour le moins étrange, et Griffiths et Wilkins relèvent (avec raison) qu'elle requiert tout de même la présence d'un mécanisme de pistage de la vérité : la seule manière de faire fonctionner un tel système serait d'avoir « *some other cognitive mechanisms which tracked truth, and which adjusted the desires in the light of the actual goals of the organism so as to ensure that the false beliefs nevertheless produced adaptive behavior.* » (*ibid*, pp. 5-6). L'argument est donc le suivant : puisque même dans un cas où la cognition a avantage à produire des croyances fausses qui, de concert avec certains désirs, donnent lieu à des comportements adaptatifs, il est nécessaire d'avoir un mécanisme de pistage de la vérité, pourquoi ne pas simplement conserver l'idée que le pistage de la vérité a été la cible de la sélection?

Selon Griffiths et Wilkins, la raison pour laquelle il est vu comme improbable que la cognition ait été sélectionnée pour ses propriétés de pistage (*tracking*) de la vérité est qu'il existe une kyrielle d'observations empiriques incompatibles avec une architecture cognitive correspondant en somme à la vision traditionnelle (et idéalisée) de la rationalité. En d'autres termes, nous avons souvent une idée déraisonnable de ce que serait (ou devrait être) la cognition sélectionnée pour le pistage de la vérité et toute observation d'une déviation de cet optimum idéalisé est considéré comme preuve que la cognition n'a pas été sélectionnée pour le pistage de la vérité. Toutefois, si l'on considère le pistage de la vérité comme une simple propriété écologique du même type que la recherche de nourriture, ce problème disparaît. Tout phénotype sélectionné pour une fonction donnée ne peut être optimisé de manière absolue : il doit respecter (au moins) deux sortes de contraintes.

1) Les contraintes de coût .

⁵⁶ Cet exemple est adapté de Plantinga (1993), cité dans Griffiths et Wilkins (2009).

Une contrainte de coût contraint l'apport de ressources énergétiques allouées à une tâche donnée. Cette allocation est contrainte parce que l'organisme doit allouer de l'énergie à plusieurs processus. Ainsi, l'énergie est distribuée de manière à assurer les chances de survie de l'espèce. Par exemple, *Antechinus stuartii* (petit marsupial australien), lors de la saison reproductive, redirige toutes ses ressources énergétiques vers la production de spermatozoïdes. Il retire donc des ressources d'autres processus (système immunitaire, etc.) parce que ses chances de survivre jusqu'à la prochaine saison d'accouplement sont tout simplement trop faibles et il est donc plus profitable, pour la survie de l'espèce, de maximiser la production de spermatozoïdes, afin de maximiser le potentiel reproductif (Griffiths et Wilkins 2009, p.7). Dans le cas de la cognition, un système cognitif qui piste (*track*) la vérité ne devrait pas dépenser une calorie de plus en opérations cognitives qu'il n'est absolument nécessaire⁵⁷.

2) Les contraintes de tâche

Une réponse optimale à une tâche cognitive est toujours soumise à une exigence de *trade-off* entre des erreurs de type 1 et des erreurs de type 2⁵⁸. Il n'existe pas de solution sans erreur (Godfrey-Smith 1991). Un organisme est continuellement confronté à la nécessité de prendre des décisions en situation d'incertitude et, par exemple, un lapin ne peut (et n'a aucun intérêt à) atteindre une certitude cartésienne concernant la présence ou non de prédateurs dans son environnement immédiat. La réponse évolutivement optimale est celle qui, de manière générale, génère un taux prédictible d'équilibre entre les erreurs de type 1 et de type 2 (proportion prévisible d'erreurs de type 1 par rapport aux erreurs de type 2 et inversement). Ce qui permet de déterminer le nombre acceptable d'erreurs de type 1 ou 2 est l'attribution de valeur à la réduction d'un type d'erreur donné et le coût relatif de l'erreur dépend d'une kyrielle de facteurs spécifiques. Il n'existe donc pas de phénotype cognitif optimal générique. Les contraintes de tâche impliquent de maximiser la performance à une tâche au prix de réduire la performance à une ou plusieurs autres.

⁵⁷ Nous pouvons à titre d'exemple penser aux heuristiques : les heuristiques ne sont pas des méthodes servant à faire des erreurs mais plutôt à obtenir le plus de vérité qu'il vaut la peine d'avoir en fonction du coût qui y est associé (Griffiths et Wilkins 2009). Voir également Gigerenzer et Todd (1999). Notons également que la prochaine section discute plus en détail du lien entre la vérité telle que présentée dans cette section et la littérature sur les heuristiques.

⁵⁸ Faux positif et faux négatif, respectivement.

Donc, selon Griffiths et Wilkins, si la cognition a été sélectionnée *pour* le pistage (*tracking*) de la vérité, elle serait sujette à des contraintes de coût et ferait face à des contraintes inhérentes aux tâches qu'elle doit accomplir, et il s'agit de la seule définition plausible d'un système évolutivement optimal de pistage de la vérité (Griffiths et Wilkins 2009). Leur argument repose sur une conception naturalisée et affaiblie de la vérité, que nous appelons « vérité faible ». La vérité faible est, tout simplement, la réduction d'un certain type d'erreur en fonction de la valeur attribuée à ce type d'erreur dans un contexte C. Cette définition a deux avantages. Premièrement, elle est tout à fait cohérente avec l'exigence de spécialisation présentée en 3.2.1 parce qu'il ne s'agit pas d'un critère universel, mais de paramètres qui contraignent les solutions possibles aux problèmes adaptatifs que les organismes (incluant les humains) rencontrent. Deuxièmement, elle permet de formuler une version « cible de la sélection » satisfaisante de l'argument évolutionniste :

Argument 1 : sélection du pistage de la vérité pragmatique

1. Notre architecture cognitive piste (*track*) tant la vérité pragmatique, que la vérité forte. (Prémisse 1, provenant de l'observation)
2. Si notre architecture cognitive avait été sélectionnée pour le pistage de la vérité pragmatique sans sélection concomitante du pistage de la vérité forte, elle aurait maximisé notre valeur adaptative (*fitness*). (Prémisse 2, provenant de l'analyse des contrefactuelles)
3. Si notre architecture cognitive avait été sélectionnée pour le pistage de la vérité forte sans qu'il n'y ait sélection concomitante du pistage de la vérité pragmatique, elle n'aurait pas maximisé notre valeur adaptative. (Contrefactuel 2)

Conclusion : notre architecture cognitive a été sélectionnée pour le pistage de la vérité pragmatique.

On pourrait nous faire l'objection suivante : l'idée que la cognition a été sélectionnée pour le pistage de la vérité pragmatique est incohérente avec l'exigence de la spécialisation défendue par les psychologues évolutionnistes, parce qu'il s'agit d'un critère général, et qu'il n'y a pas

de solution générale aux problèmes adaptatifs. L'argument que nous venons de présenter nous permet de répondre à cette objection :

Argument 2 : sélection de modules spécialisés

1. La vérité pragmatique est un paramètre qui contraint les solutions possibles aux problèmes adaptatifs que les organismes (incluant les humains) rencontrent, et requiert, afin de produire des comportements adaptatifs, d'interagir avec des mécanismes spécialisés) (Prémisse 1)
2. Si une architecture cognitive est adaptative et piste la vérité pragmatique, alors elle doit être composée au moins en partie de mécanismes spécialisés (Conditionnelle 1)
3. Notre architecture cognitive est adaptative a été sélectionnée pour le pistage de la vérité pragmatique. (Prémisse 2, provenant de l'argument 1)

Conclusion : Notre architecture cognitive est (en partie) composée de mécanismes spécialisés.

Ce second argument montre la compatibilité de la notion de vérité pragmatique avec la perspective de la modularité. De plus, il anticipe l'objection selon laquelle « cognition » est un terme trop large : le pistage de la vérité pragmatique n'est pas incohérent avec l'idée de module spécialisés, il s'agit de deux notions complémentaires (voir 3.2.2). Ainsi reformulé, l'argument évolutionniste offre de bonnes raisons de croire que certaines formes de connaissances (comprises comme des représentations permettant de réguler le comportement et dont nous devons supposer l'existence pour formuler des explications comportementales) ne dépendent pas d'un critère fort de vérité, mais bien d'un critère de vérité pragmatique. Ceci ne signifie pas qu'il n'existe pas des connaissances qui ne dépendent pas d'un critère fort de vérité (voir chapitre II), mais la différence profonde entre les deux types de vérité est selon nous suffisante pour justifier le fait de parler de deux types de connaissance différents.

Griffiths et Wilkins (2009) font un lien entre cette conception de la vérité et le concept d'*Umwelt* développé (au sens où il est utilisé ici)⁵⁹ par Lorenz (1935). Il s'agit d'un autre argument qui, nous croyons, clarifie un peu plus ce qu'implique la notion de vérité faible. Dans le « monde de Jim » (le monde tel que Jim le perçoit), lorsque Jim regarde un oiseau précis dans un groupe de choucas des tours (un oiseau de la famille des corvidés), il voit un oiseau entouré d'autres choucas. Mais, pour l'oiseau, ces autres volatiles sont des compagnons de chasse, des partenaires sexuels, etc. Les deux manières de voir ces oiseaux sont « vraies », même si Jim et le choucas ne les voient pas de la même manière (Griffiths et Wilkins 2009, p. 10). De plus, certaines représentations employées par certains organismes et répondant aux critères de réduction d'erreur ne peuvent simplement pas être phrastiques. Une sangsue qui perçoit un mammifère par la chaleur qui s'en dégage n'a pas de représentation phrastique du mammifère, mais elle « sait » qu'il y a un mammifère dans les environs et cette représentation est « vraie » pour la sangsue. Ceci prémunit la conception faible de la vérité contre l'argument qu'adopter un critère tel que la vérité exclut de l'ensemble « organismes qui peuvent connaître » tout organisme n'employant pas de représentations phrastiques (Churchland 2007, voir également section 2.2).

Cet argument est principalement théorique. Nous présentons dans la prochaine section certains résultats empiriques qui tendent à soutenir nos conclusions.

3.2.4. L'apport de la psychologie du raisonnement

Nous devons, avant tout, faire un avertissement. Nous employons les résultats qui suivent pour suggérer que l'idée de la vérité pragmatique comme critère de vérité pertinent est plausible et cohérente avec *certaines* résultats de la psychologie du raisonnement. Nous ne prétendons pas faire une revue exhaustive de la littérature.

⁵⁹ Le concept a d'abord été introduit dans la théorie sémiotique de von Uexkull (1909).

Le travail de Kahneman, Slovic et Tversky (1982) a été très influent, et plusieurs disciplines, incluant l'économie et les sciences sociales (voir entre autres Elster 1983, 1989, ainsi que Thaler et Sunstein 2009) ont recours à la notion d'heuristique. Les heuristiques sont surtout considérées comme des raccourcis cognitifs employés pour réduire le temps et les ressources requises pour, par exemple, évaluer des probabilités. Prenons par exemple l'heuristique de représentativité. Lorsque nous devons juger de la probabilité d'un événement, nous avons accès à deux types d'information : l'information sur la fréquence des occurrences d'un événement et l'information spécifique concernant le cas en question. Le premier type concerne donc le taux de base de l'événement. En l'absence d'information spécifique, c'est cette information qui devrait être employée pour évaluer la probabilité d'un événement. Toutefois, lorsque les deux types d'information sont disponibles, les gens ont tendance à accorder plus d'importance à l'information spécifique (information concernant le cas particulier), négligeant le taux de base, ce qui peut causer une surestimation des probabilités (Tversky et Kahneman 1974). Cette vision des heuristiques comme des raccourcis biaisés causant des « échecs de rationalité » est relativement répandue⁶⁰.

Il existe des perspectives plus « optimistes »⁶¹. Dans la lignée des travaux de Simon sur la rationalité limitée (1956, 1990), Gigerenzer et ses collègues (Gigerenzer, Todd et ABC Research Group. 1999, Gigerenzer et Brighton 2009) étudient le fonctionnement précis des heuristiques, tout en cadrant ces recherches dans le contexte de l'amélioration des modèles

⁶⁰ Les travaux d'Elster (1983, 1989), ont contribué à alimenter la popularité de cette perspective en sciences sociales, parce qu'il se sert de la notion d'heuristique pour expliquer l'occurrence de mauvaises pratiques administratives ou de choix politiques irrationnels.

⁶¹ Nous ne voulons pas suggérer une quelconque dichotomie entre l'interprétation de Kahneman et celle de Gigerenzer. Nous voulons simplement attirer l'attention sur le fait que Gigerenzer insiste fortement sur l'apport bénéfique des heuristiques. D'ailleurs, dans son récent ouvrage *Thinking, Fast and Slow* (2011), Kahneman est très clair sur le fait qu'il vaut mieux avoir une vision nuancée des heuristiques.

descriptifs du raisonnement humain (ce qui inclut l'adoption d'une perspective adaptative). Dans leur article programmatique de 2009, Gigerenzer et Brighton présentent des résultats expérimentaux qui soutiennent deux de leurs thèses principales : 1) l'étude des heuristiques montre que moins de computation, d'information et de temps peut dans certains cas augmenter la précision et 2) les heuristiques doivent également être évaluées sur la base de leur rationalité écologique (Gigerenzer et Brighton 2009, p.107).

La première thèse repose sur la découverte d'effets *less-is-more*⁶². Ces effets constituent une seconde justification pour l'emploi des heuristiques. La première (et conventionnelle) justification est celle du *accuracy-effort trade-off* : les heuristiques sont moins coûteuse et demandent moins d'effort au prix d'une perte de précision (*ibid*, p. 111). Les effets *less-is-more* ont été découverts lors de l'évaluation de la performance d'heuristiques du type *take-the-best*. Cette classe d'heuristiques (aussi appelée la classe des heuristiques *one-good-reason*) fonctionne en ignorant la majeure partie de l'information disponible en choisissant, parmi l'ensemble des indices (variables) disponibles, le premier qui permet de faire une inférence (*ibid*, p. 113). Les indices sont ordonnés sans regard pour les dépendances entre indices mais plutôt en employant une mesure de la corrélation entre chaque indice et le critère pertinent dans le contexte (*ibidem*). Pour prendre un exemple simple, pensons simplement au fait que prédire le lever du soleil en se basant sur le rythme circadien d'un coq est une stratégie prédictive efficace. Elle repose sur un seul indice (le chant du coq) et on peut imaginer que lors du choix de cet indice comme prédicteur, les dépendances entre les indices ont été totalement ignorées au profit des corrélations entre les indices⁶³. L'indice (rythme circadien du coq) et le critère (prédiction satisfaisante du lever du soleil) sont fortement corrélés. Lorsque nous comparons les performances de différentes stratégies (régression

⁶² Comme il s'agit de recherches récentes, nous conservons la terminologie originale anglaise.

⁶³ Un système qui s'attarderait aux dépendances entre indices tenterait de déterminer l'existence d'un lien causal entre deux variables corrélées et, surtout, la direction de ce lien causal. À l'opposé, un système qui se concentre sur la détection de simples corrélations ne tient pas compte du lien de causalité et de sa direction.

multiple, *tallying*⁶⁴ et *take-the-best*), les résultats sont surprenants. Non seulement les heuristiques de type *take-the-best* performant mieux que les stratégies plus complexes et coûteuses lorsqu'il est question de formuler des prédictions, mais il semble que plus d'information et de computation peut en fait être néfaste à la justesse des prédictions.

⁶⁴ Le *tallying* consiste à ignorer la valeur (*weight*) des indices disponibles, de manière à réduire le temps de traitement et les ressources nécessaires.

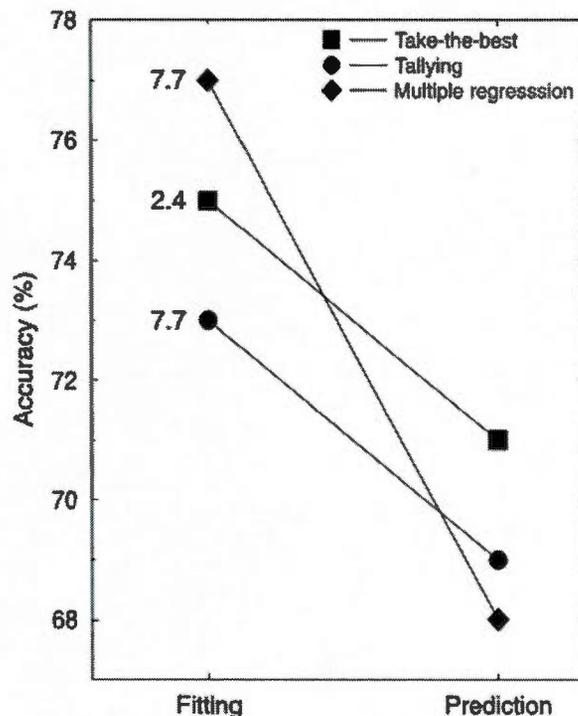


Figure. 3.1.⁶⁵ Effet *less is more*. Autant le *tallying* que *take-the-best* prédisent de manière plus précise que la stratégie de régression multiple malgré le fait qu'ils emploient moins d'information et de computation. La régression multiple, en revanche, est très performante en ce qui concerne le *fitting* (faire correspondre ses paramètres à des données connues) mais est comparativement peu performante pour la prédiction. On note également que *take-the-best* est considérablement plus frugale, n'utilisant en moyenne que 2.4 indices (contre 7.7. pour les deux autres). Ces résultats représentent 20 études incluant des tâches d'inférence psychologiques, biologiques, sociologiques et économiques (Czerlinski, Gigerenzer et Goldstein 1999).

Cet effet, appelé *less-is-more*, constitue donc une seconde justification pour l'emploi des heuristiques dans les environnements où celles-ci sont écologiquement rationnelles⁶⁶. Cette

⁶⁵ Figure tirée de Gigerenzer et Brighton (2009, p.112). Explication adaptée du même texte.

⁶⁶ Dans d'autres environnements ou circonstances, toutefois, employer des heuristiques, surtout des heuristiques très frugales du type *take-the-best*, peut avoir des conséquences (très) négatives.

justification a des conséquences normatives : elle va à l'encontre de la définition classique de la prise de décision rationnelle comprise comme le fait de prendre en considération toutes les informations disponibles et suggère que dans certains contextes des procédés biaisés sont plus efficaces (Gigerenzer et Brighton 2009, p. 111).

Ces résultats sont aussi cohérents avec l'idée qu'un organisme doit, pour réguler son comportement et se conduire adéquatement dans un environnement, prédire les états futurs de son environnement et doit donc, considérant les contraintes de coûts et les contraintes inhérentes à la tâche, simplifier l'information qu'il traite (simplifier la classe d'information pertinente permet de réduire la variance et améliorer les performances prédictives). Une manière efficace d'effectuer cette simplification est de tout simplement ignorer la majeure partie de l'information au profit d'un seul indice dont la corrélation (déterminée par l'environnement) avec le critère comportemental (déterminé par le contexte) est suffisamment fiable pour que la prédiction soit la plupart du temps correcte. Cette idée correspond, d'une part, aux inférences évolutivement valides des psychologues évolutionnistes (voir section 3.2.1) et à la conception faible de la vérité (voir section 3.2.3).

Cette dernière idée nous amène à considérer la rationalité écologique des heuristiques. Comme nous le savons, l'emploi d'heuristique est accompagné de l'émergence de biais. Sous l'interprétation classique de Kahneman et al., les biais sont des erreurs systématiques découlant des simplifications opérées par les heuristiques. Gigerenzer et Brighton (2009) suggèrent plutôt que les heuristiques sont mieux comprises du point de vue de la reconnaissance de motifs (pattern recognition) et de l'apprentissage automatique (*machine learning*), contextes où l'on trouve plusieurs exemples d'algorithmes inductifs biaisés pouvant réaliser des prédictions plus précises que des algorithmes non biaisés. Il n'est pas nécessaire d'entrer dans le détail de ces recherches (voir Geman, Bienenstock et Doursat 1992 pour une discussion détaillée de ces questions); l'important est de comprendre que ces résultats

peuvent être expliqués en décomposant l'erreur prédictive en la somme de trois composants (Gigerenzer et Brighton 2009, p.117)⁶⁷ :

$$\text{Erreur totale} = (\text{biais})^2 + \text{variance} + \text{bruit}^{68}$$

Le biais n'est donc qu'un composant de l'erreur. Ainsi, lorsqu'il est question de faire une prédiction précise (*accurate*), l'idéal serait de trouver un moyen de réduire au maximum toutes les sources d'erreur⁶⁹. Pour atteindre un faible taux d'erreur sur une large classe de problèmes, un modèle doit pouvoir accommoder une classe aussi riche de motifs (*patterns*) afin de minimiser le biais (Gigerenzer et Brighton 2009, p. 119). Une telle diversité dans la classe des motifs (*patterns*) qu'un modèle peut accommoder impose toutefois une augmentation de la variance parce que le modèle (à cause de sa grande flexibilité) accommode autant les motifs systématiques que les motifs accidentels. Si des motifs accidentels sont utilisés pour faire des prédictions, celles-ci risquent d'être inadéquates (*ibid*, p. 120). Nous sommes donc devant un dilemme : minimiser le biais requiert d'utiliser une large classe de modèles, mais combattre la variance requiert de restreindre cette classe de modèles. C'est pourquoi, selon Geman et al. (1992), les systèmes « *general-purpose* » (composés d'un algorithme complexe et flexible) ont tendance à être de mauvais prédicteurs lorsque l'information est rare (*sparse*).

Nos systèmes cognitifs sont confrontés à ce dilemme chaque fois qu'ils tentent de faire des inférences à propos du monde. Le succès des heuristiques de la classe *take-the-best* et le fait que les sciences cognitives tendent de plus en plus à confirmer l'idée que le système

⁶⁷ Formule tirée de Gigerenzer et Brighton (2009, p.117, notre traduction)

⁶⁸ Les termes « biais », « variance » et « bruit » réfèrent à trois sources d'erreur différentes. Le biais est au carré pour éliminer le négatif (le biais peut être positif ou négatif).

⁶⁹ Comme nous les verrons plus loin, les sources d'erreur (biais, variance, bruit) ont des effets différents. Notamment, le biais cause des erreurs de *fitting*, alors que la variance risque de causer des erreurs de prédiction.

cognitif humain fonctionne très bien lorsqu'il s'agit de généraliser à partir d'observations limitées (Oaksford et Chater 1998, Griffiths et Tenenbaum 2006) imposent déjà des contraintes importantes sur le type de modèle qui peut expliquer les performances cognitives humaines. Ces résultats, et la capacité du système cognitif de faire des prédictions précises (*accurate*) sur la base d'information limitée (*sparse*) suggèrent que la variance est minimisée, parce que si la variance était élevée, des motifs (*patterns*) accidentels seraient pris en considération, causant des performances prédictives faibles (Gigerenzer et Brighton 2009, p. 120). Pour expliquer ce contrôle de la variance, il faut abandonner l'idée d'un mécanisme inductif générique :

« Put simply, the bias–variance dilemma shows formally why a mind can be better off with an adaptive toolbox of biased, specialized heuristics. A single, general-purpose tool with many adjustable parameters is likely to be unstable and to incur greater prediction error as a result of high variance. » (Gigerenzer et Brighton 2009, p. 120)

Cette conclusion est cohérente avec la conception de la vérité pragmatique. Les recherches sur le fonctionnement des heuristiques suggèrent que notre système cognitif (et probablement celui des animaux non-humains) fonctionne sur la base d'heuristiques simples qui minimisent la variance afin de réduire le nombre d'erreurs prédictives. De plus, ces recherches appuient la thèse de spécialisation des psychologues évolutionnistes.

Avant de passer à la prochaine étape, il est approprié de soulever l'objection suivante : d'un point de vue écologique, les heuristiques fonctionnent bien lorsque l'environnement est coopératif. Mais les environnements naturels sont changeants et remplis de prédateurs pouvant exploiter ces biais et donc, en quoi un ensemble de mécanismes biaisés est-il adaptatif? La réponse est simple et comporte deux parties. Premièrement, les écosystèmes au sein desquels vivent les animaux non-humains sont stables, et les biais des mécanismes de traitement d'information prennent en compte la présence de prédateurs. Si un prédateur est en mesure d'exploiter un biais prédictif donné, il est raisonnable de penser qu'il y aura une

période d'ajustement au terme de laquelle le prédateur devra trouver un moyen d'exploiter un autre biais parce que ses proies changent de stratégie. Deuxièmement, les environnements sont certes changeants, et c'est justement là que réside la force des heuristiques. En ne tenant compte que d'une petite partie de l'information, les systèmes biaisés (il est raisonnable de le croire) choisissent les indices en fonction de leur récurrence, relativement à d'autres variables trop changeantes. Un système général et flexible devrait tenir compte de tous les paramètres variables d'un environnement. Un système spécialisé et biaisé ne tient compte que d'une partie de cette information.

3.2.5. Analyse des implications : comment établir la pertinence du second axe de différenciation

Rappelons la thèse centrale de cette section : la psychologie évolutionniste nous donne de bonnes raisons de croire que certaines formes de véhicules informationnels ne dépendent pas du fait d'être « vrais » au sens fort pour remplir leur fonction de régulation comportementale. Ces structures remplissent les conditions énoncées en 1.1 et sont en conséquence des candidates appropriées à la catégorisation comme « connaissance ». Suivant l'argument de Tooby, Cosmides et Barrett (2005), ces formes de connaissances ne respectent pas les principes classiques de la rationalité humaine. L'argument original était toutefois fragile, aussi nous l'avons critiqué et raffiné (sections 3.2.2 et 3.2.3) afin de montrer qu'un engagement au naturalisme dans l'explication du phénomène de la connaissance requiert l'adoption d'une forme faible et évolutivement cohérente de la vérité.

Nous avons ensuite montré que les recherches récentes sur les heuristiques motivent 1) de considérer les heuristiques comme des composantes adaptatives de notre système cognitif et 2) que la conception pragmatique de la vérité que nous adoptons est non seulement cohérente avec les recherches récentes en psychologie du raisonnement, mais en plus s'y intègre de manière élégante. Plus encore, les recherches de Gigerenzer et ses collègues nous permettent de préciser la conception de la vérité pragmatique : ce qui est important (dans un certain contexte) est de faire des prédictions fiables. Conséquemment, une architecture

cognitive employant des heuristiques biaisées (ce qui a pour effet de réduire la variance) maximise le succès prédictif, au coût d'un plus grand nombre d'erreurs de *fitting*. Comme ce qui constitue un succès prédictif varie en fonction du contexte et du domaine motivationnel, nous satisfaisons sans problème au critère de spécialisation de la psychologie évolutionniste et, parce que les heuristiques sont peu coûteuses et rapides, notre perspective est cohérente avec les contraintes inhérentes à un système cognitif qui piste (*track*) la vérité pragmatique.

Nous disposons donc d'un second axe de différenciation, ou d'hétérogénéité : certaines formes de connaissance dépendent d'un critère pragmatique de « vérité » alors que d'autres doivent satisfaire à un critère plus fort. Une connaissance employée par un ours polaire pour décider où aller chasser ne dépend pas des mêmes critères qu'un énoncé en langue naturelle affirmant l'interprétation correcte d'un passage philosophique obscur ou qu'un énoncé exprimant une explication scientifique. Dans le premier cas, la connaissance peut être une représentation non phrastique et sa « vérité » dépend du fait qu'elle est produite par un mécanisme favorisant le succès prédictif. Les autres cas doivent satisfaire des critères plus forts. De même, du point de vue de la justification, le premier cas ne requiert rien de plus que la fiabilité générale des mécanismes qui l'ont produit. Les autres cas doivent satisfaire des critères justificatifs plus complexes. Nous ne nous engageons donc pas à soutenir une forme de fiabilisme pour l'ensemble de la connaissance. Certaines formes de connaissances obtiennent leur justification de cette manière, d'autres non.

Mais les connaissances évolutives peuvent (dans certains cas seulement; il serait absurde de considérer les représentations de la sangsue comme des candidates d'implémentation phrastique⁷⁰) être représentées de manière phrastique à cause d'un gain explicatif. L'on pourrait se demander pourquoi ne pas simplement adopter les formulations phrastiques,

⁷⁰ Ces représentations pourraient être formulées de manière phrastique, mais le gain explicatif qu'une telle manœuvre occasionnerait n'est pas clair.

puisque nous n'avons besoin de rien d'autre. Le fait est qu'il existe des formes de connaissance qui ne peuvent jamais être décrites de manière phrastique.

3.3. Implications de la neuropsychologie de la vision : pertinence des caractérisations phrastiques

À ce stade, nous disposons de deux axes de différenciation des types de connaissance : les critères justificatifs et les critères de vérité. L'objectif de la présente section est d'identifier un troisième axe : la forme que prend la connaissance.

Les recherches de Goodale et ses collègues sur la patiente D.F. suggèrent l'existence de deux systèmes visuels, l'un responsable de la perception visuelle et l'autre du guidage visuel de l'action. Nous argumentons que cette dissociation fonctionnelle (qui correspond à une dissociation physiologique) suggère l'existence d'une forme de connaissance qui, fort probablement, n'est pas formulée de manière phrastique à l'intérieur du système de manière phrastique (même si elle pourrait être décrite de manière phrastique). L'analyse de ces résultats nous permettra également d'introduire une distinction entre les connaissances évolutives et ce que nous appellerons les connaissances motrices (ou visuomotrices).

3.3.1. Le cas D.F. : considérations neuroanatomiques et description

La patiente D.F., atteinte d'agnosie aperceptive, a fourni un cas des plus intéressants en neuropsychologie. Il faut d'abord préciser qu'un diagnostic d'agnosie aperceptive est rare. Les patients atteints de cette pathologie reçoivent souvent un diagnostic initial de cécité corticale⁷¹ (ou sont simplement considérés aveugles) à cause de leur apparente incapacité à

⁷¹ Perturbation générale de la perception visuelle résultant de lésions bilatérales au cortex occipital (Bérubé 1991, p.171)

naviguer dans l'environnement visuel (Bauer et Rubens 1985, p.191). Lorsqu'ils rapportent qu'ils sont en fait capables de voir, l'on administre une série de tests afin de déterminer la nature exacte de la pathologie. L'agnosie aperceptive est caractérisée par une variété de présentations différentes : les patients en souffrant peuvent exhiber des performances différentes selon la tâche et la lésion précise causant leur handicap. De manière générale, les patients atteints d'agnosie aperceptive échouent des tâches de reconnaissance à cause d'une déficience de perception. Ces patients ne peuvent coupler des objets à un échantillon, ou désigner du doigt un objet nommé par un examinateur. Ces problèmes impliquent généralement les éléments du champ visuel qui requièrent une reconnaissance de la forme ou de motifs.

En 1992, Goodale et Milner proposent le modèle des deux systèmes visuels, qui dissocie la vision-pour-l'action de la vision-pour-la-perception. Ils supposent, en s'appuyant sur des recherches en neuroanatomie, neurophysiologie et neuropsychologie, que cette distinction est implémentée dans les deux voies projetant du cortex visuel primaire (la voie dorsale, projetant vers le cortex pariétal postérieur et la voie ventrale, projetant vers le cortex inférotemporal).

Cette idée de deux voies de la perception visuelle vient de Ungerleider et Mishkin (voir Mishkin, Ungerleider et Macko 1983), qui proposent que la voie dorsale est responsable de la perception spatiale (voie du « où »), et que la voie ventrale est responsable de l'identification des objets (voie du « quoi »). Cette hypothèse est avancée sur la base de l'observation d'une dissociation des déficits visuels associés à des lésions localisées dans le cortex inférotemporal et des lésions localisées dans le cortex pariétal. Dans les études initiales de Pohl (1973) sur des singes, les lésions inférotemporales causent des déficiences sévères dans une tâche d'apprentissage de discrimination visuelle d'objets dans laquelle après avoir été familiarisé à un objet d'une paire, ils doivent choisir l'objet inconnu de la paire pour avoir une récompense. À l'inverse, les lésions pariétales postérieures ne causent pas de déficience de ce type, mais causent une déficience dans les tâches où le singe doit apprendre à choisir le morceau de nourriture le plus proche d'un marqueur géographique placé aléatoirement. Ces résultats suggèrent que le cortex inférotemporal participe surtout aux processus de perception

et de reconnaissance d'un objet sur la base de ses caractéristiques, et non de sa position dans l'espace. Inversement, le cortex pariétal postérieur semble surtout impliqué dans la perception des relations spatiales entre objets, et non des caractéristiques intrinsèques de l'objet (forme, couleur, texture, etc.) (Mishkin, Ungerleider et Macko 1983).

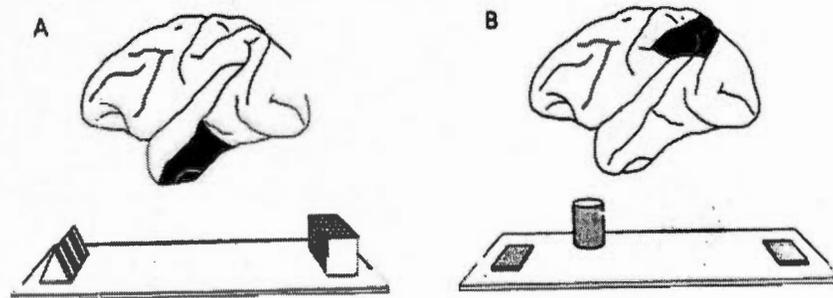


Figure. 3.2⁷². Tâches comportementales sensibles aux lésions corticales visuelles chez les singes. **A**: Discrimination d'objet. **B**: Discrimination de marqueur spatial (*landmark*). Les zones obscurcies représentent les zones touchées par les lésions (A : cortex inférotemporal; B : cortex pariétal postérieur).

Les conclusions de Pohl (1973) ont été corroborées par un certain nombre d'expériences subséquentes et de réplication de ses résultats (voir Mishkin, Ungerleider et Macko 1983 pour une discussion plus complète de ces recherches, incluant des études cliniques sur des patients humains). La conclusion générale, qui est ce qui nous intéresse, est que selon Mishkin et ses collègues il semble y avoir une forme de division du travail de la perception visuelle. La voie dorsale est responsable de la perception spatiale (relations spatiales entre objets) et la voie ventrale de la perception et de la reconnaissance des objets (identification des objets sur la base de leurs propriétés intrinsèques).

⁷² Figure et explication (adaptée) tirées de Mishkin, Ungerleider et Macko (1983).

Pour Goodale et Milner, cette différenciation n'est pas un simple cas de « division du travail », mais indique la présence de deux systèmes visuels indépendants (Westwood et Goodale 2011). Il s'agit donc d'une dissociation anatomique et fonctionnelle. Leur hypothèse se fonde entre autres sur l'étude d'une patiente (D.F.) atteinte d'agnosie aperceptive sévère résultant d'une lésion corticale localisée dans la voie ventrale (*ibidem*).

Lors d'une tâche expérimentale, la patiente doit placer des formes de bois dans des fentes orientées de diverses manières (Goodale, Milner, Jakobson et al. 1991, Goodale, Jakobson, Milner et al. 1994). Les résultats montrent qu'elle ajuste la position de sa main en fonction de l'orientation de la fente dès les premiers instants du mouvement – avant d'être en mesure de corriger son mouvement en employant de l'information visuelle ou tactile. Cette observation est fort surprenante, car D.F. est incapable de rapporter verbalement, ni manuellement (en orientant sa main et la carte de bois en fonction de l'orientation de la fente sans compléter la séquence) l'orientation de la fente. Ces résultats suggèrent que D.F. n'est pas en mesure de percevoir l'orientation de la fente de manière consciente (Westwood et Goodale 2011). Des manipulations subséquentes, visant à mesurer les aptitudes visuomotrices de D.F., révèlent que l'ouverture de sa main, lors d'un mouvement de préhension, reflète la largeur de l'objet qu'elle doit saisir (Goodale, Milner, Jakobson et al. 1991). Pourtant, elle est incapable de donner une estimation de la largeur de l'objet, dans ce cas-ci un bloc, avec son index et son pouce.

De plus, Goodale et ses collègues observent que D.F. choisit des postures de préhension appropriées lorsqu'elle doit prendre des objets de forme irrégulière. Ceci signifie en fait que l'axe d'opposition formé par le vecteur entre le point de contact de son pouce et celui de son index tend à passer par le centre de gravité de l'objet, sans quoi l'objet risque de pivoter ou d'être échappé lors de la manipulation (Goodale, Milner, Jakobson et al. 1994, Westwood et Goodale 2011). Il est aussi intéressant de constater que les points de contact initiaux de D.F. tendent à correspondre à des points de convexité maximale, ce qui diminue le risque de voir l'objet glisser de la main (Goodale et al. 1994, Westwood et Goodale 2011). Évidemment, les capacités de préhension de D.F. ne sont pas complètement préservées, mais les limites de ces

capacités ne sont pas incompatibles avec l'hypothèse des deux systèmes (voir Westwood et Goodale 2011 pour plus de détails).

Ganel et Goodale (2003) ont également démontré, chez des sujets « normaux », que la vitesse de la production d'un jugement perceptuel de la forme d'un objet est ralentie par le fait de devoir éliminer des aspects superflus (la hauteur de l'objet si la question porte sur sa largeur, par exemple), alors qu'aucun ralentissement de cette sorte n'est pas observé lorsqu'il s'agit de prendre l'objet d'une certaine manière. Ces résultats sont corroborés par ceux obtenus par Ganel, Goodale et Algom (2008). La perception visuelle des caractéristiques d'un objet serait donc guidée par des principes psychophysiques différents que le contrôle visuel de la préhension, ce qui renforce la thèse de Goodale et Milner. Notons que récemment, Goodale et Milner (sous presse) ont concédé à certains critiques que la dissociation n'est peut-être pas aussi complète qu'ils ne l'avaient supposé, mais ils maintiennent que les données dont ils disposent justifient leur interprétation.

3.3.2. Analyse des implications : comment établir le troisième axe de différenciation

L'implication principale de ces recherches est que l'idée selon laquelle la vision est un processus unifié semble être fautive (Jacob et Jeannerod 2003). Les informations visuelles que nous employons pour calibrer nos mouvements de préhension ne semblent pas être « dérivées » de la perception visuelle consciente. Si c'était le cas, il serait impossible pour D.F. de prendre un objet irrégulier sans l'échapper puisqu'il lui est impossible de percevoir consciemment la forme de l'objet. Ceci suggère qu'il existe une forme de « connaissance » visuomotrice, car le fait d'ajuster l'ouverture de la main et la position des doigts de manière appropriée à la forme de l'objet-cible requiert de l'agent qu'il en « connaisse » la forme.

Cette interprétation rapproche le type de connaissance employé par le guidage visuel de l'action de la notion de « savoir-faire » (know-how). Ce type de savoir correspond en gros à des capacités motrices. Churchland (2007) a recours à plusieurs exemples de sens commun pour avancer l'idée que la connaissance inclut également des structures informationnelles non phrastiques. Il y a par exemple de nombreuses instances de savoir-faire observées chez des

enfants en bas âge (ramper, par exemple) et même chez des adultes (marcher, attraper une balle au vol, etc.). Pour Churchland, le nouveau-né apprend à ramper hors de tout processus d'apprentissage langagier, et n'emploie pas de connaissance propositionnelle pour guider son comportement. De la même manière, les actions d'un joueur de baseball ne sont pas précédées de croyances occurrentes (*occurrent belief*), il attrape la balle de manière automatique, en se fiant à des indices fiables (bref en employant une heuristique)⁷³.

L'idée d'une différence fondamentale entre le savoir-faire et le savoir-que (connaissance propositionnelle) n'est pas nouvelle (cf. Ryle 1949, 1971). Son objectif était de s'en prendre à une perspective intellectualiste sur l'explication du comportement, et ce en montrant qu'il existe une classes de connaissances non-phrastiques (cf. Stanley et Williamson 2001)⁷⁴. L'implication que nous tirons des recherches de Goodale et ses collègues est, en quelque sorte, proche de l'idée de Ryle : il existe une forme de connaissance qui, chez l'être humain, est surtout mobilisée dans le guidage moteur de l'action, qui ne peut être ramenée à une conception phrastique de la connaissance. En effet, la vision-pour-l'action semble employer des structures informationnelles qui entretiennent un lien au monde (critère 1), guident l'action d'un organisme (critère 2) et, pour l'observateur, remplissent un rôle dans une explication scientifique plausible (critère 3). L'information employée pour la préhension n'est pas accessible consciemment et n'est pas une version simplifiée de l'information visuelle consciente, parce qu'elle est vraisemblablement employée par un système différent.

Nous disposons donc d'un second axe de différenciation (ou d'hétérogénéité) : la nature phrastique ou non. Nous l'avons déjà rencontré dans la section 3.2 (les représentations de la

⁷³ Gigerenzer et Brighton (2009) mentionnent d'ailleurs l'heuristique de fixation visuelle (*gaze heuristic*), qui consiste à simplement fixer le projectile à attraper et à ajuster sa vitesse de course en conséquence plutôt que calculer la vitesse et la trajectoire du projectile.

⁷⁴ Pour une discussion de la position intellectualiste de Stanley et Williamson (2001) selon laquelle le savoir-faire n'est qu'une espèce de savoir-que, voir la réponse de Noë (2005) – que nous adoptons.

sangsue), mais il semble que les êtres humains emploient eux aussi ce type de représentation qui, en fonction des critères définis en 3.1, correspond à une forme de connaissance.

3.4. Synthèse

Le présent chapitre avait comme objectif d'apporter un soutien empirique à l'argument que nous avons énoncé au chapitre II. Nous commencerons donc par rappeler l'argument et montrer comment les différentes considérations développées ici remplissent cette tâche.

Nous présenterons ensuite une première ébauche d'une conception hétérogène de la connaissance. Il faut insister sur le fait que celle-ci est préliminaire, provisoire, non-exhaustive et spéculative. Il ne s'agit *pas* d'une proposition substantive autant que d'un *exemple* de ce qu'il serait possible d'accomplir avec une conception de ce type.

3.4.1. Argument de nécessité explicative, prise 2

Il est utile de rappeler l'argument que nous avons énoncé en 2.4.3:

Argument de nécessité explicative à l'endroit de la connaissance

1. Pour expliquer nos capacités épistémiques, nous devons postuler la connaissance. (Prémisse 1)
2. Les constructions théoriques mises de l'avant par les théories phrastiques et non phrastiques de la connaissance n'ont que très peu⁷⁵ de propriétés communes. (Prémisse 2, venant de l'analyse métathéorique)

⁷⁵ Elles n'ont en fait aucune propriété commune, puisque, dans un contexte épistémologique moniste, l'approche non phrastique de Churchland implique l'élimination de l'approche phrastique et de son appareil conceptuel.

3. Afin d'expliquer nos capacités cognitives de manière satisfaisante, nous devons employer les deux sortes de constructions théoriques (Prémisse 3, venant de l'analyse métathéorique et de l'exposition des raisons conceptuelles en 2.3.2)
4. Donc l'explication de nos capacités épistémiques requiert d'avoir recours à des constructions théoriques n'ayant que très peu de propriétés communes. (Conclusion 1)
5. Donc, il n'existe pas de nature unique de la connaissance (ou : la connaissance n'est pas une espèce naturelle). (Conclusion 2)

Les prémisses conceptuelles (1 et 2) ont été défendues dans le chapitre 2. Le présent chapitre avait comme premier objectif de corroborer la prémisse empirique (3). Pour ce faire, nous devions montrer trois choses. Premièrement, que des formes de connaissance non phrastiques et ne dépendant pas du critère de vérité classique de connaissance existent. L'argument évolutionniste de la section 3.2 suggère qu'il existe des structures informationnelles dépendant d'un critère de vérité *pragmatique* qui, en vertu des conditions minimales exposées en 1.1, sont des instances de connaissances. De plus, l'analyse et l'interprétation des recherches de Goodale et Milner suggèrent que certains véhicules informationnels remplissant nos conditions minimales ne sont pas phrastiques. Notons que les structures employées par la vision-pour-l'action, reposent sans doute en partie sur un critère de vérité pragmatique, mais comme nous l'avons mentionné, elles requièrent quelque chose de plus. Il est important de rappeler le caractère « égocentrique » de la vérité pragmatique et de la vision-pour-l'action. Dans ces deux cas, la valeur épistémique de l'état informationnel est fonction du rôle que joue cet état *pour* un organisme⁷⁶. Le type de connaissance que suggèrent les recherches de Goodale et Milner rejoint donc l'idée de connaissances évolutives reposant sur un critère de vérité égocentrique et contextuel. Nous ne faisons pas de suggestion substantive quant à ce qu'est cette « autre chose », mais comme il est nécessaire

⁷⁶ Une piste intéressante pour approfondir cette idée est celle de la perspective biosémantique de Millikan (1989). Nous n'avons pas la place de discuter de cette théorie extrêmement riche ici, mais l'accent mis sur les mécanismes qui « consomment » les représentations, dans la théorie de Millikan, rejoint l'idée que la vérité pragmatique est « égocentrique ».

non seulement de réduire l'erreur mais aussi de compléter des calculs complexes (pour localiser les points de convexité maximale, par exemple), il y a fort à parier que la réduction d'erreur n'est qu'une partie de l'histoire. En résumé : l'identification des trois axes de différenciation (justification, vérité, forme phrastique) nous permet d'affirmer qu'il existe des formes de connaissances ne dépendant pas des critères traditionnels.

Deuxièmement, nous devons montrer qu'il existe des formes de connaissances devant être étudiées au niveau neurologique (ou psychologique) alors que d'autres étaient mieux comprises au niveau phrastique. Les connaissances évolutives, parce qu'elles s'appliquent à des situations de niveau écologique, et non seulement de performance motrice (comme c'est le cas pour la connaissance utilisée pour la préhension), sont probablement mieux étudiées au niveau scientifique.. Le niveau phrastique reste pertinent lorsqu'il est question de formes de connaissances impliquées dans des processus rendus possibles par l'utilisation du langage (voir 2.4.2).

Troisièmement, nous devons montrer que ces différents types de connaissance ne peuvent être réduits les uns aux autres, et qu'un naturalisme conséquent doit accepter l'idée qu'il existe différents types de connaissance. Comme nous venons de le mentionner, les connaissances évolutives ne dépendent pas des mêmes critères que les connaissances motrices, même si aucune ne dépend d'un critère de vérité traditionnel. Les connaissances évolutives et motrices ne sont donc pas de simples « espèces » de connaissances phrastiques; elles doivent être étudiées séparément, à l'aide des méthodes pertinentes (psychologie et neurosciences, respectivement)⁷⁷.

Nous concluons, donc, que nous avons de bonnes raisons de croire que notre argument de nécessité explicative est valide. Ceci implique qu'il n'existe pas de nature unique de la connaissance, et que par conséquent il faut rechercher une solution différente que la solution

⁷⁷ Nous ne prenons pas position sur l'approche adéquate à l'étude de ces formes de connaissance, il ne s'agit que d'une suggestion.

moniste au problème de la nature de la connaissance. Dans ce qui suit, nous proposons une première ébauche d'une conception de la connaissance qui pourrait fournir une telle solution.

3.4.2. Hétérogénéité de la connaissance : une première suggestion

Rappelons nos trois axes de différenciation : justification, vérité et forme phrastique. Nous avons, au long de ce chapitre, identifié des formes de connaissances caractérisées par le fait qu'elles dépendent de critères de justification et de vérité variables, et qui sont parfois de forme phrastique, parfois non. À l'aide de ces axes de différenciation, il serait possible de classer les formes de connaissance en fonction de leurs propriétés et des critères desquels elles dépendent. Par exemple, étudier le comportement animal requiert souvent de postuler des connaissances évolutives. Celles-ci ont certaines propriétés et dépendent de critères particuliers. Étudier les comportements préhensiles demande de postuler des connaissances motrices, mieux comprises au niveau neurologique. Ainsi, chaque type de connaissance correspondrait à une approche théorique et méthodologique particulière.

Ce que nous suggérons, en somme, est qu'une conception hétérogène de la connaissance telle que celle que nous développons permettrait d'organiser une sorte de division du travail épistémologique. Chaque forme de connaissance est mieux abordée par le biais d'une méthode particulière. Il s'agit donc d'un pluralisme méthodologique. Toutefois, ce pluralisme est « coopératif ». Parce que nous voyons les formes de connaissance plus complexes comme étant caractérisée par une augmentation progressive du nombre et de la complexité des critères épistémiques pertinents (et parfois par le remplacement d'un critère par un autre, plus exigeant). Le tableau ci-après montre à quoi ressemblerait cette division du travail. Il faut insister sur le fait que ni la conception que nous suggérons ici, ni le tableau ne sont exhaustifs. Il s'agit d'une suggestion préliminaire et provisoire.

Tableau 3.1. Typologie provisoire des formes de connaissance.

Connaissance évolutive (inférence évolutivement rationnelle) Méthodes : psychologie évolutionniste, éthologie cognitive	Vérité : pragmatique Justification : fiabiliste Forme : non phrastique (sauf pour une certaine classe de cas)
Connaissance pratique (vision-pour-l'action) Méthodes : neurosciences, sciences cognitives	Vérité : pragmatique* (réduction d'erreur + une certaine forme de correspondance) Justification : fiabiliste Forme : toujours non phrastique
Connaissance de type « sens commun » (« mes clés sont tombées à cause de la gravité », « Jim a trois chats », définitions vernaculaires) Méthode : épistémologie traditionnelle, linguistique	Vérité : forte (correspondance) Justification : forme d'inférence à la meilleure explication Forme : phrastique
Connaissance disciplinaire (interprétation correcte d'un texte philosophique, définitions techniques) Méthodes : épistémologie sociale, philosophie des sciences, sociologie des sciences	Vérité : forte* (correspondance + accord avec les normes de la discipline) Justification : Dépend des normes de la discipline Forme : surtout phrastique (mais certaines instances sont exprimées sous forme mathématique, par exemple)
Connaissance scientifique (théorie scientifique, définitions scientifiques) Méthodes : épistémologie sociale, philosophie des sciences, sociologie des sciences	Vérité : forte ** (en fonction de la conception du succès scientifique employée) Justification : dépend de plusieurs facteurs (succès explicatif, indices bibliométriques, facteurs sociaux etc.) Forme : Phrastique* (peut être également exprimée par des représentations graphiques ou des expressions mathématiques, mais dépend de manière importante de l'échange social de forme phrastique)

Cette typologie montre également l'augmentation progressive de la complexité des formes de connaissance et les méthodes appropriées à l'étude de chacune d'elles. En effet, les formes de vérité marquées d'astérisques représentent des formes plus complexes, devant remplir certaines conditions supplémentaires. Rappelons que ce tableau ne sert qu'à illustrer le genre de travail que peut accomplir une conception hétérogène de la connaissance. Dans ce qui suit, nous allons conclure le chapitre et relever certains obstacles que pourrait rencontrer une telle conception.

3.5. Conclusion

3.5.1. Obstacles à surmonter

Les obstacles principaux au développement d'une conception hétérogène de la connaissance sont 1) le fait qu'il faut démontrer que les formes différentes de connaissances ne dépendent pas des mêmes critères (et sont donc irréductibles aux formes déjà acceptées) et 2) qu'il est justifié de les considérer comme des instances de connaissances.

Nous croyons que le premier obstacle est surmonté, premièrement, par les critères minimaux de la connaissance présentés en 3.1 et, deuxièmement, par la conception pragmatique de la vérité développée en 3.2. Les critères minimaux sont compatibles avec les définitions traditionnelles de la connaissance parce qu'ils sont déjà inclus dans celles-ci. Si Jim dit « je sais que le train arrive à 15h », Jim emploie une structure informationnelle afin de guider son comportement et un observateur qui voit Jim partir de chez lui à temps pour se rendre à la station de train à l'heure devra, pour expliquer son comportement, supposer que Jim sait quelque chose qui fait en sorte qu'il part à ce moment. La différence entre la connaissance de Jim et celle employée par un lapin pour déterminer s'il peut ou non aller chercher de la nourriture à un endroit donné (en se basant sur une heuristique fiable qui fait des prédictions correctes la plupart du temps) est simplement dans le nombre et la complexité des critères à satisfaire. Ainsi, les critères minimaux, qui sont suffisants pour la connaissance employée par le lapin, ne le sont pas pour le cas de Jim, mais l'ajout de critères n'est en aucun cas problématique pour la conception hétérogène développée ici.

La conception pragmatique de la vérité permet quant à elle de surmonter le second obstacle en donnant une certaine force descriptive et normative à une conception hétérogène de la connaissance. La principale critique adressée à l'épistémologie naturaliste par le camp analytique est que naturaliser l'épistémologie évacue la dimension normative de l'épistémologie (Clarke 2004). On peut aussi imaginer une critique portant sur l'aspect descriptif : pour décrire la connaissance, nous avons besoin d'un critère de vérité. Nous ne pouvons avoir de critère de vérité assez fort sans conserver la notion analytique de vérité, et

donc l'épistémologie naturalisée échoue au point de vue descriptif aussi. La critique descriptive est rapidement écartée : la vérité pragmatique est un critère assez fort pour permettre de discriminer entre les représentations qui, accidentellement, causent un comportement approprié et celles qui produisent de manière fiable des comportements appropriés.

La critique normative n'est pas vraiment plus difficile à écarter. Comme nous l'avons vu dans la section 3.4, les effets *less-is-more* ont une implication normative importante : le fait que les règles traditionnelles de l'inférence rationnelle produisent dans certains contextes (où l'erreur prédictive est à éviter) des résultats plus faibles que certaines classes d'heuristiques simples implique qu'il faut réviser ces normes du raisonnement. En accord avec le principe kantien « doit implique peut », continuer à imposer (du moins au niveau de la connaissance individuelle) les normes de rationalité traditionnelles revient à demander d'un système des choses qu'il lui est impossible d'accomplir, ce qui rend la force normative de ces prescriptions totalement nulle. On pourrait évidemment contrer en disant qu'il s'agit d'un idéal auquel nous devrions aspirer, mais il semble que même cet idéal est moins performant dans les contextes prédictifs, qui constituent une large part de nos tâches cognitives (Gigerenzer et Brighton 2009).

3.5.2. Résumé

Le premier objectif du présent chapitre était de fournir le support empirique nécessaire à l'argument de nécessité explicative énoncé au chapitre II. Plus précisément, il s'agissait de valider la prémisse empirique de l'argument (prémisse 3). Nous avons donc, dans un premier temps, développé une notion minimale de connaissance. Ceci nous a permis d'identifier un premier axe de différenciation des types de connaissance, c'est-à-dire les critères justificatifs. Nous avons ensuite développé un argument évolutionniste pour montrer qu'il existe des formes de connaissance qui ne dépendent pas d'un critère de vérité classique, mais bien d'un critère pragmatique. Nous avons ensuite analysé les recherches de Goodale et Milner (1992, 1994, 2008) sur leur patient DF afin de suggérer qu'il existe des formes de connaissances non

phrastiques. C'est ainsi que nous en sommes arrivés à identifier deux autres axes de différenciation : la vérité et la forme (phrastique ou non). Ces considérations empiriques valident la prémisse 3 de l'argument de nécessité explicative. En conséquent, nous avons atteint le but principal du mémoire, soit de défendre l'idée qu'il n'existe pas de nature unique de la connaissance et que le débat sur la nature de la connaissance doit être dépassé.

Face à ce constat, nous avons mobilisé les considérations conceptuelles et théoriques développées au long du mémoire pour suggérer une manière de dépasser le débat, c'est-à-dire d'adopter une conception hétérogène de la connaissance. Nous avons esquissé une première ébauche de cette conception, qui ne se veut ni définitive, ni exhaustive. Nous avons ensuite examiné les implications et les avantages de cette conception. Nous concluons le mémoire par une discussion des perspectives de recherche que nous ouvre une telle conception de la connaissance.

CONCLUSION

L'objectif du présent mémoire est métathéorique. Nous avons tenté de montrer que la forme traditionnelle de la solution au problème de la nature de la connaissance est à repenser. En effet, l'idée selon laquelle il existerait une nature unique de la connaissance semble problématique parce qu'elle mène à proposer des théories de la connaissance qui souffrent d'un problème d'insuffisance explicative profond.

Les approches phrastiques font face à plusieurs problèmes. Comme le suggèrent les arguments de Churchland (1979, 1992, 2007), ces approches sont *superficielles* : elles ne peuvent rendre compte que d'une partie du phénomène de la connaissance. De plus, si on accepte l'idée que l'explication de la connaissance doit passer par l'étude du substrat physique comme le suggère Patricia Churchland (1987, 1990), il semble que les réponses que donnent les approches phrastiques aux questions des véhicules et du contenu de la connaissance (phrases et contenu propositionnel, respectivement), sont inadéquates. Toutefois, l'approche éliminativiste des Churchland (principale alternative à l'approche phrastique) se heurte également à des problèmes d'insuffisance explicative. En particulier, les définitions, qui sont non seulement des connaissances mais également des structures informationnelles qui permettent d'acquérir le concept d'une chose avec laquelle nous n'avons jamais eu de contact direct, sont un élément qui joue vraisemblablement un rôle central mais qui échappe à l'éliminativisme. Ces considérations nous ont menés à la formulation d'un argument de nécessité explicative à l'endroit de la connaissance, énoncé à la fin du chapitre II et reposant sur des conditions contextuelles et méthodologiques tirées de l'analyse de deux approches hétérogènes en philosophie de la psychologie (chapitre I).

Cet argument est la composante centrale du présent mémoire et sa prémisse clé est la prémisse 3 : pour expliquer les capacités épistémiques des animaux humains et non-humains,

nous devons avoir recours aux entités théoriques postulées par les deux principales approches naturalistes. Le chapitre III fournit un support empirique à cette prémisse en montrant que diverses formes de connaissances peuvent être différenciées sur la base des critères qu'elles doivent remplir pour acquérir une valeur épistémique. Nous avons identifié trois axes de différenciation : la justification, la vérité, et la forme phrastique ou non phrastique. Certaines formes de connaissance non phrastique requièrent simplement d'être pragmatiquement ou contextuellement vraies et sont justifiées sur la base du fait qu'elles remplissent un critère fiabiliste. D'autres doivent remplir des critères de justification plus exigeants et un critère de vérité plus « fort » (correspondantiste). L'axe de la justification est tiré d'une argumentation principalement conceptuelle, quoique les considérations liées à la psychologie du raisonnement fournissent un certain support à cette composante de notre argumentation. L'axe « vérité » repose sur un argument évolutionniste étayé par des considérations tirées des recherches de Gigerenzer (2000, 2009), et l'axe phrastique-non phrastique est tiré des recherches de Goodale et Milner sur la préservation des capacités de préhension chez une patiente atteinte d'agnosie aperceptive. Ces éléments en main, nous établissons la validité de notre argument de nécessité explicative.

Les conclusions de l'argument sont les suivantes : « connaissance » réfère à des entités n'ayant que peu de choses en commun et donc « connaissance » ne réfère pas à une espèce naturelle. Ceci implique que le problème de la nature de la connaissance ne peut recevoir de réponse moniste : il faut concevoir la question d'une autre manière. À ce stade, nous avons atteint l'objectif fixé : il est clair que nous disposons de raisons valables de remettre en question la forme traditionnelle de la solution au problème de la nature de la connaissance. Les dernières pages du chapitre III proposent une forme préliminaire d'une conception hétérogène de la connaissance, qui constitue la forme alternative de la solution au problème de la connaissance que nous voulons, à terme, mettre de l'avant. Nous reviendrons sur cette conception plus loin.

Le principal mérite du travail que nous avons accompli ici est, nous croyons, que nous avons développé une plateforme à partir de laquelle il est possible d'aborder plusieurs

problèmes importants en épistémologie et en philosophie des sciences. Dans les pages du présent mémoire, nous avons travaillé à l'intérieur d'un cadre ouvertement naturaliste. L'épistémologie naturalisée, qu'elle soit articulée de manière coopérative (comme le fait par exemple Kornblith, 2002) ou de manière plus radicale (comme dans le cas du naturalisme de remplacement) repose sur deux idées fondamentales. Premièrement, le but de l'épistémologie n'est pas de trouver des fondements privilégiés de la connaissance mais bien d'expliquer causalement son émergence et deuxièmement, les résultats de la recherche scientifique peuvent et doivent être mis à contribution. Ces deux idées sont centrales à ce que nous avons fait ici et au développement d'une conception hétérogène de la connaissance.

Une telle conception, comme nous l'avons vu à la fin du chapitre III, reconnaît l'existence d'entités fondamentalement différentes que nous pouvons appeler « connaissance » en vertu du rôle explicatif qu'elles jouent dans divers contextes. Le résultat est une perspective « pluraliste » qui « assigne » à diverses approches et disciplines une forme de connaissance particulière. Chaque forme de connaissance est donc mieux comprise par le biais d'une approche particulière. Ce type d'approche n'est pas nouveau et ressemble en quelque sorte à un constat d'échec : jamais nous ne trouverons la nature de la connaissance, mieux vaut étudier différentes choses séparément. Il s'agit toutefois d'une mauvaise interprétation de ce que nous tentons de faire. Certes, ce que nous proposons implique une forme de pluralisme méthodologique. Le nier serait simplement malhonnête. Toutefois, nous pensons que les différentes formes de connaissances que nous avons identifiées, ou plutôt que les formes de connaissance qui seront identifiées (parce que la typologie proposée est provisoire) entretiennent des liens entre elles. Ainsi, il reste beaucoup de travail à faire.

Nous avons mentionné le fait que le travail accompli ici fournit une plateforme à partir de laquelle il est possible de développer des projets de recherche spécifiques. Un exemple, et une avenue que nous comptons explorer, est le fait qu'il faut encore justifier l'idée de regrouper les diverses structures informationnelles dont nous avons discuté sous l'étiquette de « connaissance ». Nous avons déjà fourni une première justification par le biais de

l'élaboration de notre conception minimale de la connaissance (chapitre III) mais cette conception est avant tout un outil de travail et n'a pas nécessairement de force substantive. Pour mieux la justifier, il faudra montrer que les différentes formes de connaissance que nous postulons entretiennent véritablement des liens importants et que comprendre la nature et le fonctionnement des formes plus fondamentales est nécessaire à comprendre les formes plus complexes. Pour ce faire, il faudra travailler, entre autre, sur la question de comment mobiliser des résultats expérimentaux dans le cadre de débats philosophiques, et surtout de quelle manière les structures informationnelles situées au niveau neurologique peuvent s'avérer pertinentes pour la recherche épistémologique.

Une autre avenue intéressante est celle suggérée au chapitre II, lorsqu'il était question des interactions entre les diverses formes de connaissances. Une conception hétérogène adéquate devra, selon nous, fournir des moyens de penser ces interactions. En ce sens, nous ne proposons pas un simple pluralisme méthodologique, ou un relativisme.

L'objectif à long terme est donc de développer une théorie générale de la connaissance, qui soit en mesure de rendre compte de l'ensemble de ce phénomène. Il s'agit d'un projet imposant, mais nous avons, dans le présent mémoire, fourni une assise que nous croyons solide.

BIBLIOGRAPHIE

- Armstrong, D. M. 1973. *Belief, Truth and Knowledge*: London, Cambridge University Press.
- Barsalou, L. W. 1999. « Perceptual Symbol Systems ». *Behavioral and Brain Sciences* 22(4): 577-660.
- Bauer, R. et Rubens, A. B. 1985. « Agnosia ». In *Clinical Neuropsychology*. K. M. Heilman et Valenstein, E. Oxford, Oxford University Press: 187-241.
- Bechtel, W. 2008. *Mental Mechanisms : Philosophical Perspectives on Cognitive Neuroscience*. New York: Routledge.
- Bechtel, W. et Richardson, R. C. 2010. *Discovering Complexity : Decomposition and Localization as Strategies in Scientific Research*. Cambridge, Mass.: MIT Press.
- Bérubé, L. 1991. *Terminologie De Neuropsychologie Et De Neurologie Du Comportement*. Montréal: Les Éditions de la Chenelière Inc.
- Boyd, R. 1989. « What Realism Implies and What It Does Not ». *Dialectica* 43(1-2): 5-29.
- Boyd, R. 1991. « Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds ». *Philosophical Studies* 61(1-2): 127-148.
- Chisholm, R. M. 1977. *Theory of Knowledge*. Englewood Cliffs, N.J.: Prentice-Hall.
- Churchland, P. M. 1979. *Scientific Realism and the Plasticity of Mind*. Cambridge ; New York: Cambridge University Press.
- Churchland, P. M. 1988. *Matter and Consciousness : A Contemporary Introduction to the Philosophy of Mind*. Cambridge, Mass.: MIT Press.
- Churchland, P. M. 1989. *A Neurocomputational Perspective : The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press.
- Churchland, P. M. 1992. « Activation Vectors Versus Propositional Attitudes: How the Brain Represents Reality ». *Philosophy and Phenomenological Research* 52(2): 419-424.

- Churchland, P. M. 2007. *Neurophilosophy at Work*. Cambridge; New York: Cambridge University Press.
- Churchland, P. S. 1980. « Language, Thought, and Information Processing ». *Noûs* 14(May): 147-170.
- Clarke, M. 2004. *Reconstructing Reason and Representation*. Cambridge, Mass.: The MIT Press.
- Collier, J. 1996. « On the Necessity of Natural Kinds ». In *Natural Kinds, Laws of Nature and Scientific Reasoning*. P. Riggs. Dordrecht, Kluwer: 1-10.
- Cosmides, L. et Tooby, J. 1994. « Origins of Domain Specificity: The Evolution of Functional Organization ». In *Mapping the Mind: Domain-Specificity in Cognition and Culture*. L. Hirshfield et Gelman, S. Cambridge, Cambridge University Press: 85-116.
- Czerlinski, J., Gigerenzer, G. et Goldstein, D. G. 1999. « How Good Are Simple Heuristics? ». In *Simple Heuristics That Make Us Smart*. G. Gigerenzer, Todd, P. M. et Group, A. R. New York, Oxford University Press: 97-118.
- Darwin, C. 1965. *The Expression of the Emotions in Man and Animals*. Chicago,: University of Chicago Press.
- Decety, J. et Houdé, O. 2003. « Information ». In *Vocabulaire De Sciences Cognitives*. O. Houdé. Paris, Presses Universitaires de France.
- Dennett, D. C. 1987. *The Intentional Stance*: MIT Press.
- Dretske, F. I. 1981. *Knowledge & the Flow of Information*. Cambridge, Mass.: MIT Press.
- Ekman, P. 1971. « Universals and Cultural Differences in Facial Expressions of Emotion ». In *Nebraska Symposium on Motivation*. J. K. Cole. Lincoln, University of Nebraska Press. 4.
- Ekman, P. et Friesen, W. V. 1971. « Constants across Cultures in the Face and Emotion ». *Journal of Personality and Social Psychology* 17(2): 124-129.
- Elster, J. 1983. *Sour Grapes : Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press.
- Elster, J. 1989. *Solomonic Judgements : Studies in the Limitations of Rationality*. Cambridge: Cambridge University Press.

- Evans, J. S. B. T. et Frankish, K. 2009. *In Two Minds : Dual Processes and Beyond*. Oxford ; New York: Oxford University Press.
- Faucher, L. 2007. « *Faraway, So Close?* ». *Philosophiques* 34(2): 375-388.
- Feldman, R. 2001. « Naturalized Epistemology ». In *Stanford Encyclopedia of Philosophy*.
- Fodor, J. A. 1974. « Special Sciences ». *Synthese* 28(2): 97-115.
- Fodor, J. A. 1975. *The Language of Thought*. New York: Crowell.
- Fodor, J. A. 2000. *The Mind Doesn't Work That Way : The Scope and Limits of Computational Psychology*. Cambridge, Mass.: MIT Press.
- Ganel, T. et Goodale, M. A. 2003. « Visual Control of Action but Not Perception Requires Analytical Processing of Object Shape ». *Nature* 426(6967): 664-667.
- Ganel, T., Goodale, M. A. et Algom, D. 2008. « Visual Coding for Action Violates Fundamental Psychological Principles ». *Current Biology* 18(599-601).
- Geman, S., Bienenstock, E. et Doursat, R. 1992. « Neural Networks and the Bias/Variance Dilemma ». *Neural Computation* 4: 1-58.
- Gendler, T. 2008a. « Alief in Action (and Reaction) ». *Mind and Language* 23(5): 552--585.
- Gendler, T. S. 2008b. « Alief and Belief ». *Journal of Philosophy* 105(10): 634-663.
- Gettier, E. 1963. « "Is Justified True Belief Knowledge?" ». *Analysis* 23(6): 121-123.
- Gigerenzer, G. et Brighton, H. 2009. « Homo Heuristicus: Why Biased Minds Make Better Inferences ». *Topics in Cognitive Science* 1: 107-143.
- Gigerenzer, G., Todd, P. M. et ABC Research Group. 1999. *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.
- Godfrey-Smith, P. 1991. « Signal, Decision, Action ». *Journal of Philosophy* 88: 709-722.
- Goldman, A. 1979. « What Is Justified Belief? ». In *Justification and Knowledge*. G. Pappas, Boston: D. Reidel: 1-25.
- Goldman, A. I. 1976. « Discrimination and Perceptual Knowledge ». *Journal of Philosophy*(73): 771-791.
- Goldman, A. I. 1986. *Epistemology and Cognition*. Cambridge, Mass.: Harvard University Press.

- Goodale, M. A., Jakobson, L. S., Milner, A. D., et al. 1994. « The Nature and Limits of Orientation and Pattern Processing Supporting Visuomotor Control in a Visual Form Agnostic ». *Journal of Cognitive Neuroscience* 6(1): 46-56.
- Goodale, M. A. et Milner, A. D. 1992. « Separate Visual Pathways for Perception and Action ». *Trends in Neurosciences* 15(1): 20-25.
- Goodale, M. A., Milner, A. D., Jakobson, L. S., et al. 1991. « A Neurological Dissociation between Perceiving Objects and Grasping Them ». *Nature* 349(6305): 154-156.
- Goodale, M. A., Milner, A. D., Jakobson, L. S., et al. 1994. « The Nature and Limits of Orientation and Pattern Processing Supporting Visuomotor Control in a Visual Form Agnostic ». *The Journal of Cognitive Neuroscience* 6(1): 46-56.
- Goode, R. et Griffiths, P. E. 1995. « The Misuse of Sober's Selection of/Selection for Distinction. ». *Biology and Philosophy* 10: 99-108.
- Griffiths, P. E. 1997. *What Emotions Really Are : The Problem of Psychological Categories*. Chicago, Ill.: University of Chicago Press.
- Griffiths, P. E. 2004. « Emotions as Natural Kinds and Normative Kinds ». *Philosophy of Science* 71((5 Supplement: Proceedings of the 2002 Biennial Meeting of the PSA)): 901-911.
- Griffiths, P. E. 2004. « Is Emotion a Natural Kind? ». In *Philosophers on Emotion*. R. C. Solomon. Oxford and New York, Oxford University Press: 233-249.
- Griffiths, P. E. et Scarantino, A. 2005. « Emotions in the Wild: The Situated Perspective on Emotion ». In *The Cambridge Handbook of Situated Cognition*. P. Robbins et Aydede, M., Cambridge University Press.
- Griffiths, P. E. et Wilkins, J. F. 2009. «When Do Evolutionary Explanations of Belief Debunk Belief? ». *Darwin in the 21st Century*, Notre-Dame University.
- Griffiths, T. L. et Tenenbaum, J. B. 2006. « Optimal Predictions in Everyday Cognition ». *Psychological Science* 17(9): 767-773.
- Hacking, I. 1991. « A Tradition of Natural Kinds ». *Philosophical Studies* 61(1): 109-126.
- Jacob, P. et Jeannerod, M. 2003. *Ways of Seeing : The Scope and Limits of Visual Cognition*. Oxford ; New York: Oxford University Press.
- Joyce, R. 2006. *The Evolution of Morality*. Cambridge, Mass.: MIT Press.
- Kahneman, D. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

- Kahneman, D., Slovic, P. et Tversky, A. 1982. *Judgment under Uncertainty : Heuristics and Biases*. Cambridge ; New York: Cambridge University Press.
- Keil, F. C. 1989. *Concepts, Kinds and Cognitive Development*. Cambridge, Mass.: MIT Press.
- Kornblith, H. 2002. *Knowledge and Its Place in Nature*. New York: Oxford University Press.
- Kripke, S. A. 1980. *Naming and Necessity*: Harvard University Press.
- Kurzban, R., Tooby, J. et Cosmides, L. 2001. « Can Race Be Erased? Coalitional Computation and Social Categorization ». *Proceedings of the National Academy of Sciences* 98: 15387- 15392
- Lorenz, K. 1935. « Der Kumpan in Der Umwelt Des Vogels ». *J. Ornithologie* 83: 137-213.
- Machery, E. 2005. « Concepts Are Not a Natural Kind ». *Philosophy of Science* 72: 444-467.
- Machery, E. 2009. *Doing without Concepts*. Oxford ; New York: Oxford University Press.
- Marr, D. 1982. *Vision*: Freeman.
- Millikan, R. G. 1989. « Biosemantics ». *Journal of Philosophy*, 86: 281-297
- Milner, A. D., Ganel, T. et Goodale, M. A. sous presse. « Does Grasping in Patient D.F. Depend on Vision? » *Trends in Cognitive Science*. Publication électronique 16 mars 2012, Consulté: 1er avril 2012.
<http://www.sciencedirect.com/science/article/pii/S1364661312000599>
- Milner, A. D. et Goodale, M. A. 2008. « Two Visual Systems Re-Viewed ». *Neuropsychologia* 46(3): 774-785.
- Mishkin, M., Ungerleider, L. G. et Macko, K. 1983. « Object Vision and Spatial Vision: Two Cortical Pathways ». *Trends in Neurosciences* 6: 414-417.
- Moore, G. E. 1939. « Proof of an External World ». *Proceedings of the British Academy*, 25: 273-300.
- Noë, A. 2005. « Against Intellectualism ». *Analysis* 65(288): 278-290.
- Oaksford, M. et Chater, N. 1998. *Rational Models of Cognition*. Oxford, Ang.: Oxford University Press.

- Osherson, D. N. et Smith, E. E. 1981. « On the Adequacy of Prototype Theory as a Theory of Concepts ». *Cognition* 9: 35-58.
- Pinker, S. 2002. *The Blank Slate : The Modern Denial of Human Nature*. New York: Viking.
- Plantinga, A. 1993. *Warrant and Proper Function*. New York: Oxford University Press.
- Plantinga, A. 2002. « The Evolutionary Argument against Naturalism ». In *Essays on Plantinga's Evolutionary Argument Against*. J. K. Beilby. Ithaca, NY., Cornell University Press: 1-13.
- Pohl, W. 1973. « Dissociation of Spatial Discrimination Deficits Following Frontal and Parietal Lesions in Monkeys ». *Journal of Comparative and Physiological Psychology* 82(227-239).
- Poirier, P. et Beaulac, G. 2011. « Le Véritable Retour Des Définitions ». *Dialogue* 50: 153-164.
- Prinz, J. J. 2002. *Furnishing the Mind: Concepts and Their Perceptual Basis*: MIT Press.
- Pryor, J. 2000. «The Skeptic and the Dogmatist ». *Noûs* 34-4: 517-549.
- Putnam, H. 1975. *Mind, Language, and Reality*: Cambridge University Press.
- Putnam, H. 1994. « Sense, Nonsense, and the Senses: An Inquiry into the Powers of the Human Mind ». *Journal of Philosophy* 91(9): 445-517.
- Quine, W. V. 1969. « Epistemology Naturalized ». In *Ontological Relativity and Other Essays*, New York: Columbia University Press.
- Quine, W. V. 2004. *Quintessence: Basic Readings from the Philosophy of W.V. Quine*: Belknap Press of Harvard University Press.
- Ryle, G. 1949. *The Concept of Mind*: Hutchinson and Co.
- Ryle, G. 1971. « Knowing How and Knowing That ». In *Collected Papers*. G. Ryle. New York, Barnes and Nobles. 2: 212-225.
- Scoville, W. B. et Milner, B. 1957. « Loss of Recent Memory after Bilateral Hippocampal Lesions ». *Journal of neurology, neurosurgery, and psychiatry* 20: 11-21.
- Shapiro, L. A. 1992. « Darwin and Disjunction: Optimal Foraging Theory and Univocal Assignments of Content ». *Proceedings of the Philosophy of Science Association* 1992 1: 469-480.

- Simon, H. A. 1956. « Rational Choice and the Structure of Environments ». *Psychological Review* 63: 129-138.
- Simon, H. A. 1990. « Invariants of Human Behavior ». *Annual Review of Psychology* 41: 1-19.
- Sober, E. 1984a. *The Nature of Selection : Evolutionary Theory in Philosophical Focus*. Cambridge, Mass.: MIT Press.
- Sober, E. 1984b. « Force & Disposition in Evolutionary Theory ». In *Minds, Machines & Evolution*. C. Hookway. Cambridge, Cambridge University Press.
- Stanley, J. et Williamson, T. 2001. « Knowing How ». *Journal of Philosophy* 98(8): 411-444.
- Steup, M. 2008. « The Analysis of Knowledge ». In *Stanford Encyclopedia of Philosophy*.
- Thaler, R. H. et Sunstein, C. R. 2009. *Nudge : Improving Decisions About Health, Wealth, and Happiness*. New York: Penguin Books.
- Tooby, J., Cosmides, L. et Barrett, H. C. 2005. « Resolving the Debate on Innate Ideas ». In *The Innate Mind: Structure and Contents*. Peter Carruthers, Laurence, S. et Stich, S. New York, Oxford University Press: 305-337.
- Tversky, A. et Kahneman, D. 1974. « Judgment under Uncertainty: Heuristics and Biases ». *Science* 185(4175): 1124-1131.
- Von Uexküll, J. 1909. *Umwelt und Innenwelt der Tiere*. Berlin: Springer
- Webb, D. M. et Zhang, J. 2005. « Foxp2 in Song-Learning Birds and Vocal-Learning Mammals ». *J Hered* 96(3): 212-216.
- Westwood, D. A. et Goodale, M. A. 2011. « Converging Evidence for Diverging Pathways: Neuropsychology and Psychophysics Tell the Same Story ». *Vision Research* 51: 804-811.
- Williamson, T. 2000. *Knowledge and Its Limits*: Oxford University Press.