

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

LE PROJET DE NATURALISATION DE L'INTENTIONNALITÉ PAR L'ÉNACTIVISME

MÉMOIRE  
PRÉSENTÉ  
COMME EXIGENCE PARTIELLE  
À LA MAÎTRISE EN PHILOSOPHIE

PAR  
ANTOINE MILETTE-GAGNON

FÉVRIER 2024

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.12-2023). 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.»

## REMERCIEMENTS

Je tiens tout d'abord à remercier mon directeur de recherche, Pr. Pierre Poirier. Son attention sincère et son érudition, mais surtout son travail exemplaire comme directeur et mentor m'ont été précieux durant la rédaction de ce mémoire. Sans son soutien, il n'est pas garanti que j'aurais trouvé le moyen de me rendre jusqu'au bout de cette aventure.

Je tiens également à remercier Pr. Luc Faucher et Pr. Denis Fisette. Leurs commentaires m'ont permis de réorienter quelque peu la direction de ce mémoire, qui je crois s'en trouve grandement amélioré.

Je tiens à remercier ma famille, notamment mon frère Alexandre, mon père Benoit, ma grand-mère Andrée, ma tante Danièle et mon cousin Frédéric, qui ont su m'encourager dans cette épreuve qu'est la rédaction d'un mémoire de maîtrise.

Enfin, je tiens à remercier le Fonds de recherche du Québec pour son soutien généreux de mes recherches philosophiques.

## TABLE DES MATIÈRES

REMERCIEMENTS.....	ii
LISTE DES FIGURES .....	v
RÉSUMÉ .....	vi
ABSTRACT.....	vii
INTRODUCTION .....	1
CHAPITRE I TROIS TYPES D'INTENTIONNALITÉ.....	11
1.1. La conception « classique » de l'intentionnalité.....	11
1.2. L'intentionnalité représentationnelle .....	13
1.3. L'intentionnalité pragmatique.....	16
1.3.1. Une intentionnalité sans représentation .....	17
1.3.2. L'intentionnalité pragmatique chez Merleau-Ponty et Gallagher.....	18
1.4. L'intentionnalité biologique.....	20
1.4.1. Les germes d'une intentionnalité biologique : Merleau-Ponty, Schrödinger et Jonas... 21	
1.4.2. L'intentionnalité biologique comme directionnalité fondamentale du vivant.....	24
1.5. Conclusion: vers une redéfinition du problème de la naturalisation de l'intentionnalité .....	26
CHAPITRE II LA NATURALISATION DE L'INTENTIONNALITÉ.....	28
2.1. Tentatives passées de la naturalisation de l'intentionnalité représentationnelle : les théories téléosémantiques .....	29
2.1.1. Dretske: de l'information à la téléosémantique .....	29
2.1.2. Millikan: la téléosémantique expliquée par l'évolution.....	33
2.1.3. Critiques des théories téléosémantiques .....	35
2.1.4. Conclusion partielle .....	39
2.2. La stratégie d'intégration de l'énactivisme: de l'intentionnalité biologique à l'intentionnalité pragmatique.....	39
2.2.1. Les concepts majeurs de l'énactivisme.....	40
2.2.2. Le <i>sense-making</i> comme intentionnalité biologique au sein de l'énactivisme .....	45
2.2.3. L'ur-intentionnalité: une alternative radicale au <i>sense-making</i> pour la naturalisation de l'intentionnalité biologique? .....	49
2.2.5. La naturalisation de l'intentionnalité biologique: un succès?.....	52
2.2.6. L'émergence de l'agentivité: de l'intentionnalité biologique à l'intentionnalité pragmatique .....	54
2.3. Le Skilled Intentionality Framework: une alternative pour l'intégration des différents types d'intentionnalité? .....	60
2.4. Conclusion : une naturalisation partielle par l'intégration des intentionnalités biologique et pragmatique.....	64

CHAPITRE III NATURALISER L'INTENTIONNALITÉ REPRÉSENTATIONNELLE : MISSION IMPOSSIBLE POUR L'ÉNACTIVISME ? .....	66
3.1. Le problème du <i>scaling-up</i> de l'énactivisme .....	66
3.2. Une conception néo-pragmatiste de l'intentionnalité représentationnelle .....	68
3.3. Le <i>scaling-up</i> de l'énactivisme autpoïétique: le modèle du corps linguistique .....	72
3.3.1. Le <i>sense-making</i> participatif .....	72
3.3.2. Le modèle du corps linguistique: un <i>sense-making</i> participatif sous forme de dialogue ....	74
3.3.3. La grammaire énaactive: l'explication énaactiviste de la syntaxe et des symboles .....	78
3.4. Discussion du modèle du corps linguistique.....	82
3.4.1. Quelle notion de contenu intentionnel pour le modèle du corps linguistique?.....	82
3.4.2. Une dimension diachronique grâce à l'énactivisme radical?.....	86
3.5. Conclusion : une ébauche pour la naturalisation de l'intentionnalité représentationnelle.....	88
CONCLUSION.....	89
RÉFÉRENCES .....	94

## LISTE DES FIGURES

FIGURE	Page
Figure 1.1 Les trois types de contenu. ....	7
Figure 2.1 La fermeture opérationnelle (en noir) .....	43

## RÉSUMÉ

Ce mémoire a comme thème le problème de la naturalisation de l'intentionnalité de l'esprit par la philosophie et les sciences cognitives. Plus spécifiquement, le mémoire s'intéresse à la stratégie de naturalisation d'une famille théorique en particulier : l'énactivisme. L'énactivisme est un nouveau courant en sciences cognitives qui soutient que l'esprit n'est pas seulement l'affaire du cerveau, mais plutôt le résultat d'une interaction réciproque entre le cerveau, le corps et l'environnement. À cet effet, l'énactivisme s'oppose généralement au concept de représentation mentale. La question que le mémoire pose est alors la suivante : Comment l'énactivisme entend-il naturaliser l'intentionnalité de l'esprit? La thèse principale de ce mémoire est que la stratégie de l'énactivisme consiste à intégrer trois types d'intentionnalité dans son explication : l'intentionnalité biologique, l'intentionnalité pragmatique et l'intentionnalité représentationnelle. Dans le premier chapitre, je définis ces types d'intentionnalité de la manière suivante : l'intentionnalité biologique est la directionnalité fondamentale du vivant qui vise son monde extérieur, l'intentionnalité pragmatique est la directionnalité des actions qui vise des objets en vue de projets moteurs, et enfin l'intentionnalité représentationnelle est celle dont la directionnalité dépend de représentations (mentales ou non). Dans le second chapitre, qui porte sur la naturalisation de l'intentionnalité en elle-même, j'aborde dans un premier temps les stratégies téléosémantiques de réduction de l'intentionnalité représentationnelle à des fonctions biologiques pour y montrer les limites de ces méthodes. Puis, j'explique ce qu'est le cadre énonciviste et je montre que la stratégie de ce cadre théorique se fonde sur la thèse de continuité entre l'esprit et le vivant. De ce fait, tout vivant possède fondamentalement une intentionnalité biologique, et, comme nous l'expliquons, cette intentionnalité biologique peut se complexifier pour devenir une intentionnalité pragmatique. Enfin, dans le dernier chapitre, j'aborde les difficultés de l'énactivisme pour aborder les phénomènes cognitifs qui impliquent normalement des représentations, notamment le langage. Je présente une ébauche d'explication du langage par les énoncivistes en spécifiant que la proposition s'inspire des approches néo-pragmatistes (l'intentionnalité représentationnelle tire son origine des normes sociales), mais que le projet nécessite du travail supplémentaire pour permettre la naturalisation complète de l'intentionnalité.

Mots clés : Intentionnalité, sciences cognitives, énoncivisme, sense-making, néo-pragmatisme, naturalisation, langage, représentation.

## ABSTRACT

The theme of this dissertation is the problem of the naturalization of the intentionality of mind by philosophy and cognitive science. More specifically, the dissertation focuses on the naturalization strategy of one theoretical family in particular: enactivism. Enactivism is an emerging field in cognitive science that argues that the mind is not just the brain's business, but rather the result of a reciprocal interaction between the brain, the body and the environment. To this end, enactivism is generally opposed to the concept of mental representation. The question that the dissertation asks is then the following: How does enactivism intend to naturalize the intentionality of the mind? The main thesis of this dissertation is that the strategy of enactivism is to integrate three types of intentionality into its explanation: biological intentionality, pragmatic intentionality and representational intentionality. In the first chapter, I define these types of intentionality in the following way: biological intentionality is the fundamental directionality of the living being which aims at the external world, pragmatic intentionality is the directionality of actions which aims at objects in view of motor projects, and finally representational intentionality is that whose directionality depends on representations (mental or not). In the second chapter, which deals with the naturalization of intentionality itself, I first discuss teleosemantic strategies for reducing representational intentionality to biological functions and show the limits of these methods. Then, I explain what the enactivist framework is and I show that the strategy of this theoretical framework is based on the continuity thesis between mind and life. As a result, all living beings fundamentally possess a biological intentionality, and, as we explain, this biological intentionality can become pragmatic. Finally, in the last chapter, I discuss the difficulties of enactivism in addressing cognitive phenomena that normally involve representations, notably language. I present one enactivist explanation of language, specifying that the proposal draws on neo-pragmatist approaches (i.e., representational intentionality originates in social norms), but I show that the project requires further work to fully naturalize intentionality.

Keywords : Intentionality, cognitive science, enactivism, sense-making, neo-pragmatism, naturalization, language, representation.

## INTRODUCTION

L'esprit est souvent décrit comme possédant deux propriétés essentielles: l'intentionnalité et la conscience phénoménale. L'intentionnalité, c'est-à-dire la propriété de certains états d'être dirigés vers d'autres états, ou encore d'être à propos de ceux-ci, sera le point focal de ce mémoire. Ce concept, (ré)introduit dans le canon philosophique occidental par le philosophe autrichien Franz Brentano au XIXe siècle, est souvent désigné (notamment par Brentano lui-même) comme étant la marque du mental, c'est-à-dire que les états mentaux se distinguent des états physiques parce qu'ils sont intentionnels. Suite à Brentano, la philosophie de l'esprit s'est emparée de cette notion afin d'en étudier les propriétés singulières, notamment le fait que certains états intentionnels puissent être en relation avec des choses qui n'existent pas (une propriété qui ne se retrouve pas dans le reste du monde physique). Ainsi, des philosophes comme Jerry Fodor, John Searle ou Daniel Dennett ont tenté chacun à sa façon de formuler une théorie complète de l'intentionnalité de l'esprit.

Les XIXe et XXe siècles voient l'émergence de sciences expérimentales comme la psychologie, la biologie et, plus tard, ce que l'on nommera les sciences cognitives qui cherchent à étudier l'esprit humain et ses propriétés. Dans ce contexte, le naturalisme philosophique (l'idée selon laquelle le monde naturel comprend l'ensemble des choses qui existent, y compris l'esprit humain, et donc l'intentionnalité) devient dominant, et plusieurs concepts philosophiques feront face à une nouvelle contrainte: celle de la naturalisation, c'est-à-dire la démonstration que le concept concorde au moins (naturalisme faible) avec les sciences naturelles, voire s'y réduit (naturalisme fort). La notion d'intentionnalité n'a pas échappé à cette contrainte, et c'est précisément ce problème qui sera abordé dans ce mémoire: comment peut-on expliquer les propriétés intentionnelles de l'esprit d'une manière qui soit cohérente avec les explications issues des sciences naturelles? Pour le reste de ce mémoire, ce projet portera le nom de naturalisation de l'intentionnalité, et c'est ce problème que ce mémoire abordera. Jerry Fodor (1989), l'un des grands défenseurs de ce programme de naturalisation, va arguer que le projet, si mené à terme, permettrait de résoudre l'un des « Grands Mystères Métaphysiques » : comment expliquer l'existence de la signification (*meaning*) dans l'ordre de la nature.

Fodor et certains de ses contemporains ont cru qu'il serait possible de réussir un tel projet en réduisant l'intentionnalité du mental à celle d'une notion nouvellement issue des sciences cognitives de l'époque: la représentation. Historiquement, le concept d'intentionnalité a été très souvent associé au concept de représentation mentale (un état mental possédant un contenu représentationnel), si bien que chez certains

auteurs, les deux concepts sont considérés comme des synonymes (voir Jacob, 2003<sup>1</sup>). Cela a donné naissance à la théorie représentationnelle de l'esprit (aussi appelé cognitivisme), c'est-à-dire que les propriétés intentionnelles de l'esprit peuvent être expliquées par l'existence de structures - des représentations mentales - qui sont intentionnelles et possèdent également un corrélat neuronal. Fodor place ce programme de naturalisation - expliquer et spécifier grâce aux sciences naturelles comment un état peut être intentionnel - en continuité avec ce qu'il nomme le « programme de Grice », c'est-à-dire la réduction des propriétés sémantiques des langues naturelles et des artefacts culturels aux propriétés intentionnelles des états mentaux des locuteurs (Fodor, 1989). Autrement dit, on pourrait selon lui procéder à une double réduction: d'abord, expliquer comment les propriétés intentionnelles des artefacts culturels et du langage dérivent des propriétés intentionnelles de l'esprit et, ensuite, expliquer comment les propriétés intentionnelles de l'esprit dérivent des propriétés intentionnelles des représentations mentales, lesquelles pourraient être expliquées en termes des propriétés non intentionnelles (physiques et/ou biologiques) du cerveau et, ultimement, du monde de la nature. Fodor (1987) va indiquer qu'expliquer l'intentionnalité en des termes naturels revient à l'expliquer « in non-intentional, non-semantic, non-teleological, and in general, non-question begging vocabulary » (p. 126).

Si les différents projets de naturalisation de l'intentionnalité (notamment ceux en lien avec la théorie représentationnelle de l'esprit) ont été abondamment commentés, et sont généralement considérés comme des échecs (notamment Ramsey, 2007, que nous commenterons dans le deuxième chapitre), certaines nouvelles approches commencent à gagner en popularité au sein des sciences cognitives et de la philosophie de l'esprit. Ces théories alternatives vont remettre en cause l'idée selon laquelle l'intentionnalité (voire l'esprit) doit être expliquée en termes de représentations mentales, et vont mettre l'accent sur d'autres sources d'explications, comme les besoins vitaux des organismes, le corps vécu ou encore la théorie des systèmes dynamiques.

Ce mémoire se concentre sur une famille de théories « alternatives » s'imposant de plus en plus en sciences cognitives: l'énactivisme (Di Paolo, 2005; Di Paolo *et al.*, 2018; Di Paolo *et al.*, 2017; Thompson, 2007; Varela, Thompson & Rosch, 1991). Il suffit pour l'instant de dire que l'énactivisme est une famille de théories de la cognition voulant faire figure d'alternative au cognitivisme en avançant que la cognition est le résultat d'une interaction réciproque entre l'organisme et son environnement. Selon l'énactivisme, le cerveau n'est plus le centre privilégié pour comprendre l'esprit, mais l'analyse

---

<sup>1</sup> La *Stanford Encyclopedia of Philosophy*, qui fait souvent office de référence dans le monde philosophique anglo-saxon, est on ne peut plus explicite à cet égard, et ce dès l'ouverture du texte « Intentionality »:

In philosophy, intentionality is the power of minds and mental states to be about, to *represent*, or to stand for, things, properties and states of affairs. To say of an individual's mental states that they have intentionality is to say that they are *mental representations* or that they have contents. (Nos italiques)

doit également inclure le corps (à travers le système sensorimoteur) et l'environnement de l'organisme. L'énactivisme s'inspire notamment de la biologie et de la phénoménologie.

En prenant compte des différents thèmes abordés précédemment, la problématique du mémoire sera donc la suivante: Comment l'énactivisme entend-il naturaliser l'intentionnalité de l'esprit?

L'intérêt que l'on porte à cette famille de théories n'est pas fortuit ou arbitraire. En effet, ces dernières mettent l'accent sur l'origine naturelle de l'esprit. En effet, l'énactivisme considère l'esprit comme étant un processus dynamique (plutôt qu'une substance ou chose abstraite - comme un programme d'ordinateur) dont la nature et l'organisation rejoignent celles du vivant.<sup>2</sup> L'objectif du mémoire sera de montrer comment cette famille de théories entend aborder, dans un contexte naturaliste, l'intentionnalité comme marque du mental et ainsi mettre en lien l'esprit, la biologie, et leurs fondements philosophiques (ontologiques, épistémologiques) respectifs. Plus encore, dans ce mémoire, j'argumenterai que l'entreprise de naturalisation de l'intentionnalité de l'énactivisme diffère de la plupart des projets de naturalisation de l'intentionnalité entrepris par le cognitivisme, notamment parce que les énaïvistes refusent d'identifier l'intentionnalité à la représentation. En affirmant que la représentation ne devrait pas être le concept central de la cognition, les énaïvistes élargissent le concept d'intentionnalité pour revenir au sens premier de directionnalité. Si certains auteurs énaïvistes plus radicaux rejettent complètement la notion de représentation (Hutto & Myin, 2013; 2017), la plupart des énaïvistes acceptent que *certain*s phénomènes intentionnels de l'esprit soient aussi représentationnels. Les adhérents à cette famille philosophique vont alors proposer que d'autres types d'intentionnalité (non représentationnelle) permettent de saisir certains phénomènes de l'esprit, notamment au sein des interactions entre le cerveau, le corps et l'environnement.

Il est important de préciser qu'il est à strictement parler inexact d'utiliser une formule comme « l'intentionnalité »: il n'existe pas en effet qu'un seul concept d'intentionnalité, mais une grande panoplie de notions plus ou moins apparentées. Un passage de Schlicht (2018) est particulièrement éloquent à cet égard:

The debate about intentionality is still very much alive, with an abundance of approaches to its naturalization still on the table, including causal (Fodor 1987), structuralist/isomorphic (Cummins 1997; Shea 2014; O'Brien and Opie 2004), teleological (Dretske 1986; Millikan 1989; Neander 2016), instrumentalist (Dennett 1987), fictionalist (Sprevak 2013), measurement-theoretic (Matthews 2007), function-theoretic (Egan 2014), eliminativist (Churchland 1981), and pragmatist (Brandom 1994) positions and no consensus is in sight. To complicate matters, the renaissance of phenomenological philosophy and progress in the empirical mind sciences have led to a proliferation of notions of intentionality, including e.g. affective intentionality (Slaby *et al.* 2011),

---

<sup>2</sup> Par ailleurs, des rapprochements explicites entre ces deux familles de théories ont déjà été proposés récemment (Kirchhoff *et al.*, 2018 ; Ramstead *et al.*, 2020).

phenomenal intentionality (Kriegel 2013), motor intentionality (Kelly 2002; Jacob and Jeannerod 2003; Sinigaglia 2008), enactive intentionality (Gallagher 2017), skilled intentionality (Van Dijk and Rietveld 2017), biological intentionality (Thompson 2007), neural intentionality (Damasio 2011), nano-intentionality (Fitch 2008), and even Ur-intentionality (Hutto and Myin, 2017). (Schlicht & Starzak, 2021, p. 90)

Évidemment, il serait mal avisé dans ce contexte que de tenter de commenter toutes ces notions d'intentionnalité (exercice fastidieux qui dépasse largement le cadre de ce mémoire). Plutôt que de commenter tous les types, je propose d'identifier trois grands groupes de notions d'intentionnalité qui, j'argumente, sont présents dans les écrits énaïvistes même s'ils ne portent pas toujours le nom d'« intentionnalité ».

Je nommerai ces trois types d'intentionnalité ainsi: 1) l'intentionnalité représentationnelle, 2) l'intentionnalité pragmatique et enfin 3) l'intentionnalité biologique. L'intentionnalité représentationnelle, comme son nom l'indique, désigne les notions d'intentionnalité qui font appel à une notion de représentation (c'est-à-dire un véhicule possédant un contenu représentationnel). L'intentionnalité pragmatique, quant à elle, désigne les relations de directionnalité entre le corps et l'environnement qui ne sont pas médiées par des représentations mais par des actions. Ce sont en effet les actions de l'organisme, soutenues par le système sensorimoteur et les possibilités d'actions offertes par l'environnement, qui manifestent cette forme d'intentionnalité. Enfin, l'intentionnalité biologique désigne l'idée selon laquelle l'existence même d'un organisme vivant implique que celui-ci soit dirigé vers son environnement immédiat, le monde ambiant gagnant ainsi une valeur, et la direction de l'organisme vers le monde gagnant des normes. Ce sont les relations organisme-environnement qui sont intentionnelles en ce troisième sens.

L'un des objectifs principaux du mémoire sera ainsi de montrer comment ce triple arrangement des notions d'intentionnalité est nécessaire pour comprendre la manière dont l'énaïvisme cherche à naturaliser l'intentionnalité de l'esprit, voire peut-être, nous y reviendrons au dernier chapitre, pour la naturalisation de l'intentionnalité en général. En effet, j'argumenterai que, contrairement aux programmes de naturalisation de l'intentionnalité qui ont eu cours dans les années 1980, et qui étaient des programmes réductionnistes (réduction de l'intentionnalité à des substrats physiques), les tentatives de naturalisation effectuées par l'énaïvisme sont des tentatives d'intégration des différents types d'intentionnalité, de la plus simple (l'intentionnalité biologique) jusqu'à l'intentionnalité la plus complexe (l'intentionnalité représentationnelle), en passant par une intentionnalité intermédiaire (l'intentionnalité pragmatique).

Qu'est-ce que cette intégration implique dans le contexte de ce mémoire? Concrètement, j'utilise le terme « intégration » pour signifier que l'énaïvisme propose d'expliquer d'abord l'intentionnalité

biologique, pour ensuite étendre son cadre théorique pour expliquer l'intentionnalité pragmatique, et, enfin, l'étendre encore pour expliquer l'intentionnalité représentationnelle. L'explication de l'intentionnalité représentationnelle se base ainsi sur une sophistication de l'explication de l'intentionnalité pragmatique, qui elle-même se base sur une sophistication de l'explication de l'intentionnalité biologique. Toutefois, comme nous le verrons, il ne s'agit jamais de réductions, puisque chaque nouveau niveau voit émerger une normativité qui lui est propre et qui peut entrer en conflit avec la normativité plus basique. Nous verrons dans le second chapitre ce que tout cela signifie exactement.

Pour l'instant, il suffit de comprendre que cette manière de concevoir la naturalisation de l'intentionnalité se distingue fortement des approches cognitivistes du problème, qui ont tendance à se concentrer uniquement sur l'intentionnalité représentationnelle et pour lesquelles le défi de la naturalisation de l'intentionnalité consiste à expliquer les conditions physico-biologiques de l'existence des représentations mentales et du contenu représentationnel. Ces différents projets de naturalisation reviennent plus ou moins alors à des tentatives de réduction des représentations à des entités et relations physiques comme exposé plus tôt. Nous croyons que l'énactivisme, qui suit notamment des idées que Merleau-Ponty a développées dans *La Structure du Comportement* (1942), a cherché à fournir une explication naturalisée de l'intentionnalité de l'esprit sans pour autant chercher une réduction de celle-ci au biologique, voire au neurologique. Le fil rouge du mémoire s'inscrit dans la lignée de ce que Thompson (2007) nomme un « naturalisme non réductionniste », c'est-à-dire un naturalisme qui considère que l'esprit et le vivant se trouvent dans un même continuum, ces deux phénomènes étant eux-mêmes dans un continuum avec les choses physiques (cette vision est aussi appelée la thèse de continuité entre l'esprit et le vivant, voir Kirchoff & Froese, 2017; Thompson, 2007), tout cela sans pour autant réduire ou éliminer les domaines supérieurs vers le physique.

La thèse principale de ce mémoire sera la suivante: l'énactivisme propose une redéfinition du problème de la naturalisation de l'intentionnalité de l'esprit, qui sera conçue non plus comme une réduction physicaliste des représentations (mentales), mais plutôt comme une intégration des intentionnalités biologique, pragmatique et enfin représentationnelle, du biologique jusqu'au représentationnel.

Pour ce faire, l'un des objectifs du mémoire sera de montrer que les notions d'intentionnalité, aussi variées et complexes qu'elles puissent être, ne se limitent pas à des notions impliquant des représentations. Cette grande diversité des notions exige un retour à l'élément fondamental caractérisant l'intentionnalité: la *directionnalité*. Cette simple définition ouvre un immense champ de possibilités quant à sa naturalisation. En effet, si l'on présume que toute intentionnalité est nécessairement représentationnelle, on met la charrue conceptuelle devant les bœufs philosophiques, c'est-à-dire que, comme le dit Ramsey (2017), on règle une question empirique (« Y'a-t-il quelque chose qui agit comme des représentations dans tous les systèmes intentionnels? ») par une question métaphysique

(« L'intentionnalité étant nécessairement représentationnelle, il doit y avoir quelque chose comme une représentation dans les systèmes intentionnels. Comment concilier cela avec le naturalisme? »). Si l'on élimine le critère représentationnel, on élargit les possibilités de naturalisation de l'intentionnalité, sans nier que certains types d'intentionnalité nécessaires pour expliquer la cognition humaine peuvent être représentationnels.

Pour accompagner cet élargissement de la notion d'intentionnalité, le mémoire utilisera également une version élargie de la notion de contenu intentionnel. Généralement, la notion de contenu intentionnel est définie comme étant ce vers quoi l'état intentionnel est dirigé (Jacob, 2003). Dans la tradition qui a inspiré les sciences cognitives, comme la notion d'intentionnalité est conçue de manière représentationnelle, la notion de contenu est également conçue de manière représentationnelle, et donc la directionnalité propre à l'intentionnalité est à comprendre en termes représentationnels. Et puisque dans les sciences cognitives traditionnelles, les états mentaux sont conçus comme propositionnels (i.e., des attitudes propositionnelles), le contenu des représentations mentales est également conçu comme propositionnel (la théorie représentationnelle de l'esprit, cf. Fodor 1981), le contenu d'une représentation étant alors spécifié par ses « conditions de vérité », c'est-à-dire l'état dans lequel le monde doit être pour que la représentation soit « vraie ». Toutefois, dans ce mémoire, j'utiliserai une notion plus large de contenu qui permettra de rendre compte de différents types d'intentionnalités. En effet, la notion de « contenu » provoque certaines tensions au sein de la littérature énoncative. Par exemple, Hutto & Myin (2013; 2017) refusent carrément d'utiliser cette notion pour parler de la cognition, la réservant à l'usage de la description des manifestations socio-culturelles. D'autres énoncateurs, comme Thompson (2018), suggèrent au contraire que la notion de contenu doit être plus large que sa version représentationnelle, et qu'il existe effectivement du contenu non représentationnel. Selon Thompson (qui se base sur Merleau-Ponty et Walter Freeman), la notion de contenu implique des conditions de satisfaction, au sens large que le contenu est sujet à des normes. Pour accommoder cet usage élargi de la notion de contenu, j'utiliserai une terminologie très utile développée dans Rolla & Huffermann (2022), lesquels définissent la notion de contenu en parlant de conditions de satisfaction, mais en spécifiant trois types de contenu et trois types de conditions de satisfaction. Les trois types de contenu sont 1) le contenu basique; 2) le contenu représentationnel; 3) le contenu propositionnel. Voici un schéma pour visualiser cette idée:

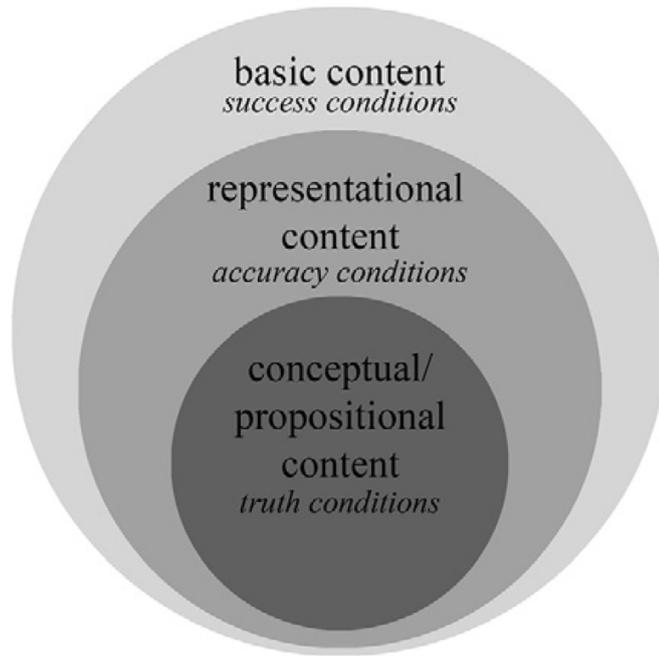


Figure 1.1 Les trois types de contenu. Schéma tiré de Rolla & Huffermann (2022), p. 354

Le contenu basique est le contenu de n'importe quel phénomène cognitif qui peut être déterminé par des conditions de succès. Cette notion de succès est normative, au sens où les actions des organismes peuvent réussir ou échouer. Par exemple, pour reprendre l'exemple classique de l'énonctivisme, le contenu de l'état « cognitif » d'une cellule motile qui cherche à remonter un gradient de glucose pour se nourrir est le glucose, et celui-ci sera jugé « *successful* » si la cellule remonte de fait le gradient de glucose. Cette notion s'applique également au niveau sensorimoteur avec des conditions de succès plus complexes. Le contenu représentationnel est défini quant à lui selon des conditions d'exactitude (*accuracy conditions*), c'est-à-dire un certain degré d'isomorphisme entre la représentation et ce qui est représenté. Cette notion est également normative, en ce sens qu'une représentation peut mé-représenter (un problème qui sera illustré et traité en détail à la section 2.1). Enfin, le contenu propositionnel est défini par des conditions de vérité, c'est-à-dire qu'un état mental propositionnel (ou une phrase) est vrai si et seulement si le monde est tel que la proposition le décrit (« la neige est blanche » est vrai ssi la neige est blanche; cf Tarski 1931). Notons que selon Rolla & Huffermann (2022), les conditions de vérité impliquent les conditions d'exactitude, qui impliquent les conditions de succès.

Avant de présenter l'organisation du mémoire, quelques mots sur l'éléphant dans la pièce: la conscience phénoménale. En effet, en ouverture de l'introduction, j'ai affirmé que l'esprit possédait deux qualités essentielles, c'est-à-dire l'intentionnalité et la conscience phénoménale. À ce stade-ci, la problématique, telle que formulée, n'aborde pas explicitement la question, et le lecteur est en droit de se demander quel

traitement sera réservé à la conscience phénoménale. Pourquoi ne pas intégrer la question de la conscience phénoménale dans la problématique?

Plusieurs raisons me motivent à ne pas traiter explicitement la question de la conscience phénoménale. D'une part, la question de la conscience a fait couler beaucoup d'encre depuis plusieurs centaines d'années, tant du côté des philosophes que des scientifiques. C'est que la conscience phénoménale - l'*effet que cela fait* - est très difficile à étudier, si bien que la problématique entourant son explication a été surnommée le *hard problem of consciousness* (Chalmers, 1995), à mettre en opposition avec les « problèmes faciles » de la conscience dont traitent la psychologie et les sciences cognitives. Pour compliquer les choses, la question du lien entre les notions d'intentionnalité et de conscience phénoménale n'est pas encore résolue non plus. En effet, certains vont argumenter que l'intentionnalité de l'esprit et des états mentaux est indépendante du caractère phénoménal de ceux-ci. Certains états intentionnels ne seraient pas dotés d'un caractère phénoménal (par exemple, des expressions qui ne réfèrent à rien mais qui ont quand même un sens sémantique, comme un « carré rond ») tandis que certains états phénoménaux ne seraient pas dotés d'un caractère intentionnel, comme la douleur. Kriegel (2013) nomme ce programme le *Naturalist-Externalist Research Program*<sup>3</sup>, c'est-à-dire l'explication de l'intentionnalité comme étant une relation naturelle entre un état interne et un état externe, un programme représenté notamment par Dennis Stampe (1977), Fred Dretske (1981), Ruth Millikan (1984), McGinn (1989) et Neander (1995), entre autres. Évidemment, la question est beaucoup plus complexe que ne le laissent suggérer ces quelques lignes, mais dans le contexte de ce mémoire de maîtrise, je me contenterai d'énoncer que s'il existe un certain lien entre l'intentionnalité et la conscience phénoménale, l'intentionnalité comprise dans toutes ses formes et non réduite à la représentation mentale n'implique pas nécessairement un caractère phénoménal. Cette position, qualifiée (péjorativement) de « séparatiste » par Horgan & Tienson (2002), a l'avantage de permettre une certaine flexibilité dans le traitement de notre problématique, c'est-à-dire que le mémoire ne tentera pas de répondre à des questions difficiles comme l'émergence de la conscience phénoménale, ou encore s'il existe des « niveaux de conscience » (par exemple entre les différentes espèces animales).

Deuxièmement, bien que les énaïvistes s'intéressent autant aux questions de l'intentionnalité et de la conscience, celles-ci ont essentiellement été séparées dans leurs études. En effet, même si la question de la conscience phénoménale et de l'expérience est primordiale chez la plupart des énaïvistes (on peut penser au sous-titre du manifeste énaïviste *The Embodied Mind, : Cognitive Science and Human Experience*), la conscience est néanmoins une notion qui a été traitée différemment de l'intentionnalité

---

<sup>3</sup> Dans ce programme, Kriegel inclut Dennis Stampe (1977), Fred Dretske (1981), Ruth Millikan (1984), McGinn (1989) et Neander (1995), entre autres.

comme telle par Varela et ses collègues. Ainsi, il paraît justifié de traiter de la question de l'intentionnalité indépendamment de celle de la conscience.

Enfin, il s'agit bêtement d'une question de focus du mémoire. En effet, si la question de la conscience phénoménale est passionnante en soi, elle mérite l'attention entière et pleine du chercheur voulant explorer les différentes problématiques lui étant associées. Ainsi, il s'agit en quelque sorte d'un choix éditorial de me concentrer sur la question de la naturalisation de l'intentionnalité, rejoignant plus précisément mes intérêts philosophiques.

Le reste du mémoire sera organisé en trois chapitres. La trame narrative des trois chapitres sera essentiellement similaire, c'est-à-dire que chaque chapitre s'ouvrira sur un problème spécifique sous le thème de la naturalisation de l'intentionnalité.

Le premier chapitre, intitulé « Trois grands types d'intentionnalité », sera consacré à définir et caractériser les trois grandes familles d'intentionnalités retrouvées dans la littérature, soit l'intentionnalité représentationnelle, l'intentionnalité pragmatique et l'intentionnalité biologique. Le but du chapitre sera simplement d'exposer les caractéristiques de chacun des types afin de mettre la table pour la suite du mémoire.

Le deuxième chapitre, intitulé « La naturalisation de l'intentionnalité », sera en quelque sorte le chapitre principal du mémoire et attaquera de front le problème de la naturalisation de l'intentionnalité. Il exposera brièvement, en premier lieu, certaines tentatives « traditionnelles » de naturalisation de l'intentionnalité, notamment les projets téléologiques de Dretske et Millikan, afin d'illustrer les limites de ce genre d'entreprise. Ensuite, j'exposerai la stratégie d'intégration employée par les différentes théories énoncivistes, principalement les travaux de Ezequiel Di Paolo. En effet, en accord avec la thèse de continuité entre l'esprit et la vie, les différentes théories prendront comme point de départ une explication naturelle de l'intentionnalité biologique grâce à la notion d'*autonomie* (Di Paolo & Thompson, 2014). Cette intentionnalité biologique prendra le nom de *sense-making* au sein de l'énoncivisme (Di Paolo, 2005; Weber & Varela, 2002). Par la suite, je m'intéresserai à la naturalisation de l'intentionnalité pragmatique par les schémas sensorimoteurs (Di Paolo *et al.*, 2017).

Le troisième chapitre poursuivra en quelque sorte là où le deuxième chapitre nous a laissés. En effet, il abordera la question du *scaling-up problem* (Gallagher, 2017). Cette expression désigne une série de critiques tirant leur origine des travaux de Andy Clark sur la « *representation-hungry cognition* », l'idée selon laquelle certains phénomènes cognitifs nécessitent absolument des représentations pour être expliqués. La question se pose alors ainsi: est-ce que l'énoncivisme peut surmonter le *scaling-up problem*? Le troisième chapitre abordera une tentative de solution de la part de l'énoncivisme. Je

m'intéresserai plus spécifiquement à l'ouvrage *Linguistic Bodies* de Di Paolo, Cuffari et de Jaegher (2018). Cet ouvrage propose un modèle énonctiviste du langage, le « modèle du corps linguistique ». J'argumenterai que ce modèle cherche à expliquer l'intentionnalité représentationnelle de manière similaire aux théories néo-pragmatistes de l'intentionnalité (représentationnelle) (Haugeland, 1990).

## CHAPITRE I

### TROIS TYPES D'INTENTIONNALITÉ

Avant de plonger au cœur de la question de la naturalisation de l'intentionnalité par l'énactivisme, il importe de bien comprendre les termes du problème qui nous préoccupe. Ce premier chapitre aura comme objectif principal de comprendre et définir ce qui est entendu par « intentionnalité » au sein de la littérature en sciences cognitives.

Le terme « intentionnalité » et les discussions entourant sa définition ont généré un grand nombre de débats philosophiques depuis la fin du XIXe siècle, et plus particulièrement pendant le XXe siècle. De manière très sommaire, il est possible de définir l'intentionnalité comme étant la directionnalité de certains états. Toutefois, un survol rapide de la littérature permet de comprendre que les termes « intentionnalités » et « directionnalité » ne sont pas univoques. Ainsi - et en gardant en tête qu'il n'est pas possible de répertorier toutes les théories sur l'intentionnalité existant et de commenter les nombreuses nuances les distinguant - le chapitre propose de classer les différents types d'intentionnalités que l'on retrouve dans la littérature en trois grands groupes: 1) l'intentionnalité représentationnelle; 2) l'intentionnalité pragmatique et; 3) l'intentionnalité biologique.

Le plan pour le reste du chapitre est assez simple. Je m'attarderai d'abord à résumer sommairement la définition classique de l'intentionnalité (telle que ressuscitée par Brentano au XIXe siècle), puis je présenterai les trois grands groupes d'intentionnalité ainsi que quelques tentatives de naturalisation pour chacun des types afin de mettre la table pour le projet de naturalisation de l'intentionnalité par l'énactivisme qui sera détaillé au prochain chapitre.

#### 1.1. La conception « classique » de l'intentionnalité

Le terme « intentionnalité » désigne, au sens classique mis de l'avant par Franz Brentano (et généralement accepté par la communauté philosophique; voir Jacob, 2003), la propriété de certaines choses d'être *à propos* d'autres choses, les exemples paradigmatiques étant des états mentaux comme les croyances et les désirs (qui ont pour objet quelque chose d'autre qu'eux-mêmes). Brentano, au XIXe siècle, a réintroduit en philosophie la notion scolastique d'« intentionnalité » par le biais de son analyse des états mentaux caractérisés par l'inexistence intentionnelle, c'est-à-dire la propriété des choses mentales d'exister dans l'acte mental lui-même (n'ayant donc pas besoin d'exister dans le monde réel<sup>4</sup>).

---

<sup>4</sup> La notion d'*inexistence* intentionnelle porte parfois à confusion dans la littérature contemporaine en philosophie de l'esprit. En effet, le sens d'inexistence est parfois assimilé à celui de la *non-existence*, c'est-à-dire que

L'intentionnalité caractérise la relation entre les choses mentales et les choses du monde et c'est cette caractéristique qui distingue les choses mentales des choses physiques, ces dernières n'étant pas intentionnelles. L'idée générale de Brentano est résumée dans ce passage très célèbre de la *Psychologie d'un point de vue empirique* (1874):

Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not all do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on. This intentional inexistence is characteristic exclusively of mental phenomena. No physical phenomenon exhibits anything like it. We can, therefore, define mental phenomena by saying that they are those phenomena which contain an object intentionally within themselves. (Brentano, 1995/1874, p. 68)

Ce passage est encore énormément cité dans la littérature contemporaine concernant l'intentionnalité, si bien qu'il constitue un passage presque obligé pour quiconque cherchant à aborder cette notion. S'il existe encore aujourd'hui des défenseurs d'une conception véritablement néo-brentanienne de l'intentionnalité plus fidèle à la théorie de Brentano<sup>5</sup>, la plupart des chercheurs contemporains en philosophie de l'esprit ne se contentent que de citer ce passage dans le but de mettre en lumière la directionnalité ainsi que l'*aboutness* (le caractère « à propos ») des choses mentales. Cette notion intuitive de la directionnalité de l'intentionnalité est particulièrement saillante dans la dernière phrase du premier paragraphe : « In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on. » (Brentano, 1874, p. 68). Dans ces actes, selon Brentano, il y a un « objet intentionnel », la chose qui est visée par ces actes. Les croyances et les désirs sont des exemples paradigmatiques d'états mentaux intentionnels, puisque les croyances et les désirs ont toujours un objet (on croit quelque chose, on désire quelque chose).

Un problème classique avec la notion d'intentionnalité émerge dès que l'on observe que l'on peut aisément croire à des choses qui n'existent pas (comme le Père Noël) ou désirer des choses futures (comme un diplôme de maîtrise) et qu'on se demande comment expliquer que des états mentaux puissent être dirigés vers des choses non existantes (comme des états futurs, ou des choses imaginaires). Sans rentrer dans les détails techniques, il est possible de comparer la version « immanente » de l'intentionnalité de Brentano avec la version « relationnelle » héritée de Husserl (Fisette & Poirier,

---

l'inexistence intentionnelle décrirait la relation à des objets imaginaires ou non-existant. Il faudrait plutôt comprendre *inexistence* comme in-existence, c'est-à-dire « existence-dans ». C'est pour cela que l'on peut dire que les objets *inexistent* dans l'acte selon Brentano. (Voir à ce sujet Fisette & Poirier, 2000, p. 169, note 1.)

<sup>5</sup> Voir notamment les travaux de Tim Crane (1998) et de Kriegel (2013) pour une théorie de l'intentionnalité plus fidèle aux enseignements de Brentano.

2000). La version immanente de l'intentionnalité considère que l'acte mental, ordinaire, vise un objet extraordinaire, qui « inexiste intentionnellement » dans l'acte. La version « relationnelle » va considérer que l'intentionnalité des états mentaux est une propriété sémantique plutôt qu'ontologique. Les objets vers lesquels sont dirigés les états mentaux sont ordinaires (et uniques), mais la relation entre l'état mental et son objet va varier en fonction du mode de présentation de cet objet. En effet, un objet (unique) ne sera jamais entièrement apprécié « d'un bloc » par l'esprit : il sera toujours présenté d'une certaine manière, sous un certain aspect. Pour Husserl, ce n'est plus l'objet de l'acte mental qui est « spécial », c'est plutôt la relation entre cet acte et son objet. Encore une fois, sans affirmer que les sciences cognitives sont profondément husserliennes, il est possible d'affirmer que la version relationnelle de l'intentionnalité de l'esprit a prévalu sur la version immanente et que les différentes théories contemporaines de l'intentionnalité sont (presque) toutes relationnelles.

Bref, l'intentionnalité, dans son sens le plus fondamental, est la propriété de certains états d'être dirigé vers des choses ou vers un phénomène physique, cette chose ou ce phénomène étant prise comme l'objet intentionnel de l'acte (par exemple, le son étant l'objet intentionnel de l'audition). L'intentionnalité est conçue traditionnellement comme la marque de l'esprit, c'est-à-dire que seuls les états mentaux sont intentionnels.

## 1.2. L'intentionnalité représentationnelle

Jusqu'à présent, j'ai exposé une conception sommaire de l'intentionnalité de l'esprit centrée sur l'idée selon laquelle l'intentionnalité réfère à une « directionnalité » de l'esprit, c'est-à-dire que les états mentaux sont dirigés vers des choses autres qu'eux-mêmes. Cette directionnalité est comprise selon les termes d'une relation entre les états mentaux et des choses du monde (qui peuvent exister ou ne pas exister). Dans cette section, j'exposerai les notions les plus populaires d'intentionnalité, c'est-à-dire les notions qui font appel à la notion de représentation. Si l'on soutient que certains états ont des propriétés intentionnelles et qu'il est possible de naturaliser ces propriétés, comment s'y prendre? L'une des stratégies ayant dominé la seconde moitié du XXe siècle consiste à développer une théorie représentationnelle de l'esprit et d'expliquer l'intentionnalité de l'esprit grâce à celle des représentations mentales qui les constituent et qui ont des propriétés intentionnelles et causales.

De manière générale, une représentation est une chose qui tient lieu (stand-in) d'autres choses. Dès lors, une représentation mentale est un état mental qui tient lieu d'autres choses. Une représentation est dite porter un contenu représentationnel, c'est-à-dire ce qui est représenté par la représentation (Haugeland, 1990). Par exemple, la photo d'un chien est une représentation du chien dont la photo a été prise; ma pensée de ce même chien en est également une représentation, mentale cette fois-ci. Il est possible de remarquer que le concept de représentation rejoint en grande partie la formulation générale de

l'intentionnalité (en termes de directionnalité) que nous avons exposée dans la section précédente. Il est possible de lier les deux concepts en affirmant que les représentations mentales sont dirigées vers leur contenu. Cette conception de l'intentionnalité sera nommée intentionnalité représentationnelle pour le reste du mémoire.

Les conceptions représentationnelles de l'intentionnalité sont très populaires chez les auteurs contemporains de philosophie de l'esprit.<sup>6</sup> Si la notion de représentation mentale peut être saisie de manière intuitive, son examen plus minutieux révèle certaines ambiguïtés qu'il est important de clarifier. D'une part, lorsque l'on parle de « représentation », il faut faire attention à savoir si l'on parle d'un explanandum (le phénomène à expliquer) ou d'un explanans (ce qui sert à expliquer). Nous associerons, dans le cadre de ce mémoire, le terme « intentionnalité représentationnelle » au phénomène (à expliquer, explanandum) de la représentation (c'est-à-dire la propriété de l'esprit ou de tokens physiques de pouvoir être dirigé vers autre chose), tandis que le terme « représentation (mentale) » sera réservé à la désignation des construits théoriques (les représentations mentales portant un contenu représentationnel et/ou propositionnel) servant à expliquer le (donc appelés à faire partie d'explanans du) phénomène de l'intentionnalité. Cette distinction est à mettre en lien avec la distinction personnel/sous-personnel (Dennett, 1969; voir aussi Bennett & Hacker, 2003 pour une critique), c'est-à-dire qu'il faut distinguer l'intentionnalité représentationnelle explicite au niveau personnel (par exemple, lorsque l'on pense à quelque chose et qu'on se le représente en tête) et les représentations mentales sous-jacentes qui sont censées expliquer comment l'on peut penser à des choses et agir en lien avec ces pensées. L'intentionnalité se révèle au niveau personnel (par exemple, lorsque l'on a une croyance ou un désir), et dans les théories les plus populaires en philosophie de l'esprit, cette intentionnalité est expliquée par des représentations mentales qui portent le contenu des croyances ou des désirs.

Fodor est l'un des défenseurs les plus célèbres d'une théorie représentationnelle de l'esprit, qu'il développera en détail dans *The Language of Thought* (1975) et perfectionnera dans *A theory of content* (1990). La théorie représentationnelle de l'esprit (TRE) telle que développée par Fodor soutient que pour expliquer l'intentionnalité des états mentaux (notamment celle des attitudes propositionnelles comme les croyances et les désirs), il faut s'intéresser aux fonctions des états mentaux qui lient les stimuli sensoriels et les réponses comportementales (on ouvre la « boîte noire » des béhavioristes). Fodor soutient que les états mentaux qui agissent comme intermédiaires sont des représentations mentales possédant à la fois des propriétés sémantiques (un contenu représentationnel) et syntaxiques (des règles

---

<sup>6</sup> La réputée *Stanford Encyclopedia of Philosophy* fait même équivaloir l'intentionnalité et la représentation : « To say of an individual's mental states that they have intentionality is to say that they are mental representations or that they have contents. » (Jacob, 2003) Évidemment, nous ne désirons pas construire un argument d'autorité quant à la définition de l'intentionnalité, mais nous souhaitons simplement mettre en exergue la popularité de la définition « représentationnelle » de l'intentionnalité.

pour être combinées à d'autres représentations mentales). L'ensemble des représentations mentales forment ce que Fodor appelle un langage de la pensée qui possède une intentionnalité originelle, et c'est de cette intentionnalité des états mentaux que l'intentionnalité des propositions en langues naturelles est dérivée<sup>7</sup> (Fodor, 1975; 2008).

Il est important de noter que Fodor défend un réalisme intentionnel, c'est-à-dire qu'il existe réellement des états ayant des propriétés intentionnelles. Par exemple, il existerait réellement des croyances ou des désirs, et ces croyances et ces désirs auraient réellement des propriétés intentionnelles. Ces croyances et ces désirs sont des attitudes propositionnelles, c'est-à-dire des attitudes par rapport à une proposition (voir Nelson, 2000 pour une explication simple). Cette position est à contraster avec 1) l'instrumentalisme (Dennett, 1987), qui soutient qu'il n'existe pas réellement d'états possédant des propriétés intentionnelles, mais qu'il est utile de prétendre que oui puisque cette fiction possède un grand pouvoir prédictif quant aux comportements de nos congénères (c'est ce que Dennett nomme la posture intentionnelle, un niveau d'explication supérieur et irréductible aux explications mécaniques/causales); et avec 2) l'éliminativisme (Churchland, 1981), qui soutient que l'ensemble des descriptions intentionnelles des comportements forment une théorie du sens commun qui est inadéquate et qui devrait être éliminée du langage scientifique.

La théorie de Fodor est loin d'être la seule version existante du computationnalisme. En effet, un grand nombre de théories en sciences cognitives vont reprendre, d'une manière ou d'une autre, le concept de représentation pour tenter d'expliquer ce qui est proprement « cognitif » au sein d'un système. Pour reprendre les mots de Newell (1980), la représentation a une telle importance historique au sein du paradigme principal des sciences cognitives<sup>8</sup> qu'elle peut être comparée à « what the theory of evolution is to all of biology, the cell doctrine to cellular biology, the notion of germs to the scientific concept of disease (...) » (p. 136). Ainsi, sans rentrer dans les détails de l'ensemble des théories représentationnelles existantes (un travail disproportionné, voire impossible, surtout compte tenu des contraintes d'un mémoire de maîtrise), il est possible d'affirmer avec assurance que la représentation

---

<sup>7</sup> La formulation originelle-dérivée a également été reprise par Haugeland (1990). Elle se rapproche mais se distingue de celle de Searle (1980), pour qui les artefacts culturels ne seraient pas intrinsèquement intentionnels, mais leur intentionnalité dépendrait plutôt d'un esprit capable de faire la connexion entre l'artefact et ce qu'il représente (la distinction étant intrinsèque-relatif à l'observateur). Pour une plus ample discussion à ce sujet, voir Haugeland (1990).

<sup>8</sup> L'idée selon laquelle les sciences cognitives ont formé un seul paradigme est parfois remise en question. En effet, Chemero (2009), dans sa monographie *Radical embodied cognitive science*, se montre sceptique quant au fait que les sciences cognitives aient atteint le statut de « science normale » (au sens que Kuhn a développé dans *La structure des révolutions scientifiques*). Pour Chemero, les sciences cognitives, si elles sont prises en un bloc, sont encore à l'état pré-scientifique, c'est-à-dire que les théories déployées sont souvent défendues par des arguments *a priori* de type hégélien plutôt que par l'expérimentation. Ainsi, l'idée de *représentation* serait devenue un *a priori* du fonctionnement et de la conduite des sciences cognitives, et il serait impossible de songer à faire de la science cognitive autrement. Plus de détails sur cette critique seront amenées à la section 1.3.

est le principal moyen utilisé en science cognitive pour tenter d'expliquer le fonctionnement de l'esprit, et par le fait même ses propriétés intentionnelles.

Par ailleurs, peu importe la théorie étudiée, il est très important de noter que les représentations doivent avoir la possibilité de mal représenter le monde réel. Autrement dit, la possibilité de la mé-représentation est une caractéristique essentielle de la représentation. Par exemple, un système cognitif peut se méprendre en voyant un animal et se le représenter comme un loup, alors que c'est un gros chien. L'un des grands défis de la naturalisation de l'intentionnalité représentationnelle consiste justement à fournir les conditions naturelles qui permettent de déterminer comment et quand les organismes se trompent dans l'utilisation de leurs représentations. Développer une théorie pouvant expliquer comment un système cognitif peut-il mal utiliser des représentations est l'une des motivations derrière les projets biologiques-informationnels de Dretske (1981; 1988; 1995) et Millikan (1984; 1989), qui seront abordés dans le prochain chapitre. Ces deux projets veulent fonder la théorie représentationnelle de l'esprit sur des bases informationnelles et biologiques afin d'offrir une explication naturelle à l'intentionnalité de l'esprit.

### 1.3. L'intentionnalité pragmatique

Au sein de la dernière section, je me suis affairé à montrer d'une part la conception représentationnelle de l'intentionnalité (des représentations mentales portant un contenu intentionnel). D'autre part, j'ai exploré quelques tentatives de naturalisation de ce type d'intentionnalité. Je me suis notamment intéressé aux solutions de Fodor, Dretske et Millikan qui cherchent à baser l'intentionnalité sur des entités naturelles, à savoir l'information et les téléofonctions. Toutefois, nous avons vu que ces solutions ne sont pas entièrement satisfaisantes, entre autres parce qu'elles ont tendance à présupposer l'existence de l'intentionnalité qu'elles sont censées naturaliser.

Si cette vision représentationnelle de l'intentionnalité a été historiquement la plus populaire, surtout dans l'écosystème philosophique anglo-saxon du XXe siècle, elle n'est pas la seule. En effet, tous ne sont pas d'accord avec l'interprétation de la notion d'intentionnalité comme étant nécessairement l'affaire d'un contenu représentationnel ou propositionnel. Pour contrer cette vision de l'esprit centrée sur la représentation, différentes théories ont tenté de proposer une conception de l'intentionnalité (la propriété de l'esprit d'être dirigé vers autre chose) centrée autour de l'action de l'organisme (ou de l'individu) sur son environnement. Nous regroupons différents concepts d'intentionnalité développés par plusieurs théories sous le nom d'« intentionnalité pragmatique ». Cette expression cherche à capturer un thème commun de ces théories, à savoir que la directionnalité de l'esprit n'implique pas nécessairement un contenu sémantique, mais plutôt des engagements pratiques dans le monde.

### 1.3.1. Une intentionnalité sans représentation

À première vue, l'expression « intentionnalité pragmatique » peut sembler étrange, surtout à la suite de ce qui a été dit précédemment au sujet du contenu représentationnel. Après tout, même la réputée *Stanford Encyclopedia of Philosophy* fait correspondre les deux notions: « In philosophy, intentionality is the power of minds and mental states to be about, *to represent*, or to stand for, things, properties and states of affairs. » (Jacob, 2003, nos italiques). Par l'expression « intentionnalité pragmatique », nous ne désignons pas une notion d'un auteur en particulier, mais plutôt un ensemble de notions apparentées dans leur tentative d'expliquer la relation de l'esprit au monde sans recourir au concept de représentation, mais plutôt au concept d'action, et ce, en mettant l'accent sur l'accès direct au monde par l'organisme, en outre par le biais du corps dont la contribution active à la directionnalité de l'intentionnalité est parfois négligée par les tenants de la théorie représentationnelle et/ou computationnelle de l'esprit (qui n'y voit qu'une simple réalisation des fonctions de l'esprit, laquelle serait non pertinente dans la détermination de ses caractéristiques).<sup>9</sup>

Pourquoi s'opposer à l'utilisation des représentations pour expliquer l'intentionnalité de l'esprit? À ce stade de mon mémoire, il est possible d'identifier deux raisons principales, une épistémologique et l'autre ontologique. Une première raison pourrait être le constat de l'échec des tentatives de naturalisation des représentations et la conclusion que la notion de représentation elle-même n'est pas un bon candidat pour expliquer (de manière naturaliste) l'intentionnalité de l'esprit. Cet argument est de nature épistémologique, mais il semble plutôt marginal dans la littérature anti-représentationnaliste. En effet, la deuxième raison, de nature ontologique, est plus fondamentale: elle critique la conception représentationnaliste comme étant une affaire de médiation entre l'esprit et le monde (Zahavi, 2018). Selon cette critique, l'esprit est toujours directement dirigé vers le monde; il n'y a pas d'intermédiaire représentationnel jouant le rôle de médiateur donnant à l'esprit accès au monde. En proposant des représentations comme intermédiaire entre l'esprit et le monde, on ferait fausse route sur la nature même de l'esprit et de notre expérience.

Pourquoi alors parler d'intentionnalité si l'on cherche à éviter (ou du moins à limiter) l'usage des représentations dans la caractérisation de l'esprit? Nous croyons que l'enjeu revient à capturer la caractéristique de la directionnalité de l'esprit, c'est-à-dire l'idée selon laquelle l'esprit est dirigé vers les choses du monde.<sup>10</sup> Il est important de préciser que refuser les représentations mentales comme

---

<sup>9</sup> Il existe néanmoins certaines approches computationnelles qui cherchent à mettre de l'avant le rôle du corps matériel dans les activités de la cognition tout en conservant le concept de représentation. Ces approches prennent souvent le nom de *embodied cognitive science* (voir Chemero, 2009 ; Clark, 2013).

<sup>10</sup> L'enjeu rejoint, d'une certaine manière, les débats entre l'internalisme et l'externalisme de l'esprit. (Clark & Chalmers, 1998 ; Godfrey-Smith, 2013; Ramstead *et al.*, 2021) En effet, si le débat entre l'internalisme et l'externalisme concerne les « limites » de la cognition (i.e. ce qu'on peut qualifier d'« esprit »), notre mémoire

explication de l'intentionnalité de l'esprit n'équivaut pas à nier l'existence du phénomène de l'intentionnalité de l'esprit.<sup>11</sup> Poser l'existence de l'intentionnalité de l'esprit, c'est poser qu'il est nécessaire d'expliquer les comportements des organismes comme étant plus qu'une simple succession d'événements et de mouvements mécaniques; poser l'existence de cette intentionnalité n'implique toutefois pas que celle-ci soit nécessairement représentationnelle.

### 1.3.2. L'intentionnalité pragmatique chez Merleau-Ponty et Gallagher

De quoi aurait l'air une intentionnalité non représentationnelle ou pragmatique? Pour répondre à cette question, les travaux de Merleau-Ponty, notamment *La Structure du Comportement* (1942) et *Phénoménologie de la perception* (1945), deviennent des références incontournables.<sup>12</sup> En effet, Merleau-Ponty, dans la *Phénoménologie de la perception*, va (ré)introduire un concept que nous pouvons placer sous le parapluie de l'intentionnalité pragmatique: l'*intentionnalité motrice*. L'intentionnalité motrice désigne chez Merleau-Ponty une compréhension antéprédicative et pré-réflexive du monde par le biais du corps, qui est placé au centre des interactions entre l'esprit et le monde.

Merleau-Ponty détaille les contours de l'intentionnalité motrice en citant l'exemple de Schneider, un soldat ayant été blessé au combat et ayant des problèmes neurologiques. Schneider est incapable d'accomplir un mouvement abstrait. Par mouvement abstrait, on veut dire par exemple de « lever le bras à une certaine hauteur ». Un mouvement concret serait par exemple de ramasser un objet devant soi. Pour Schneider, accomplir un mouvement abstrait est très difficile, et ce n'est qu'en gesticulant au hasard qu'il peut espérer accomplir le mouvement demandé. Merleau-Ponty décrit ainsi la situation du malade:

Ce qui lui manque n'est ni la motricité, ni la pensée, et nous sommes amenés à reconnaître entre le mouvement comme processus en troisième personne et la pensée comme représentation du mouvement une anticipation ou une saisie du résultat assurée par le corps

---

s'intéresse plutôt au *lien* entre le système cognitif et le monde, à la manière dont l'esprit entre en relation avec le monde.)

<sup>11</sup> Cette position niant l'existence de quelque chose comme l'intentionnalité existe néanmoins et se retrouve dans la littérature sous le nom d'éliminativisme (voir par exemple les travaux de Paul et Patricia Churland, même si leur position s'adresse plus directement à l'intentionnalité représentationnelle, voir 1981). Une autre version de cette négation de l'intentionnalité pourrait s'apparenter à une position behavioriste forte, où seuls les comportements sont étudiés et la notion même d'*esprit* est évitée.

<sup>12</sup> Les travaux de Merleau-Ponty sont de beaucoup redevables aux travaux de Heidegger, notamment l'ontologie existentielle du Dasein et de l'être-au-monde développée dans *Être & Temps* (1927). Nous ne nions pas du tout cette interprétation, mais pour des raisons d'espace et de proximité conceptuelle, nous limiterons notre analyse des origines de l'intentionnalité non-représentationnelle aux travaux de Merleau-Ponty, puisque ce sont avec ceux-ci que les théories à l'étude dans ce mémoire, l'énactivisme et l'inférence active, sont plus intimement apparentées.

lui-même, un « projet moteur » (*Bewegungsentwurf*), une « intentionnalité motrice » sans lesquels la consigne demeure lettre morte. (Merleau-Ponty, 1945, p. 141)

Merleau-Ponty prend la peine de préciser que les demandes faites à Schneider ont une signification sémantique pour lui, mais elles n'ont pas de signification motrice. L'argument de Merleau-Ponty consiste à dire que si ni la compréhension verbale, ni la pensée, ni la motricité elle-même ne sont affectées mais que le malade est tout de même incapable d'accomplir la tâche, c'est qu'il manque quelque chose comme un « projet moteur » ou une « intentionnalité motrice », c'est-à-dire que l'action elle-même a (normalement) une signification indépendante de toute compréhension sémantique à son propos. Autrement dit, l'esprit (par le biais du corps) est dirigé vers le monde, mais sans l'intermédiaire d'une représentation.

Le concept est parfois nommé « intentionnalité opérante » dans la littérature énaïviste, et c'est pour éviter les multiples désignations que je propose « intentionnalité pragmatique » pour parler du concept en général. Je réserve « intentionnalité motrice » pour désigner le concept de Merleau-Ponty. C'est justement parce que la directionnalité de l'intentionnalité est l'affaire d'une action que nous donnons le nom d'intentionnalité pragmatique à différents concepts, dont l'intentionnalité motrice de Merleau-Ponty et l'intentionnalité opérante de Gallagher. Gallagher décrit succinctement l'intentionnalité opérante ainsi:

(...) the concept of operative intentionality attempts to capture the fact that the experiencing agent is intentionally engaged with the world through actions and projects that are not reducible to simple mental states, but involve an intentionality that is motoric and bodily. (Gallagher, 2017, p. 78)

On retrouve l'importance de la directionnalité dans cette définition de l'intentionnalité, mais cette directionnalité est plutôt le fait d'actions dirigées vers le monde par des actions et des projets plutôt que le fait d'actes mentaux dirigés vers le monde par l'intermédiaire d'une représentation portant un contenu représentationnel<sup>13</sup>. Comme le précise Thompson (2007), ce type d'intentionnalité n'est pas l'affaire d'un réflexe ni de l'intellect, mais plutôt l'affaire d'une compétence (*skill*).

Gallagher argue également que c'est par ce type d'intentionnalité que l'on peut comprendre la signification des actions des autres. Par exemple (Miyahara, 2011), on pourrait imaginer un conducteur

---

<sup>13</sup> En plus de l'intentionnalité motrice, Merleau-Ponty présente également un autre concept intentionnel: l'arc intentionnel. L'*arc intentionnel* désigne l'ensemble des rapports (ou des horizons) sous lesquels est tendue la conscience afin d'être une conscience unifiée; l'arc comprend par exemple le passé, les possibilités futures, le milieu humain (socio-culturel), l'environnement physique, etc. (Merleau-Ponty, 1945, p. 170) Encore une fois, Merleau-Ponty utilise une pathologie en exemple pour montrer la nature de ce type d'intentionnalité. Chez un patient malade, les différentes expériences ne peuvent être facilement mise en lien entre elles. La vie se présente alors comme une succession ininterrompue d'expériences, et c'est l'arc intentionnel devrait normalement soutenir l'unité de ces expériences.

voyant un piéton sur le bord du trottoir qui regarde activement de gauche à droite. Le conducteur se prépare alors à freiner, ou du moins passe son pied au-dessus de la pédale de frein au cas où. Si quelqu'un lui demandait pourquoi il a fait ça, il pourrait dire que le piéton semblait vouloir traverser. Toutefois, ce serait tenter d'expliquer le comportement a posteriori grâce à des raisons ou des désirs. Plus originellement, placer le pied sur la pédale de frein revient simplement à expérimenter l'intentionnalité (pragmatique) du piéton. Merleau-Ponty utilise la notion d'intercorporéité pour décrire cette situation. Dans la Phénoménologie de la perception, donne également en exemple l'intentionnalité érotique pour expliquer l'intercorporéité. Pour Merleau-Ponty (1945), dans une situation érotique, c'est le corps qui s'adresse directement à l'autre corps, qui est dirigé vers lui, sans qu'il y ait un intermédiaire représentationnel entre les deux. Pour Merleau-Ponty et Gallagher, c'est ce type d'intentionnalité qui devrait être considéré comme étant l'intentionnalité non dérivée (« primary and non-derived intentionality », Gallagher, 2017, p. 79; « la motricité comme intentionnalité originale », Merleau-Ponty, 1945, p. 171).

Le but de cette section est de montrer qu'il est possible de retrouver dans la littérature philosophique et scientifique un ensemble de notions apparentées qui peuvent être qualifiées d'« intentionnelles » même si elles ne sont pas représentationnelles. À la lumière de ce que nous avons présenté sommairement, nous pouvons comprendre ce que nous appelons l'intentionnalité pragmatique comme étant la directionnalité des actions (qui relèvent d'une certaine forme de cognition) vers le monde externe, et ce sans qu'il n'y ait besoin d'une médiation par des représentations ou encore d'une réduction à un béhaviorisme.

#### 1.4. L'intentionnalité biologique

Dans les précédentes sections, j'ai abordé différents concepts d'intentionnalité présents dans la littérature en philosophie et en sciences cognitives. D'une part, j'ai présenté les grandes lignes de l'intentionnalité représentationnelle ainsi que certaines tentatives de naturalisation par le biais de l'information et des fonctions biologiques. D'autre part, j'ai mentionné que certains penseurs et chercheurs qui, pour diverses raisons, s'opposaient au concept d'intentionnalité représentationnelle - ou du moins à son hégémonie - et ont proposé d'autres concepts d'intentionnalité pour rendre compte de l'intentionnalité de l'esprit, concepts que nous avons rassemblés sous le terme -parapluie d'intentionnalité pragmatique.

Toutefois, notre présentation des différents concepts d'intentionnalité ne pourrait être complète sans aborder un autre type d'intentionnalité, encore plus fondamental selon certains auteurs. C'est ce que je nommerai « intentionnalité biologique ». Comme c'était le cas pour l'intentionnalité pragmatique, l'intentionnalité biologique n'est pas le concept d'un auteur particulier, mais plutôt une catégorie de

concepts plus ou moins apparentés reprenant la même idée fondamentale. Cette idée, c'est que l'organisme vivant, pour pouvoir survivre et maintenir son intégrité, doit s'ouvrir vers le monde externe et que cette ouverture est sélective et dirigée. Ce sont cette sélectivité et cette directionnalité de l'ouverture sur le monde externe qui font les bases d'un comportement intentionnel, c'est-à-dire non pas une simple réactivité biochimique, mais un comportement qui est dirigé vers un environnement *significatif*.

D'entrée de jeu, je tiens à clarifier que ce que je nomme « intentionnalité biologique » n'est pas à rapprocher d'une vision téléfonctionnaliste de l'intentionnalité. En effet, dans le téléfonctionnalisme, notamment celui de Millikan (1984; 1989), il est question de la naturalisation (biologique-évolutionniste) d'une intentionnalité représentationnelle, c'est-à-dire l'explication biologique des conditions dans lesquelles une représentation peut agir comme représentation et acquérir un contenu. L'intentionnalité biologique est encore plus fondamentale que les intentionnalités représentationnelle et pragmatique, puisqu'elle concerne l'origine de la directionnalité du comportement dans le monde naturel. Bref, l'idée principale derrière le concept d'intentionnalité biologique est que l'activité d'un organisme vivant témoigne d'une ouverture dirigée de manière spécifique vers le monde. Qu'est-ce que cela veut dire?

Pour répondre à cette question, je m'intéresserai particulièrement aux réponses qu'ont tenté d'apporter les travaux de Merleau-Ponty et Jonas en philosophie ainsi que ceux de Schrödinger en physique. Le premier s'est intéressé au lien entre le physique, l'organique et le psychique. Le second s'est intéressé à la question des normes naturelles, et de la téléologie intrinsèque. Le troisième s'est intéressé au phénomène du vivant comme étant celui de l'ordre généré à partir de l'ordre.

#### 1.4.1. Les germes d'une intentionnalité biologique : Merleau-Ponty, Schrödinger et Jonas

L'idée d'une intentionnalité biologique s'inscrit de pair avec les questionnements entourant la définition du vivant au XXe siècle. En effet, la catégorie du « vivant », relativement nouvelle au début du XXe siècle, n'est pas sans problèmes théoriques. C'est l'époque où la biologie a délaissé le vitalisme - théorie où le vivant est défini par une « force vitale » irréductible aux forces physiques - en raison des progrès des sciences physico-chimiques. Ces progrès amènent la biologie à définir le vivant comme étant un phénomène strictement physique. Néanmoins, si le vivant est un phénomène strictement physique, comment distinguer ce qui est vivant de ce qui ne l'est pas? Et comment intégrer le domaine de la psychologie dans une compréhension physico-chimique du vivant?

Merleau-Ponty, dans *La Structure du Comportement* (1942), va chercher à répondre à ce genre de questions, notamment au troisième chapitre dédié aux relations entre ce qu'il nomme les ordres

physique, vital (biologique/vivant), et humain (mental/psychologique). L'un des buts principaux de Merleau-Ponty est de redéfinir la notion de comportement au-delà de la conception béhavioriste, alors dans son âge d'or. On peut notamment voir au sein de cette œuvre une genèse du concept d'intentionnalité pragmatique, par exemple lorsque Merleau-Ponty décrit le comportement non pas comme une « série de points géographiques », mais des « mouvements considérés dans leur articulation intérieure (...) comme une mélodie cinétique douée d'un sens » (Merleau-Ponty, 1942, p. 196).

En plus de l'intentionnalité pragmatique, il est également possible de retrouver chez Merleau-Ponty la genèse d'un autre type d'intentionnalité: l'intentionnalité biologique. En effet, l'intégration entre le physique, le vivant et le mental souhaitée par Merleau-Ponty vise à dépasser deux écueils: le matérialisme (dont le réductionnisme rend très mal compte du mental) et le vitalisme-spiritualisme (qui explique mal la continuité entre les différents ordres). Pour accomplir cette intégration, Merleau-Ponty introduit la notion de *structure*, c'est-à-dire une forme dont il est impossible de décrire les éléments individuels sans faire référence à tous les autres. Par exemple, un champ électrique ou une goutte d'huile placée au milieu d'une masse d'eau ne sont pas simplement un amoncellement d'ondes ou de particules, mais bien un système dynamique dont la description ne peut se réduire à celle de ses parties. Toutefois, si une structure physique demeure un équilibre en tension obtenu à l'égard des conditions extérieures, l'action de ces forces extérieures a toujours pour effet de réduire cet état de tension et d'acheminer le système vers le repos. Par exemple, une tornade (qui est une structure, ou une forme) est générée par l'action de masses d'air chaud et d'air froid créant un différentiel de pression. La tornade s'écoule vers l'équilibre et le système retombe au repos.

L'idée de Merleau-Ponty est de montrer l'originalité des structures vitales (vivantes) par rapport aux structures physiques, sans pour autant tomber dans le vitalisme ni dans le matérialisme. Merleau-Ponty présente la spécificité des systèmes organiques ainsi:

Nous parlons de structures organiques, lorsque l'équilibre est obtenu, non pas à l'égard de conditions présentes et réelles, mais à l'égard de conditions seulement virtuelles que le système amène lui-même à l'existence, - lorsque la structure, au lieu de procurer, sous la contrainte des forces extérieures, une détente à celles dont elle est traversée, exécute un travail hors de ses propres limites et se constitue un milieu propre. (Merleau-Ponty, 1942, p. 221)

Ce passage contient plusieurs idées importantes pour le développement d'un concept d'intentionnalité biologique. Il est dit que le système maintient son équilibre à l'égard de conditions « virtuelles » qu'il amène lui-même à l'existence. Ces conditions sont virtuelles dans la mesure où elles sont une possibilité, soit celle du maintien (ou non) de l'intégrité du système organique. Ces conditions sont déterminées par le système lui-même, et ne sont pas exprimables par une loi, mais plutôt par une norme, la norme de sa

propre continuation. L'impératif de ce maintien de l'ordre peut être exprimé tant d'un point de vue thermodynamique que phénoménologique.

En se basant sur une description thermodynamique, Erwin Schrödinger, dans le chapitre 7 de *What is Life?* (1944), caractérise le vivant comme étant un système qui perpétue l'ordre (« *the order-from-order principle* »), contrairement à des systèmes physiques non organiques qui peuvent produire un ordre momentané à partir du désordre (« *order-from-disorder principle* »), mais qui ont tendance à retourner au désordre. Autrement dit, une structure organique est une structure (physique) qui maintient un état de tension ; cet état de tension doit être maintenu par l'activité de la structure organique en raison de la tendance universelle des systèmes physiques d'augmenter l'entropie (la dissipation de l'énergie disponible pour effectuer un travail), suivant la seconde loi de la thermodynamique. Autrement dit, un système organique doit accomplir un certain travail pour combattre l'entropie et maintenir son intégrité structurelle, et ce travail n'est pas gratuit: cela demande du temps, de l'énergie et des ressources. Une structure organique est donc une structure active plutôt que passive.

La même idée (celle du maintien d'un état de tension) se retrouve également dans la « biologie philosophique » de Hans Jonas, même si ce dernier s'oppose à la mathématisation du vivant. En effet, dans son essai « *Is God a mathematician?* » (1966), le philosophe présente son intuition selon laquelle c'est sa propre expérience vécue – inscrite corporellement - qui lui permet de comprendre le phénomène du vivant: « On the strength of the immediate testimony of our bodies we are able to say what no disembodied onlooker would have a cause for saying (...) – the point of life itself: its being self-centered individuality, being for itself (...) » (1966, p. 79). Selon Jonas, cette expérience vécue nous amène à comprendre que le vivant se voit inscrit d'une « téléologie intrinsèque », en ce sens que la vie devient sa propre fin. Cette « téléologie intrinsèque » s'accomplit par un processus d'auto-affirmation d'une identité métabolique, c'est-à-dire d'une forme qui participe à son propre maintien.

Pour Jonas, cette identité métabolique ne se réduit donc pas à une identité logico-physique  $A=A$  où les choses restent identiques à elles-mêmes par simple vertu de la continuité des dimensions spatio-temporelles dans lesquelles elles se situent. En effet, une forme physique, comme une vague, ne cherche pas à se maintenir face à l'augmentation de son désordre : la « forme » a plutôt le caractère d'un « accident » momentané de la matière, résultat du mouvement des particules d'eau suivant une loi. La « forme » de la vague ne force pas les molécules à bouger d'une certaine manière et une fois la vague passée, les particules retrouvent un état désordonné (Jonas, 1966).

À propos de la forme organique, Jonas précise qu'elle se maintient par le changement de sa matérialité. En effet, le philosophe explique que contrairement à une machine où chaque pièce conserve son identité par vertu de la simple continuation de sa matérialité spatio-temporelle, l'organisme métabolique doit

continuellement renouveler et changer sa matérialité afin de conserver sa forme, sa structure (Jonas, 1966). Cela rejoint l'idée de Merleau-Ponty voulant qu'un système présentant une forme conserve les mêmes propriétés si tout change en son sein sans que la relation entre ses éléments change.

#### 1.4.2. L'intentionnalité biologique comme directionnalité fondamentale du vivant

À partir de ces idées de Merleau-Ponty, Schrödinger, et Jonas, il nous est possible de comprendre le lien entre le problème du vivant et le problème de l'intentionnalité. J'ai précédemment exposé que l'intentionnalité (qu'elle soit représentationnelle ou pragmatique) consistait en une propriété de certains systèmes d'être dirigé (de manière normative) vers des choses autres qu'eux-mêmes, généralement des choses du monde. L'argument derrière l'existence d'une intentionnalité biologique revient à dire qu'en vertu des propriétés des structures organiques, quelque chose comme une directionnalité devient une propriété nécessaire du vivant: elle émerge de lui. En effet, nous avons vu qu'une structure organique, qui doit lutter contre l'augmentation de son entropie pour se perpétuer, doit nécessairement faire « transaction » avec son environnement. Schrödinger nommait ce phénomène, de manière quelque peu métaphorique, « l'extraction de l'ordre depuis l'environnement » (Schrödinger, 1944, p. 73). Autrement dit, l'organisme vivant, en tant que système qui tend à se maintenir dans un état de tension perpétuelle (*non-stable thermodynamic equilibrium*), a besoin d'énergie et de matière venant de son environnement, c'est pourquoi celui-ci doit être dirigé vers celui-ci.

Toutefois, il est important de préciser que l'ouverture de l'organisme vers le monde est sélective. En effet, un système étant ouvert de manière indifférenciée à son environnement se dissiperait comme le font les systèmes physiques non vivants qui ne peuvent que subir les forces extérieures les ramenant à l'équilibre thermodynamique. Chez les systèmes vivants, l'ouverture se doit d'être sélective puisque tous les éléments ne contribuent pas de la même façon au maintien de l'état de tension perpétuelle de l'organisme. De cette interaction différenciée entre l'organisme et son environnement émergent des valences (par là c'est bon, par là c'est mauvais) déterminées en fonction de la perpétuation de l'organisme, laquelle devient la valeur suprême selon laquelle toutes les interactions seront évaluées. Comme le dit Merleau-Ponty dans l'extrait étudié dans la section précédente, la structure organique va se « constituer un milieu propre ». Cela fait dire à Merleau-Ponty que le système organique « mesure lui-même l'action des choses sur lui et délimite lui-même son milieu par un processus circulaire qui est sans analogue dans le monde physique » (Merleau-Ponty, 1942, p. 225). Autrement dit, l'environnement physique externe devient un milieu, et le monde ambiant devient parcouru de valences. C'est par cette interaction qu'il peut effectuer le renouvellement matériel de sa forme. L'activité de l'organisme est ainsi toujours dirigée vers le monde en fonction des différentes valences qui émergent de l'interaction environnement-organisme. C'est précisément cette relation directionnelle normative entre l'organisme

et son milieu que je nomme intentionnalité biologique: l'intentionnalité de base qui émerge de l'activité de n'importe quelle structure organique, et sans laquelle celle-ci se dissiperait.

À ce stade-ci, le lecteur est en droit de se poser quelques questions. En effet, quel est le lien entre cette notion générale de directionnalité et d'ouverture au monde et ce que la tradition philosophique nomme l'esprit ou le mental ? De plus, quelle est l'utilité de cette caractérisation somme toute assez générale de l'intentionnalité ? N'est-il pas évident que les organismes doivent être « dirigés vers » le monde pour survivre ? J'avance que la pertinence de cet argumentaire est multiple. D'une part, cela relève du constat de la possibilité d'ancrer les discussions sur la nature fondamentale de l'intentionnalité (et, par le fait même, de l'esprit) dans le vocabulaire naturaliste de la théorie des systèmes dynamiques ouverts pour proposer une naturalisation concrète et plausible de l'intentionnalité de l'esprit. D'autre part, parler d'intentionnalité biologique permet de faire un lien avec la thèse de continuité entre l'esprit et le vivant (Kirchhoff, 2018; Kirchhoff & Froese, 2017; Thompson, 2007). Cette thèse soutient que l'esprit et le vivant possèdent le même ensemble de propriétés organisationnelles (Clark, 2001; Godfrey-Smith, 1996) et que l'esprit est « préfiguré » (*foreshadowed*) par le vivant. Cette thèse, à plusieurs égards, suit les enseignements de Jonas, qui ouvre *The Phenomenon of life* sur ces propos:

A philosophy of life comprises the philosophy of the organism and the philosophy of mind. This is itself a first proposition of the philosophy of life, in fact its hypothesis, which it must make good in the course of its execution. For the statement of scope expresses no less than the contention that the organic, even in its lowest forms prefigures mind, and that mind even in its highest reaches remains part of the organic. (Jonas, 1966, p. 1)

Pourquoi prendre la peine d'introduire ce troisième type d'intentionnalité? Ne serait-il pas possible de rendre compte de ces phénomènes à l'aide de l'intentionnalité pragmatique? À mon sens, l'intentionnalité biologique se distingue de l'intentionnalité pragmatique puisque cette directionnalité de base n'est pas nécessairement accomplie par un système sensorimoteur mais est plutôt le reflet de la nature fondamentale de l'organisme, à savoir celle d'un être luttant contre l'entropie thermodynamique et se perpétuant lui-même. Autrement dit, l'intentionnalité biologique est l'affaire des normes vitales de base des organismes, tandis que l'intentionnalité pragmatique est l'affaire d'un système sensorimoteur permettant à l'organisme d'agir sur son monde afin de satisfaire les normes vitales, même s'il n'y réussit pas toujours. En ce sens, l'intentionnalité biologique est plus fondamentale par rapport à l'intentionnalité pragmatique.

Un exemple pourrait aider à clarifier la distinction entre les deux types d'intentionnalité: Imaginons un lion qui chasse un zèbre dans la savane. Ici, un premier type d'intentionnalité est à l'œuvre, à savoir l'intentionnalité biologique qui fait du zèbre une proie *pour* le lion *en fonction* de ses propres besoins et de ses normes vitales (l'obtention de protéines dans la viande du zèbre pour permettre à l'organisme du lion de se perpétuer dans ses opérations biochimiques). Le lion est *ouvert* au zèbre sans le recours à une

représentation du concept « proie » qui serait appliqué « au » zèbre. Le zèbre *est* une proie par la simple existence du lion. Lorsque le lion tente de s'approcher du zèbre, un second type d'intentionnalité est à l'œuvre : l'intentionnalité pragmatique. En effet, pour avancer prudemment et s'approcher du zèbre, puis choisir le moment d'attaque, l'action du lion cible non seulement le zèbre comme objet physique, mais également le sol, l'herbe, le vent pour éviter de se faire sentir, etc. Ainsi, l'intentionnalité pragmatique est l'affaire d'un *projet sensori-moteur*, ce projet étant lui-même justifié par l'intentionnalité biologique fondamentale, à savoir le besoin pour le lion de se nourrir. Ainsi, même si la cognition du lion jouit de deux types d'intentionnalité en gros caractérisables comme « biologiques », il existe néanmoins une distinction entre celles-ci qui, je soutiens, est pertinente et nécessaire dans le cadre du projet de naturalisation de l'intentionnalité par l'énactivisme.

Bref, l'idée derrière le concept d'intentionnalité biologique est que l'activité du vivant, en tant que structure cherchant à maintenir son intégrité face à l'augmentation naturelle de son entropie, témoigne d'une directionnalité minimale, et cette directionnalité vers le monde externe prend la forme d'une évaluation activement accomplie par le vivant, qui voit son environnement se transformer en milieu doté de valences.

#### 1.5. Conclusion: vers une redéfinition du problème de la naturalisation de l'intentionnalité

Tout au long de ce chapitre, j'ai catégorisé trois types d'intentionnalité: l'intentionnalité représentationnelle, l'intentionnalité pragmatique et l'intentionnalité biologique. D'une part, l'intentionnalité représentationnelle conçoit l'intentionnalité comme l'affaire de représentation, c'est-à-dire des structures qui possèdent un contenu représentationnel. D'autre part, l'intentionnalité pragmatique conçoit l'intentionnalité comme étant la signification d'une action pour un organisme. Enfin, l'intentionnalité biologique conçoit l'intentionnalité comme l'affaire de la directionnalité basique du vivant vers son environnement.

Revenons au problème de départ de notre mémoire: la naturalisation de l'intentionnalité. Dans les sections précédentes, nous avons vu quelques tentatives de naturalisation de l'intentionnalité se concentrant uniquement sur l'intentionnalité représentationnelle et qui argumentent que le défi de la naturalisation de l'intentionnalité consiste à expliquer les conditions physico-biologiques de l'existence des représentations mentales et du contenu représentationnel. C'est pourquoi j'ai exposé trois grands types d'intentionnalité dans ce chapitre. D'une part, le but était de montrer que les notions d'intentionnalité, aussi variées et complexes qu'elles puissent être, ne se limitaient pas aux versions représentationnelles. La grande diversité des notions nous a amené à revenir à l'élément fondamental caractérisant l'intentionnalité: la directionnalité.

Suite à ce que j'ai exposé au long de ce chapitre, le reste de mon mémoire sera consacré à argumenter en faveur d'une redéfinition le problème de la naturalisation de l'intentionnalité qui ne se concentre pas uniquement sur l'intentionnalité représentationnelle, mais qui intègre véritablement les trois types d'intentionnalité. Ainsi, le but ne sera plus de naturaliser l'intentionnalité représentationnelle, mais plutôt d'offrir une théorie intégrant l'intentionnalité biologique, pragmatique et représentationnelle de manière à ne laisser ni saut mystérieux ni réduction abusive.

Dans le prochain chapitre, j'argumenterai que c'est la stratégie employée par les différentes théories énoncivistes. En effet, en accord avec la thèse de continuité entre l'esprit et la vie, les différentes théories prendront comme point de départ une explication naturelle de l'intentionnalité biologique grâce à la notion d'autonomie, une amélioration du concept d'autopoïèse (Di Paolo & Thompson, 2014). Cette intentionnalité biologique prendra le nom de *sense-making* au sein de l'énoncivisme (Weber & Varela, 2002; Di Paolo, 2005). Par la suite, je m'intéresserai à la naturalisation de l'intentionnalité pragmatique par les schèmes sensorimoteurs (Di Paolo, 2017).

## CHAPITRE II

### LA NATURALISATION DE L'INTENTIONNALITÉ

Le précédent chapitre avait pour but d'introduire le problème de la naturalisation de l'intentionnalité en expliquant ce qu'il fallait naturaliser. La proposition principale du chapitre était de redéfinir le problème de naturalisation de l'intentionnalité comme étant l'affaire d'une intégration de plusieurs types d'intentionnalités, de la plus simple (l'intentionnalité biologique), en passant par l'intentionnalité pragmatique, jusqu'à l'intentionnalité représentationnelle. L'idée de l'intégration des différents types d'intentionnalité doit être comprise comme une continuité - de manière analogue avec la thèse de la continuité entre l'esprit et la vie (Thompson, 2007; Kirchhoff, 2018). La signification de cette continuité sera précisément l'objet de ce chapitre, qui s'intéressera plus précisément à comment l'énactivisme entend naturaliser les différents types d'intentionnalité.

Le deuxième chapitre sera divisé en deux sections principales. Dans un premier temps, je présenterai rapidement comment le problème de la naturalisation de l'intentionnalité a été abordé historiquement en me concentrant sur les théories téléosémantiques (Dretske, 1981; 1988; Millikan, 1989). Les théories téléosémantiques conçoivent le projet de naturalisation comme étant celui de l'explication biologique (voire physique) de l'intentionnalité représentationnelle. Autrement dit, pour naturaliser l'intentionnalité, il faut naturaliser les représentations mentales. La stratégie téléosémantique consiste à expliquer le concept de représentation grâce à une téléologie ayant une origine ontogénétique ou phylogénétique. Je me concentrerai sur les approches téléosémantiques pour deux raisons. D'une part, ces approches ont eu une influence considérable sur la question de la naturalisation de l'intentionnalité jusqu'au point d'être parfois considéré comme les approches « les plus prometteuses » (Cash, 2008; Hutto & Satne, 2015; Ritchie, 2014; voir aussi Rosenberg, 2013). D'autre part, ces approches accordent une grande importance à la dimension biologique de l'explication de l'intentionnalité (Godfrey-Smith, 2006). Ces deux caractéristiques font de la téléosémantique un candidat idéal pour comprendre non seulement comment a été traitée historiquement la question de la naturalisation de l'intentionnalité, mais pour comprendre également comment l'énactivisme se distingue des théories biologiques l'ayant précédé. Je présenterai également des critiques des théories téléosémantiques.

Dans un second temps, j'exposerai la théorie énonciviste et la manière dont elle propose de naturaliser les différents types d'intentionnalité. La stratégie énonciviste de naturalisation de l'intentionnalité est guidée par deux idées principales: 1) la cognition devrait être expliquée dans des termes principalement anti-représentationalistes; 2) il existe une continuité entre l'esprit et le vivant. Ces deux idées amènent des auteurs énoncivistes à proposer que l'intentionnalité dans son sens fondamental est une propriété

essentielle du vivant et que c'est grâce à cette propriété fondamentale du vivant qu'il sera possible d'expliquer d'autres formes d'intentionnalité, i.e. l'intentionnalité pragmatique et l'intentionnalité représentationnelle. Dans la section 2.2, j'illustrerai premièrement les concepts fondamentaux de l'énaclivisme pour ensuite expliquer comment cette théorie utilise ces concepts pour naturaliser l'intentionnalité biologique et l'intentionnalité pragmatique. En raison de ses difficultés particulières, le cas de l'intentionnalité représentationnelle sera réservé au troisième chapitre.

## 2.1. Tentatives passées de la naturalisation de l'intentionnalité représentationnelle : les théories téléosémantiques

Comme nous avons vu précédemment, l'une des stratégies les plus populaires pour naturaliser l'intentionnalité consiste à le faire grâce à la notion de représentation (mentale). Ainsi, si l'on accepte cette idée, on peut expliquer comment l'esprit est dirigé vers des objets: c'est la représentation mentale qui porte un contenu représentationnel, et c'est la relation entre la représentation et son contenu représentationnel qui est intentionnelle.

Toutefois, affirmer que l'intentionnalité est le fait de représentations mentales n'est pas suffisant, en soi, pour assurer une véritable naturalisation. En effet, l'un des défis des théories représentationnelles de l'esprit consiste à expliquer, dans des termes non intentionnels, comment des choses physiques peuvent posséder ou acquérir des propriétés intentionnelles (Fodor, 1987). À cet égard, les travaux de Dretske et Millikan sont particulièrement intéressants pour comprendre comment on a tenté d'expliquer de manière naturaliste les propriétés intentionnelles de l'esprit, et ainsi répondre au défi de Fodor.

### 2.1.1. Dretske: de l'information à la téléosémantique

Le projet de Dretske cherche à expliquer les propriétés des représentations mentales grâce à la notion d'*information*. Pour Dretske, l'intentionnalité des représentations mentales dérive de relations informationnelles qui sont utilisées au sein de systèmes biologiques et/ou cognitifs, et le travail de Dretske consiste à expliquer comment la fonction de représentation peut dériver des propriétés intentionnelles naturelles de l'information sans faire intervenir un interprète.

Dans le sens qui nous concerne ici, l'information est une notion issue de la théorie de la communication de Shannon (1948) et désigne, dans son sens statistique premier, la réduction de l'incertitude associée à un ensemble de possibilités générées par un événement. Le gain d'information se produit lorsque l'événement permet d'éliminer des possibilités, et ce gain peut être transmis ou communiqué par une source vers un récepteur. Par exemple, si on lance un dé, il y a six possibilités, et le fait que le dé tombe sur un trois réduit les possibilités à un, ce qui réduit l'incertitude associée à cet événement. Il y a donc

un gain d'information lorsque les possibilités sont réduites de six à une. D'entrée de jeu, Dretske (1981) prend la peine de spécifier que la notion d'information de Shannon est statistique (c'est-à-dire qu'elle est en fait une moyenne), et non pas sémantique. L'idée de Dretske sera donc de fonder correctement une théorie sémantique de l'information sur laquelle il sera possible de construire les propriétés sémantiques de la cognition.

Selon Dretske, l'information est une quantité objective, c'est-à-dire qu'elle existe indépendamment des observateurs (même si Dretske reconnaît lui-même que tout le monde n'est pas d'accord avec lui sur ce point). Un événement porte de l'information sur un autre événement ssi il y a une véritable relation nomologique entre les deux événements, c'est-à-dire que l'un est la cause de l'autre. Par exemple, les anneaux d'un arbre portent de l'information sur l'âge de l'arbre, indépendamment de toute observation, puisque le passage du temps entraîne nécessairement la croissance de l'arbre (toute chose étant égale par ailleurs). En termes de probabilité, A porte de l'information sur B si la probabilité de B sachant A est de 1. Dretske (1988) nomme cette relation informationnelle indication<sup>14</sup> (Dretske, 1988; Fisette & Poirier, 2000; voir également Peirce, 1974<sup>15</sup>) et elle est considérée par plusieurs, dont Dretske, comme étant la plus basique des formes d'intentionnalité existante. Le contenu informationnel d'un état est alors ce avec quoi l'état co-varie nomologiquement.

Rappelons que l'idée principale derrière la naturalisation de l'intentionnalité représentationnelle est d'expliquer l'intentionnalité de l'esprit par l'intentionnalité des représentations mentales, puis de naturaliser les propriétés de ces dernières. Une représentation mentale est dite *intensionnelle*, avec un s, c'est-à-dire qu'elle échoue aux deux tests d'extensionnalité<sup>16</sup>. Elle doit également pouvoir mé-représenter (i.e. représenter faussement ou se tromper quant à) l'état du monde. Pour Dretske, c'est cette possibilité d'erreur qui marque la distinction entre l'information et l'intentionnalité représentationnelle.

---

<sup>14</sup> Hutto & Myin (2013) établissent également une distinction entre ce qu'ils nomme l'*information-comme-covariance* et l'*information-comme-contenu*.

<sup>15</sup> Il est possible que le lecteur pense à la notion d'*indice* en sémiotique, c'est-à-dire le second type de signe selon la trichotomie de Peirce. L'indice est un signe dont la signification est tirée d'une relation causale ou corrélative entre le *representamen* (le signifiant) et l'objet (le signifié). Toutefois, il faut faire attention aux amalgames, puisque la théorie de Peirce est beaucoup plus complexe qu'une simple relation d'indication et implique toujours un troisième élément dans la relation entre le signifiant et le signifié: l'interprétant. Les discussions entourant le statut ontologique de l'interprétant dépassent le cadre du mémoire, mais il est possible de préciser rapidement qu'un interprétant n'est pas une *interprétation* cognitive faite par un sujet. Un interprétant est *ce qui permet* de fonder la signification entre le signifiant et le signifié. Pour une introduction plus détaillée à la doctrine de Peirce, voir Short (2007).

<sup>16</sup> Techniquement, l'intensionnalité (avec un s) désigne ce qui échoue aux deux tests d'extensionnalité, à savoir la *substitution des identiques* (on peut remplacer deux expressions sémantiques ayant le même référent et conserver la même valeur de vérité) et la *généralisation des existentiels* (si A est extensionnel et que l'on utilise l'expression « A est P »,  $\exists x(P(x))$  est vraie). (Fisette et Poirier, 2000)

La question qui se pose est alors la suivante: en présupposant que l'information est la forme la plus basique d'intentionnalité dans la nature et compte tenu du fait que l'information en elle-même ne peut pas être fautive ou erronée, comment expliquer la capacité de mé-représentation, et par le fait même le passage d'une relation informationnelle à une relation représentationnelle?

Une première solution s'apparente à ce que Fodor (1987) (voir également Fisette et Poirier, 2000) a nommé la théorie causale rudimentaire. Grosso modo selon cette idée, si un objet A cause le token A', on peut dire que A' représente A. Tel que soulevé par Fisette et Poirier, les problèmes de cette approche simpliste sont évidents. D'une part, si le token A' peut être causé par des objets non-A (un cas de mé-représentation), on ne peut pas dire que c'est parce qu'ils ont causé A' que A' les représente. D'autre part, comme on s'attend à ce que les représentations soient générales, on imagine que A' peut représenter toute occurrence de A. Or, la plupart des occurrences de A ne causeront jamais directement A' (puisque l'organisme a des limites spatio-temporelles).

Fisette et Poirier (2000) précisent que des auteurs comme Fodor (1990) et Dretske (1988) vont surmonter ce problème en prenant une posture nomologique, c'est-à-dire que s'il existe une loi telle que les occurrences de A causent un token A', on peut dire que A' représente A. Cet ajout a l'avantage de régler le problème des occurrences non rencontrées par l'organisme. L'important est le type d'occurrence de pomme A qui peut causer « pomme », A'. Toutefois, un autre problème pour ce type de solution est le problème de la disjonction. Le problème de la disjonction est le problème de la sous-détermination du contenu d'une représentation. En effet, si l'on croit qu'une représentation A' représente A, il n'y a rien qui nous empêche de penser qu'elle représente que B ou C peuvent causer des occurrences de A'. Notons que la disjonction s'applique pour ce qu'on peut appeler des homologues sensoriels, c'est-à-dire des choses qui peuvent générer des états sensoriels similaires. Le problème est donc de s'assurer que A' représente bel et bien A et non pas A-ou-B, A-ou-B-ou-C, etc.

Dretske (1988; 2000) et Millikan (1989) vont élaborer une solution téléologique au problème de la disjonction: une relation informationnelle peut devenir représentationnelle à condition que l'information ait (ou acquiert) une fonction de représentation dans un système biologique et/ou cognitif. Dretske (1988) nomme ce type de système représentationnel un système représentationnel naturel. Pour comprendre les caractéristiques d'un système représentationnel *naturel*, il les contraste avec deux autres types de systèmes représentationnels *conventionnels*.

Le premier type de système représentationnel conventionnel (type 1) désigne les systèmes symboliques. Dans ce genre de système, les éléments qui le composent (les symboles) n'ont pas de pouvoir indicateur intrinsèque. Leur intentionnalité représentationnelle est toujours dérivée. Un symbole, selon Dretske, est doublement conventionnel en ce sens que tant la relation d'indication (par exemple que le mot

français « Pomme » indique une pomme) que l'utilisation de cette relation d'indication dépend d'un interprète. Comme le résume Dretske, « we give [symbols] a job to do, and we do it for them » (Dretske, 1988, p. 54).

Le second type de système représentationnel conventionnel (type 2) est un système où des relations d'indication naturelle (de l'information) se voient assigner une fonction de représentation par un interprète. Dretske (2000) donne des exemples de relations informationnelles qui acquièrent une fonction représentationnelle de cette manière: celles qu'on utilise pour construire des instruments de mesure. En effet, un thermomètre fonctionne grâce au lien nomologique entre la température et l'expansion du mercure. L'expansion du mercure porte de l'information quant à la température, et cette relation informationnelle est utilisée par le concepteur du thermomètre pour concevoir l'instrument de mesure. Ainsi, c'est un observateur qui attribue une fonction à un lien informationnel, et c'est de là que peut être dérivée une signification représentationnelle (par exemple, celle exprimée par le thermomètre et qu'on peut exprimer en retour par la phrase « Il fait 20 degrés. »). Dretske précise que d'un point de vue fonctionnel, la signification représentationnelle portée par le thermomètre devient indépendante de la cause directe de l'augmentation du volume du mercure. En effet, un thermomètre est censé indiquer la température, même s'il peut être utilisé dans des conditions où ce n'est pas un changement de température qui cause l'expansion du mercure. Si le mercure venait à prendre de l'expansion pour une autre raison que celle de l'augmentation de la température, le thermomètre échouerait à indiquer la température réelle, et donc à la représenter. Par exemple, si on apportait le thermomètre en haute altitude, la différence de pression atmosphérique affecterait également l'expansion du mercure. S'il est vrai que le thermomètre indique la modification de la pression atmosphérique (ainsi qu'une panoplie d'autres propriétés physiques indirectes), ce n'est pas sa fonction. Ainsi, si on amenait le thermomètre en haute altitude sans le calibrer correctement pour prendre en compte la modification de la pression atmosphérique, on pourrait dire du thermomètre qu'il mé-représente la température puisque, malgré la non-correspondance entre la température représentée et la température réelle, il conserve sa fonction d'indiquer la température et est utilisé de la sorte. Ainsi, la signification représentationnelle de l'instrument est séparée des causes de ses changements physiques, tout comme sa fonction est séparée de sa performance. C'est pourquoi pour Dretske, la solution au problème de la mé-représentation (et, au final, de la naturalisation de l'intentionnalité représentationnelle) se trouve bel et bien dans la téléologie, puisque c'est grâce à l'assignation d'une fonction que l'on peut passer de la relation informationnelle à la relation représentationnelle.

On comprend alors que l'explication de l'existence des systèmes représentationnels naturels consiste à déterminer comment l'information peut se voir attribuer intrinsèquement et naturellement une fonction dans un système biologique et/ou cognitif, sans faire intervenir un interprète (puisque'il serait circulaire

de se baser sur un interprète possédant des capacités représentationnelles pour expliquer les capacités représentationnelles d'un interprète). Selon Dretske:

If an information-carrying element in a system could somehow acquire the function of carrying information, and acquire this function in a way that did not depend on our intentions, purposes, and attitudes, then it would thereby acquire (...) the power to misrepresent the conditions it had the function of informing about. (Dretske, 2000, p. 216)

Pour résoudre le problème de l'attribution de fonctions naturelles dans un système biologique, deux pistes principales de solutions s'offrent aux chercheurs: l'une phylogénétique (l'évolution) et l'autre ontogénétique (le développement et l'apprentissage). Dretske croit que la meilleure solution est une explication ontogénétique de l'acquisition naturelle de la fonction de représentation par une relation informationnelle. Imaginons un animal qui ne possède pas de comportement instinctif face à un nouveau prédateur (noté C pour contexte) dans son environnement. L'animal peut acquérir de l'information (percevoir) quant à la présence du nouveau prédateur, mais ne présente pas un comportement d'évitement qui lui permettrait de survivre. Pour que l'animal puisse survivre, il faut que l'indicateur interne de C soit intégré par un mécanisme d'apprentissage dans une boucle d'action qui permet de générer l'évitement. L'idée est que si l'indicateur interne de C est intégré au comportement d'évitement parce qu'il indique C, il gagne alors la fonction d'indiquer C dans la boucle d'action qui génère l'évitement. Dans les mots de Dretske:

Internal elements that supply needed information acquire the function of supplying it by being drafted (in this case, through a learning process) into the control loop because they supply it. A supplier of information acquires the function of supplying information by being recruited for control duties because it supplies it. (Dretske, 2000, p. 218)

### 2.1.2. Millikan: la téléosémantique expliquée par l'évolution

Millikan (1984; 1989) développe quant à elle une théorie phylogénétique. Un système biologique peut acquérir la fonction d'utiliser l'information pour créer une représentation s'il a évolué en ce sens grâce aux mécanismes évolutifs comme la sélection naturelle<sup>17</sup>. C'est ce qu'elle nomme la biosémantique ou encore la téléosémantique. Mon intérêt pour la théorie de Millikan découle du fait que celle-ci sera reprise chez les énoncivistes radicaux (Hutto & Myin, 2013; 2017) qui tentent de fonder l'intentionnalité naturelle sans utiliser de notion de contenu représentationnel. La solution de Millikan, qui se rapproche de Dretske sur certains aspects, s'en distingue en raison du vocabulaire spécifique qu'elle introduit, notamment les notions de fonction propre, d'explication normale et de condition normale. Une fonction

---

<sup>17</sup> Il est intéressant de noter que Dretske (1990) rejette l'argument évolutionniste de Millikan, argumentant qu'il ne peut pas expliquer l'intentionnalité représentationnelle des croyances, désirs, plans, etc. Plutôt, Dretske va arguer que l'*apprentissage* peut expliquer comment un indicateur interne gagne la fonction d'indiquer une condition C, et que cette indication peut être utilisée par le système cognitif pour l'associer avec le comportement approprié. Toutefois, les détails de sa solution dépassent les besoins d'exposition des théories téléosémantiques dans le cadre de ce mémoire.

propre est la fonction d'un système telle que déterminée par l'histoire évolutive de ce système en accord avec son design. L'explication normale est ce qui permet d'expliquer comment une fonction a historiquement été performée, dans les (peut-être rares) cas où la fonction a correctement performé. Les conditions normales sont les conditions qui permettent de formuler une explication normale complète, c'est-à-dire les conditions nécessaires pour qu'une fonction puisse performer normalement.

Millikan établit que pour qu'il y ait des représentations au sein d'un système biologique et/ou cognitif, il faut que ce système soit composé d'au moins deux parties ayant des fonctions différentes: l'appareil producteur de la représentation et l'appareil consommateur de la représentation, qui l'interprète. Pour Millikan, seul l'appareil consommateur est nécessaire pour comprendre comment se créent les représentations et comment se fixe leur contenu représentationnel, puisque c'est lui qui utilise les représentations produites. L'idée est que si l'on comprend ce qui permet à l'appareil consommateur de prendre comme (*taking as*) des signes naturels comme portant une information spécifique (a, b ou c plutôt que e, f ou g, par exemple), nous pouvons dériver une sémantique de l'appareil consommateur (en présupposant que l'interprétation du consommateur dérive systématiquement de la structure des signes). Ce qu'un signe naturel porte véritablement comme information, ce qu'il indique, n'importe pas pour l'appareil consommateur: la fonction de l'appareil producteur sera alors de produire des signes qui ont une signification pour l'appareil consommateur (Millikan, 1989, p. 286). La naturalisation de l'intentionnalité représentationnelle dépend alors de la compréhension de la manière dont l'appareil consommateur interprète les représentations de l'appareil producteur comme des représentations. Puisque Millikan ne se concentre que sur la consommation des signes, elle offre deux critères pour qu'un consommateur puisse utiliser un état interne comme une représentation:

1) L'adéquation entre la représentation et ce qui est représenté devient une condition normale de la fonction propre de l'appareil consommateur dans sa réponse à la représentation.

2) Des transformations dans la forme de la représentation impliquent des variations par rapport à ce qui est représenté. Ces transformations peuvent être décrites systématiquement, formant ainsi une sémantique des représentations.

L'exemple célèbre de Millikan (1989) concerne les danses que les abeilles performent pour communiquer l'emplacement de certains lieux de pollinisation. Pour Millikan, les mécanismes d'interprétation des abeilles (ce qui doit potentiellement consommer la représentation) ne rempliront pas complètement leur fonction propre d'aider à recueillir du nectar selon l'explication normale à moins que le lieu indiqué par la danse corresponde réellement au lieu où se trouve le nectar. On retrouve ici la première condition, c'est-à-dire que l'exactitude de la représentation constitue une condition normale d'explication de la fonction de l'abeille. De plus, comme les variations du tempo et de l'angle de la

danse indiquent respectivement la distance et la direction du lieu, Millikan argumente que la seconde condition est respectée, c'est-à-dire qu'il est possible d'établir une règle de correspondance entre les variations dans la représentation et les variations dans le domaine représenté. Autrement dit, on aurait affaire à une sémantique des danses des abeilles. Bref, selon Millikan, c'est par la téléologie que l'on peut naturaliser l'intentionnalité représentationnelle: un état en représente un autre si une partie du système cognitif a comme fonction propre d'utiliser d'autres états comme représentations afin d'accomplir d'autres fonctions, et que l'adéquation entre la représentation et l'état représenté constitue une condition normale de l'explication normale de la fonction en question.

En bref, l'idée principale derrière les projets de naturalisation de l'intentionnalité de Dretske et Millikan consiste à utiliser les relations naturelles entre les choses du monde (l'information), puis expliquer comment cette information peut être utilisée par un système cognitif et ainsi acquérir une fonction représentationnelle.

### 2.1.3. Critiques des théories téléosémantiques

Si les théories téléosémantiques ont été historiquement très populaires, elles ne font pas l'unanimité en vertu d'un certain nombre de problèmes les affligeant. D'une part, Fodor (1990) rejette l'aspect téléologique de l'explication du contenu des représentations, puisque selon lui cela ne permet pas de résoudre le problème de la disjonction. Fodor se montre particulièrement critique de Millikan, puisque selon lui la sélection naturelle, si elle peut expliquer pourquoi un organisme peut développer des mécanismes pour créer des représentations, ne peut pas expliquer le contenu représentationnel spécifique de ces représentations ni la fonction individuelle de chaque représentation dans l'économie cognitive de l'organisme. En clair, on ne réussit pas, selon lui, à fixer le contenu des représentations par un appel à la téléologie. Par ailleurs, Dretske lui-même explique que si l'approche de Millikan peut être intéressante pour expliquer le développement des systèmes représentationnels, elle ne peut pas être utilisée pour déterminer pourquoi une créature peut utiliser une représentation ici et maintenant (Dretske, 2000). C'est pourquoi il a proposé sa solution ontogénétique présentée dans la section précédente.

Une autre critique célèbre face aux théories téléosémantiques consiste à imaginer un homme du marais (« *swampman* »). Imaginons un éclair frappant un marais et créant une réplique molécule-par-molécule d'un être humain. Cette copie serait absolument identique à un être humain tant du point de vue physique que du comportement. Toutefois, l'homme du marais n'aurait aucune histoire développementale ni évolutive, et aucun design (puisque l'éclair a frappé le marais par chance). En suivant les théories téléosémantiques, l'homme du marais ne peut pas avoir de fonction propre ni de critères de réussite ou d'échec puisque les fonctions sont soit données par un concepteur, soit par l'évolution. Ainsi, comme les états de l'homme du marais n'ont aucune fonction (puisque leur création vient de la chance), ces

états ne peuvent pas être des représentations selon la théorie téléosémantique. Cette critique a été initialement développée par Davidson (1987) pour tenter d'anticiper des objections à son explication historique du contenu représentationnel des représentations, mais elle a été reprise plus généralement pour argumenter que les explications téléologiques ne sont pas adéquates pour rendre compte de la fixation du contenu représentationnel des représentations<sup>18</sup>.

Enfin, la solution évolutionniste de Millikan est également critiquée comme étant insuffisante pour fonder une notion de représentation. Un argument qui revient souvent dans la littérature est que l'explication de Millikan confond le succès d'une fonction biologique, dans une optique phylogénétique, et l'exactitude<sup>19</sup> d'une représentation produite par un appareil cognitif (Burge, 2010; Stich, 1990). Burge précise que la détection (exacte) d'objets externes ne peut pas être sélectionnée par la sélection naturelle puisqu'il ne s'agit pas d'une fonction biologique. Ce qui est sélectionné selon Burge, c'est la capacité du détecteur à générer une réponse favorisant la survie et la reproduction. Prenons l'exemple d'un détecteur sensoriel qui aurait comme fonction de détecter les prédateurs dans le but de générer une réponse spécifique (la fuite). Même si le détecteur s'active parce qu'il y a une erreur dans l'identification du prédateur, on ne pourrait pas dire que le détecteur échoue à sa tâche, puisque sa fonction est justement de s'assurer que l'animal puisse prendre la fuite en cas de danger. Ainsi, on ne pourrait pas expliquer la mé-représentation du prédateur par un échec de la fonction du détecteur. Selon Burge: « There is (...) a root mismatch between representational error and failure of biological function. » (2010, p. 303)

Si la théorie de Millikan est critiquée, la théorie de Dretske l'est aussi. William Ramsey, dans *Representation Reconsidered* (2007), présente une longue critique des travaux de Dretske. Si Ramsey pense que Dretske met le doigt sur le bon problème (comment expliquer le passage de l'information pure à la représentation), il n'est pas en accord avec sa solution.

Tout d'abord, il est important de comprendre que Ramsey formule le problème de la naturalisation de l'intentionnalité représentationnelle d'une manière bien particulière: ce qu'il a nommé le « défi de la description de tâche » (*job's description challenge*). L'idée de description de tâche est assez simple: on détermine a priori ce qu'une représentation doit accomplir au sein de toute théorie cherchant à expliquer la cognition, et l'on vérifie empiriquement, pour une théorie représentationnaliste de la cognition donnée, si les entités théoriques nommées « représentation » par cette théorie répondent à ces critères.

---

<sup>18</sup> Neander (1996) et plus récemment Papineau (2022) répondent de manière élaborée aux différentes objections, mais le détail de leur réponse dépasse le cadre du mémoire. L'important est de saisir que les solutions téléosémantiques, si elles apportent une solution possible au problème de la naturalisation de l'intentionnalité, ne font pas l'unanimité.

<sup>19</sup> Dans son texte, Burge parle de *vérité* des représentations. Toutefois, je pense que le terme « vérité » est utilisé au sens d'exactitude comme défini dans la terminologie de Rolla et Huffermann (2022).

Sprevak (2011), qui résume la proposition de Ramsey, met en lumière trois aspects de la description de tâche de la représentation mentale: 1) l'intentionnalité originelle (la représentation doit être dirigée vers des choses autres qu'elle-même sans l'intermédiaire d'un interprète), 2) la causalité (la représentation doit interagir causalement avec d'autres états cognitifs) et enfin 3) l'efficience causale en tant que contenu (le pouvoir causal de la représentation doit dépendre de son contenu représentationnel). Autrement dit, une entité au sein d'un système cognitif ne peut être une représentation mentale que si elle accomplit sa fonction parce qu'elle est une représentation mentale, c'est-à-dire parce qu'elle respecte les trois conditions de la description de tâche.

Selon Ramsey, on doit distinguer le problème de la détermination de la fonction représentationnelle d'un état de celui de la détermination de son contenu représentationnel (même s'il y a certainement des chevauchements entre les deux problèmes). Selon lui, les philosophes et théoriciens font souvent l'erreur de présumer implicitement que la détermination du contenu d'une représentation explique ipso facto la manière dont cette représentation peut acquérir la *fonction* d'être une représentation de prime abord. La critique principale de Ramsey à l'égard de Dretske concerne justement l'explication fonctionnelle de Dretske. Pour Ramsey, le simple fait qu'une relation informationnelle soit utilisée par un système ne fait pas automatiquement de ce système un système représentationnel. Pour Ramsey, lorsque Dretske cherche à fonder la notion de représentation sur la notion d'information, il commet l'erreur suivante:

Dretske's account of representation appears to assume that if a given structure is incorporated into a system's processing because it nomically depends on a certain state of affairs, it automatically follows that it is being used to stand for (or relay information about) that state of affairs. (Ramsey, 2007, p. 134)

L'argument de Ramsey se base sur l'interprétation sémantique (ou réaliste) de l'information, c'est-à-dire que si A dépend de relations nomiques avec B, A porte un contenu informationnel à propos de B<sup>20</sup>. Selon Ramsey, Dretske échouerait à expliquer et défendre le saut entre les relations nomiques et l'information sémantique, et surtout il échouerait à expliquer le rôle causal spécifique de l'information (et non pas des relations nomiques). Ramsey prend l'exemple du thermostat présenté par Dretske comme un système représentationnel artificiel. Le thermostat a comme fonction de maintenir une température stable et utilise une tige métallique qui se dilate plus ou moins selon la température pour ce faire. La question n'est pas tant de savoir s'il y a une intentionnalité originelle au sein du thermostat (il est entendu que non, étant donné que le thermostat est un artéfact dont la fonction est attribuée par un concepteur). La question est plutôt de savoir si la tige métallique a une efficience causale en tant que possédant un contenu représentationnel spécifique (la troisième condition décrite par Sprevak, 2011), et Ramsey argumente que l'explication de Dretske ne permet pas de respecter cette condition. Ramsey soutient que

---

<sup>20</sup> Ramsey s'intéresse également à ce qu'il nomme l'interprétation *déflationniste* de l'information selon laquelle l'information n'est que l'utilisation de relations nomiques. Pour les besoins du mémoire, l'argument de Ramsey est sensiblement le même pour les interprétations sémantiques et déflationnistes de l'information.

Dretske ne réussit pas à montrer en quoi la partie métallique du thermostat (qui réagit à la température de la pièce) a véritablement un rôle représentationnel dans l'économie fonctionnelle du thermostat au-delà d'être un simple répondeur fiable qui cause mécaniquement l'ajustement de l'appareil. La tige métallique fait office de médiateur causal fiable entre la température et le thermostat, comme le ferait la gâchette d'un fusil. Si on est incapable d'expliquer le rôle causal spécifique de l'information que la tige est censée porter de manière distincte de la relation nomique entre la tige et la température, il n'y aurait pas de sens à dire que la tige métallique est une représentation de quoi que ce soit (comme nous n'aurions pas de raison de dire que la gâchette d'un fusil « représente » le coup de feu. Ramsey ne prétend pas que l'explication est impossible, mais plutôt que les nombreux exemples donnés par Dretske pour tenter de décrire des « représentations minimales » ne sont, au final, pas des représentations.<sup>21</sup> Autrement dit, on ne réussit pas à relever le défi de la description de tâche.

Hutto et Myin, dans *Radicalizing Enactivism* (2013), critiquent également les multiples tentatives cognitivistes de naturalisation des représentations. Hutto et Myin se décrivent comme des énaïvistes radicaux, c'est-à-dire qu'ils refusent toute notion de représentation mentale. Hutto et Myin vont distinguer l'information-comme-covariance, une simple covariation entre A et B, et l'information-comme-contenu, l'idée selon laquelle l'information porte un contenu ayant des propriétés représentationnelles ou propositionnelles (des conditions d'exactitudes ou de vérité). Si Hutto & Myin n'ont aucun problème avec la notion d'information-comme-covariance, ils soutiennent que la covariance est insuffisante pour fonder quelque notion de contenu représentationnel ou propositionnel que ce soit, puisque le contenu représentationnel ne peut pas exister, selon eux, sans la présence d'un interprète. Cet argument amène Hutto & Myin à critiquer la solution téléosémantique de Dretske. En effet, ils argumentent que le simple fait qu'un organisme utilise des relations de covariance ne transforme pas ces relations de covariance en contenu représentationnel. Logiquement selon eux, s'il n'y a pas de contenu représentationnel dans les relations de covariance (puisque'ils refusent l'interprétation réaliste de l'information), aucun contenu représentationnel ne peut être extrait ou acquis de ces relations de covariance (ce qu'ils nomment le « *No acquired content principle* », p. 73).

De plus, les énaïvistes radicaux critiquent également l'explication évolutionniste du contenu représentationnel de Millikan. En effet, ils affirment que la stratégie de Millikan (déterminer le contenu représentationnel grâce à la fonction propre de la partie consommatrice des représentations) rencontre

---

<sup>21</sup> Il faut également faire attention à ne pas introduire par mégarde notre intentionnalité comme observateur. En effet, on pourrait dire que le thermostat représente la température puisqu'il a été conçu pour cela, et que le fabricant a utilisé les relations nomiques entre la température et la tige de métal pour lui *donner la fonction* d'indiquer la température au thermostat. Le problème de Ramsey avec cette méthode d'argumentation est qu'elle nécessite quand même la présence d'un observateur externe, et que *pour* le thermostat (qui fait ici office de système cognitif), il n'est pas clair pourquoi la tige métallique serait une représentation, même si la tige accomplit une certaine fonction au sein du thermostat.

le même problème que Dretske. On doit présupposer que le contenu informationnel existe indépendamment de l'interprète, qui peut le « consommer » pour guider son comportement, une voie que Hutto & Myin refusent d'emprunter et qui, selon eux, est sans issue pour les raisons mentionnées précédemment. Comme nous verrons plus tard, Hutto & Myin (2013; 2017) proposeront une version énonciviste de la téléosémantique de Millikan qu'ils nommeront téléosémiotique, c'est-à-dire une théorie qui reprend l'explication téléologique évolutionniste de Millikan, mais en la débarrassant de ses ambitions sémantiques.

#### 2.1.4. Conclusion partielle

Bref, nous avons vu que différentes tentatives de naturalisation de l'intentionnalité au travers des représentations n'ont pas engendré de solutions satisfaisantes et consensuelles. En effet, les différentes théories, que ce soit celles de Fodor, Dretske ou Millikan, ne réussissent pas à expliquer de manière naturelle comment certains états physiques peuvent avoir ou acquérir des propriétés intentionnelles. Soit l'explication est tout bonnement absente (lorsqu'ils échouent le défi de la description de tâche de Ramsey), ou encore ils présupposent l'existence du contenu intentionnel dont ils sont censés expliquer l'émergence.

Bien que cette courte présentation ne soit pas exhaustive, elle illustre bien l'attitude traditionnelle face au problème de la naturalisation de l'intentionnalité et permet de mettre en lumière plusieurs de ses aspects. D'une part, les projets de naturalisation de Dretske et Millikan se concentrent exclusivement sur la naturalisation de l'intentionnalité représentationnelle. D'autre part, cette naturalisation prend généralement la forme d'une réduction, c'est-à-dire que Dretske et Millikan cherchent à fonder les propriétés sémantiques des représentations grâce à des notions comme l'information (jugée indépendante de l'existence des observateurs) tout en comprenant la représentation comme l'affaire d'une fonction biologique. Enfin, les théories de Dretske et Millikan ont en commun l'intérêt pour une explication biologique des représentations, c'est-à-dire que tous deux cherchent à fournir une explication dans le contexte d'un organisme, voire du contexte plus large de l'évolution et de la sélection naturelle.

#### 2.2. La stratégie d'intégration de l'énoncivisme: de l'intentionnalité biologique à l'intentionnalité pragmatique

Après avoir présenté les stratégies téléosémantiques (et leur critique) pour résoudre le problème de la naturalisation de l'intentionnalité, je présenterai dans cette section la théorie énonciviste de la cognition ainsi que la stratégie des énoncivistes pour naturaliser l'intentionnalité de l'esprit. Comme indiqué dans l'introduction, celle-ci est à deux volets: 1) refuser l'identité intentionnalité-représentation en retournant au sens fondamental de directionnalité pour élaborer des notions d'intentionnalité non

représentationnelle, puis 2) proposer une intégration graduelle et émergente des différents types d'intentionnalité, de la plus simple à la plus complexe.

L'approche énative de la cognition a été formulée dans l'ouvrage majeur de Varela, Thompson et Rosch, *The Embodied Mind*, publié en 1991. L'objectif était de réformer les sciences cognitives en y (ré)intégrant l'étude du corps vécu (*lived body*) ainsi que l'expérience humaine, se réclamant directement des travaux de Merleau-Ponty (Varela *et al.*, 1991). Toutefois, trois décennies après la publication originale de *The Embodied Mind*, le terme « énavivisme » ne désigne plus une seule théorie, mais plutôt une famille de théories ayant repris différents éléments de l'ouvrage de Varela, Thompson et Rosch<sup>22</sup>.

Dans cette section, je me concentrerai principalement sur les écrits d'auteurs s'inscrivant dans la lignée de l'énavivisme « classique », notamment les travaux de F. Varela et Thompson, mais aussi ceux de E. Di Paolo. La principale raison de ce choix est que ces travaux permettent de mettre en lumière la stratégie de naturalisation de l'intentionnalité axée sur l'intégration des trois grands types de manière convaincante et détaillée. Dans un premier temps, je développerai les concepts fondamentaux de l'énavivisme. Je m'efforcerai alors à démontrer comment les concepts complémentaires d'autonomie et de *sense-making* permettent de rendre compte d'une intentionnalité biologique, la première étape de la naturalisation. Ensuite, je m'attarderai à la théorie de la vie sensorimotrice de Di Paolo *et al.* (2017) qui, j'avance, permet d'expliquer le passage de l'intentionnalité biologique à l'intentionnalité pragmatique (la deuxième étape de la naturalisation), notamment grâce au concept d'agentivité. Il ne restera alors « que » l'étape de la naturalisation de l'intentionnalité représentationnelle, problème ardu qui méritera son chapitre dédié et qui sera traité de manière subséquente.

### 2.2.1. Les concepts majeurs de l'énavivisme

L'énavivisme est une famille de théories de la cognition proposant que l'esprit soit le résultat de l'interaction dynamique entre le cerveau, le corps et l'environnement plutôt que le résultat de la représentation du monde externe par le cerveau (Thompson, 2007; Varela, Thompson & Rosch, 1991). L'approche énative cherche notamment à expliquer le rôle primordial du corps et de l'action dans la cognition (Di Paolo & Thompson, 2014). L'énavivisme est à mettre en lien avec un regroupement de théories communément réunies sous l'étiquette « Cognition 4E » (*Embodied, Embedded, Enactive* et

---

<sup>22</sup> Dans *Radicalizing Enactivism*, Hutto & Myin (2013) distinguent trois grandes branches : 1) l'énavivisme *sensorimoteur* suivant les travaux d'Alva Noë (2004), 2) l'énavivisme *autopoïétique* suivant le plus fidèlement les travaux originels de Varela *et al.* (1991) et enfin 3) l'énavivisme *radical* de Hutto & Myin (Hutto & Myin, 2013; 2017). La formulation est restée dans la littérature, mais je l'éviterai parce qu'elle n'est pas très productive et porte plus à confusion qu'autre chose. Pour un argumentaire plus détaillé, voir également Thompson (2018), qui juge assez sévèrement cette catégorisation de Hutto & Myin.

*Extended*, avec parfois l'addition d'un cinquième E pour *Ecological*; voir Ward *et al.*, 2017; Kiverstein & Rietveld, 2018; James, 2020)<sup>23</sup>.

L'énactivisme met l'accent sur le caractère biologique de l'esprit ou plutôt, comme Thompson l'explique, sur la « continuité entre la vie et l'esprit » (Thompson, 2007, p. ix), une idée connue dans la littérature sous le nom de *mind-life continuity thesis* (voir Kirchoff & Froese, 2017). Selon cette thèse, les propriétés de l'esprit découlent d'un enrichissement des propriétés auto-organisationnelles du vivant. Pour présenter les grandes lignes de l'énactivisme, nous nous concentrerons sur les concepts d'autopoïèse, d'autonomie, de *sense-making* et d'adaptativité.

L'autopoïèse (Di Paolo & Thompson, 2014; Maturana & Varela, 1980) est le concept fondamental de l'énactivisme. L'autopoïèse (littéralement « auto-crédation ») est une propriété structurelle d'un système physique qui survient lorsque les composantes du système: 1) participent de manière récursive dans les processus de création des composantes qui les ont produites et 2) établissent le réseau de composantes comme une unité dans l'espace où se déroulent les processus de production et qui se distingue de son milieu. Le système résultant forme un réseau distinct capable de se produire et de se maintenir en créant ses propres composantes. L'exemple paradigmatique prisé par les énéactivistes est celui de la cellule biologique. Chez la cellule, la membrane cellulaire est formée et entretenue grâce à des protéines fabriquées par certaines composantes de la cellule (comme les ribosomes), et en retour cette membrane permet la création d'un milieu interne permettant la continuation des processus la créant (la fabrication de protéines) tout en distinguant le système du milieu environnant.

Toutefois, la notion d'autopoïèse, se voulant lors de sa formulation originale une description strictement physique de l'organisation structurelle du vivant, s'est enrichie durant les dernières décennies, notamment en raison des influences phénoménologiques qui ont traversé l'énactivisme. En effet, lorsque Weber & Varela (2002) décrivent la notion d'autopoïèse, ils vont avancer que le système autopoïétique n'est pas simplement un réseau de processus auto-producteurs, mais qu'il est possible de parler de l'instauration d'un certain *point de vue* pour ainsi parler d'un organisme :

Autopoiesis proposes an understanding of the radical transition to the existence of an individual, a relation of an organism with it-self, and the origin of "concern" based on its ongoing self-produced identity. One could envisage the circularity metabolism-membrane entirely from the outside (this is what most biochemists do). But this is not to deny that there is, at the same time, the instauration of a point of view provided by the self-construction. (Weber & Varela, 2002, p. 116)

---

<sup>23</sup> « *Embodied* » signifie que la cognition est *incarnée*, c'est-à-dire qu'elle implique le corps de manière non triviale. « *Embedded* » signifie que la cognition est *enchâssée* dans le contexte environnemental. « *Extended* » (le plus controversé des termes) signifie que la cognition *s'étend littéralement* au-delà des limites du cerveau.

L'idée qu'une cellule puisse avoir un « point de vue » et des « besoins » (*concerns*) ne revient pas à dire que la cellule possède une expérience phénoménale équivalente à celle d'un être humain, ou même à celle d'un animal. Lorsque l'on parle de « point de vue », c'est pour indiquer qu'il est possible de considérer la cellule non seulement d'une façon strictement physique, mais également du point de vue interne de la cellule selon ses besoins. Cette manière de concevoir l'organisme selon son propre point de vue et ses propres besoins n'est pas entièrement nouvelle dans l'histoire de la biologie. En effet, comme expliqué dans le premier chapitre, les germes de cette conception du vivant se trouvaient déjà chez Kant, qui parlait des organismes vivants comme de « téléologies intrinsèques ». Cette idée a été poursuivie par Jonas sous la forme de l'identité métabolique.

C'est en s'inspirant de ces discussions téléologiques que des auteurs énaclivistes comme Di Paolo & Thompson (2014) vont introduire un nouveau concept pour expliquer l'organisation du vivant: l'*autonomie*. L'idée derrière le concept d'autonomie est de mieux formaliser l'aspect téléologique de l'autopoïèse tout en conservant ses propriétés organisationnelles. Ainsi, non seulement le système vivant est auto-organisateur, mais il crée ses propres règles d'existence, ses propres normes.

Un système autonome présente deux propriétés : la fermeture opérationnelle et la précarité. La fermeture opérationnelle reprend grosso modo les propriétés organisationnelles de l'autopoïèse, c'est-à-dire que la fermeture désigne le réseau récursif de processus qui créent une unité distincte de l'environnement. Un schéma, tiré de Di Paolo & Thompson (2014) permet d'illustrer le concept de fermeture opérationnelle:

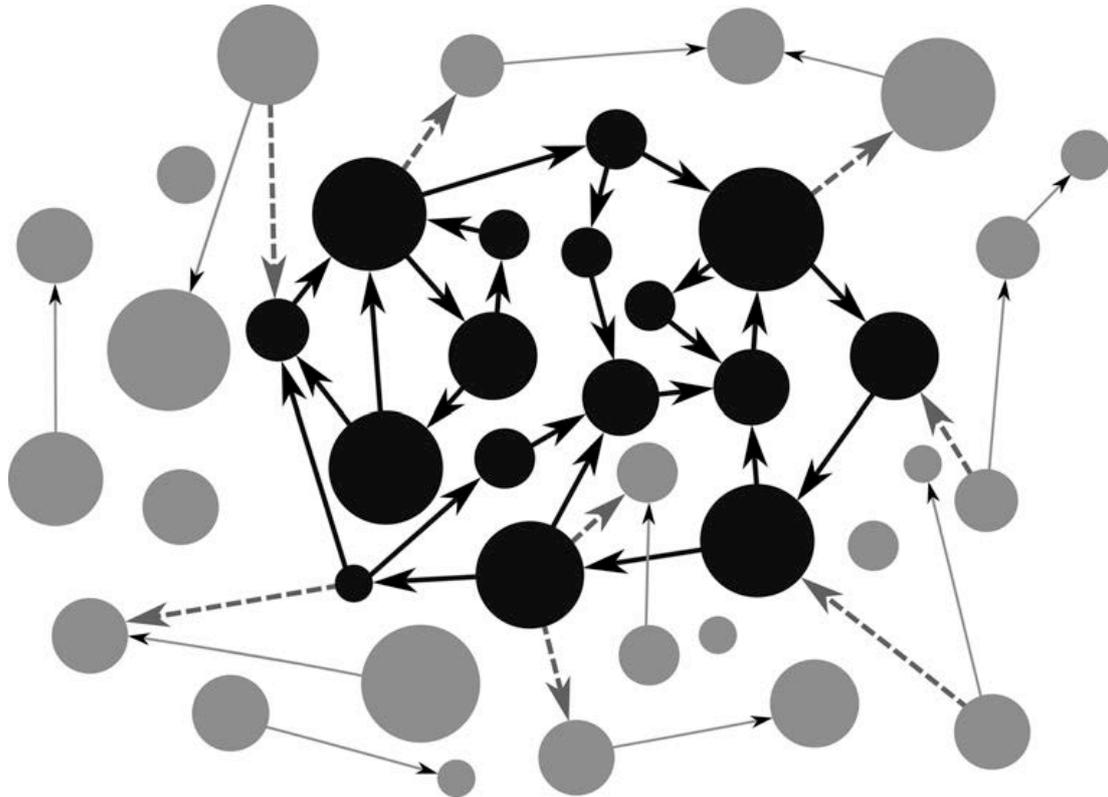


Figure 2.1 La fermeture opérationnelle (en noir). Les différents noyaux noirs sont connectés entre eux par des flèches signifiant l'inter-influence. Tiré de Di Paolo et Thompson (2014), p. 70

Ce schéma représente différentes composantes (les points) ainsi que leurs relations d'influence mutuelle (les flèches indiquant la direction de l'influence). Les composantes en noir font partie de la fermeture opérationnelle, puisque chaque élément dans le réseau reçoit une flèche (signe d'influence) et renvoie une flèche à un autre élément du réseau. Le réseau opérationnellement clos reprend ainsi les propriétés de l'autopoïèse: les composantes sont interdépendantes et elles forment une unité distincte de son milieu. Ce schéma permet également de clarifier le sens de « fermeture » : le schéma a pour but d'illustrer l'aspect récursif et interdépendant des composantes du réseau opérationnellement fermé. Un système « fermé » n'est pas absolument isolé de son environnement, au contraire. On peut le voir facilement grâce aux relations entrantes et sortantes du réseau illustré sur le schéma. Toutefois, les éléments externes ne font pas partie du réseau, puisque leur relation avec des membres du réseau est unidirectionnelle. Par exemple, la cellule qu'est un phytoplancton a clairement besoin du soleil pour exister, et l'extinction du soleil signifierait la mort de la cellule. Toutefois, le soleil est absolument indépendant de l'existence de la cellule, et c'est pourquoi cela n'aurait pas de sens de tenter d'intégrer le soleil dans le système « cellule ».

Néanmoins, pour rendre l'idée de fermeture opérationnelle non triviale, Di Paolo et Thompson ajoutent l'idée de précarité. Un réseau est dit précaire si ses processus ne peuvent se perpétuer sans le soutien

des éléments du réseau opérationnellement fermé (contrastant avec des processus non précaires qui peuvent se perpétuer avec le seul soutien des conditions externes). On peut comprendre que le réseau, comme unité distincte du milieu, est plus fragile que le milieu en question, et la rupture d'un ou de plusieurs processus du réseau opérationnellement clos peut mener à une cascade catastrophique de ruptures de processus internes, menant à la destruction du réseau et la dissolution de l'organisme au sein de son milieu (la mort). Ce concept de précarité rapproche également l'énaclivisme de la thermodynamique et rejoint les définitions du vivant de Schrödinger. En effet, comme il a été dit dans le chapitre précédent, la seconde loi de la thermodynamique affirme que l'entropie (le désordre ou la désorganisation des états) de tout système physique a tendance à augmenter. Autrement dit, le système vivant, qui est un système (auto-)organisé, est fragile puisque certaines interactions avec son environnement peuvent carrément le désorganiser (donc mener à sa perte) en interrompant le réseau récursif de processus biologiques.

C'est précisément la précarité de la fermeture opérationnelle, et donc ultimement les considérations thermodynamiques propres au vivant, qui permettent de comprendre le sens de l'identité d'un système autonome. En effet, l'identité d'un système autonome est dynamique, c'est-à-dire que pour maintenir son organisation, le réseau est constamment en train de se créer et de s'auto-organiser. C'est cette recherche constante d'auto-organisation, couplée à la notion de précarité, qui permet d'introduire le concept de *norme naturelle* dans le cadre énacliviste. Il est important de comprendre que le concept de norme est inséparable du concept d'identité (métabolique). En effet, comme le système autonome est à la fois auto-organisateur et précaire, ses interactions avec l'environnement ont une valeur pour l'organisme puisque certaines choses (externes) contribuent au maintien de l'identité du système tandis que d'autres y nuisent. Pour survivre, l'organisme doit pouvoir discriminer les différents éléments de son environnement en fonction de normes de viabilité que le système établit lui-même par sa propre existence (d'où l'expression d'auto-nomie, se donner ses propres normes). C'est là la simple poursuite de l'intuition de Kant quant à la téléologie intrinsèque. Comme l'organisme est sa propre fin, le maintien de son existence devient la norme ultime d'évaluation du monde externe, et les « conditions virtuelles » d'existence dont parlait Merleau-Ponty (1942) dans *La Structure du Comportement* établies par l'organisme sont maintenues par l'activité constante de ses différentes composantes et des différentes relations entre elles. Le moyen (le maintien des relations entre les composantes face à l'augmentation de l'entropie) se confond avec la fin (le maintien de l'existence du réseau autonome).

Évidemment, parler de normes n'aurait aucun sens si l'identité du système autonome s'apparentait à l'identité physico-chimique ( $A = A$ ) d'un système inorganique. En effet, parler de normes sous-entend que le non-respect de ces normes est à éviter. Bien qu'il soit toujours possible de parler mathématiquement des conditions « normales » d'existence d'un système non organique (la « norme » prise ici au sens statistique), parler d'une norme pour une tornade n'a pas du tout le même sens que pour

un système vivant. En effet, la tornade ne cherche pas à se maintenir, et une notion de norme de viabilité n'a aucun sens pour la tornade (en fait, la question ne se pose même pas). Toutefois, pour la cellule (ou tout système autonome), la question se pose évidemment, et il y a une différence capitale entre le maintien de l'existence et son non-maintien. Ainsi, les différences dans le milieu font une différence pour l'organisme (pour paraphraser l'expression de G. Bateson<sup>24</sup>), et c'est pourquoi les notions d'identité métabolique et de normes vont de pair : l'idée d'une identité métabolique sous-entend que les relations entre l'organisme et son environnement sont normées.

On voit ainsi pourquoi Di Paolo et Thompson (2014), sur les traces de Weber et Varela (2002), voient dans cette théorie énonciviste la possibilité d'une naturalisation non réductionniste de la téléologie intrinsèque proposée par Kant ainsi que de l'identité métabolique de Jonas (1966). En effet, Di Paolo et Thompson (2014) argumentent que la notion d'autonomie peut être vérifiée empiriquement, c'est-à-dire qu'il est possible pour un biologiste de vérifier si un système physique est opérationnellement clos et précaire. L'autonomie devient alors la description scientifique de l'organisme qui lutte activement pour assurer sa survie (le maintien de son organisation) selon des normes de viabilité qu'il établit lui-même par l'activité de son système opérationnellement clos et précaire (son phénotype).

### 2.2.2. Le *sense-making* comme intentionnalité biologique au sein de l'énoncivisme

La première étape de la stratégie énonciviste de naturalisation de l'intentionnalité consiste à expliquer dans des termes opératoires comment un être vivant acquiert une intentionnalité biologique telle que définie dans le premier chapitre. Si le concept d'autonomie est cardinal dans le cadre énonciviste, il reste du travail à accomplir afin de véritablement naturaliser l'intentionnalité biologique. Y'a-t-il quelque chose s'apparentant à une intentionnalité biologique dans le cadre énonciviste ? Dans cette section, j'avance que l'énoncivisme soutient l'existence d'une intentionnalité biologique et qu'il cherche à la naturaliser à l'aide du concept de *sense-making*.

La notion de *sense-making* fait référence à l'action de l'organisme qui instaure (naturellement) une perspective, faisant en sorte que ses interactions avec le monde gagnent une « signification vitale » (Di Paolo, 2005; Merleau-Ponty, 1942). Selon Weber & Varela (2002), le *sense-making* est complémentaire à l'autonomie. L'autonomie consiste en l'établissement d'une identité métabolique précaire que l'organisme cherche activement à préserver. Le maintien et l'activité constante des différents processus de production de l'organisme demandent du temps, de la matière et de l'énergie. Autrement dit, si l'organisme est autonome, il n'est pas autosuffisant (Di Paolo & Thompson, 2014).

---

<sup>24</sup> L'expression « *A difference that makes a difference* » provient de l'anthropologue américain Gregory Bateson (1979) qui l'utilisait pour désigner l'*information*.

Cette dépendance à l'environnement externe force l'organisme à interagir avec celui-ci, mais pas n'importe comment.

Le *sense-making* est la manière dont l'organisme agit dans l'environnement en fonction de ses propres normes de viabilité (définies par l'autonomie), transformant l'environnement physique en milieu. Plus exactement, le *sense-making* est une activité d'évaluation du milieu selon les normes de viabilité de l'organisme autonome. Les éléments du milieu peuvent être bénéfiques, nuisibles ou neutres, et l'organisme doit pouvoir discriminer ces éléments. C'est justement cette activité de *sense-making*, de « faire-sens », qui constitue l'aspect cognitif primordial du vivant et, réciproquement, l'aspect vital de la cognition. Thompson (2007) résume ainsi l'argument énonciste pour expliquer le lien entre le vivant et le *sense-making*:

1. Le vivant = autopoïèse et cognition. Tout système vivant est à la fois autopoïétique et cognitif.
2. L'autopoïèse implique l'émergence d'un soi incarné. Un système physique autopoïétique, en vertu de sa fermeture opérationnelle, produit et réalise un soi individuel sous la forme d'un corps vivant, l'organisme.
3. L'émergence d'un soi implique l'émergence d'un monde. L'émergence d'un soi est également nécessairement la co-émergence d'un domaine d'interaction propre à soi, c'est-à-dire un Umwelt.
4. L'émergence de soi et du monde = *sense-making*. L'environnement de l'organisme est constitué par le sens que l'organisme donne au monde. L'environnement est un lieu de signification et de valence résultant de l'action de l'organisme.
5. *Sense-making* = énonciste. Le *sense-making* est une conduite viable de l'organisme. Cette conduite est orientée vers l'environnement et sujette à sa signification et les valences<sup>25</sup>. Les valences ne sont pas « déjà là » (*out there*), mais sont énoncées et constituées par les êtres vivants. Vivre implique le *sense-making*, ce qui équivaut à l'énonciste. (Thompson, 2007, p. 158, nous traduisons)

Ainsi, le *sense-making* est la conséquence nécessaire de l'autonomie d'un système autopoïétique. Thompson (2004) va écrire explicitement que le « 'sense-making' is reminiscent of the phenomenological notion of intentionality » (p. 389-390), comprise non pas comme un « à propos » (*aboutness*) représentationnel, mais plutôt dans le sens d'une visée en vue d'une satisfaction ou d'un besoin (les « conditions de succès » décrites par Rolla et Huffermann, 2022). J'avance que cette interprétation rejoint le sens d'intentionnalité biologique développé au premier chapitre, et cela fait du *sense-making* la notion toute désignée pour naturaliser l'intentionnalité biologique. Comment s'y prendre?

---

<sup>25</sup> Ce passage peut sembler contradictoire avec ce qui a été dit précédemment, mais il est possible d'être charitable et d'interpréter ce passage comme voulant dire que la conduite est sujette à la signification et aux valences de l'environnement *telles qu'elles émergent* de la co-interaction entre l'organisme et l'environnement.

Si l'article séminal de Weber & Varela (2002) fait du *sense-making* une conséquence nécessaire de l'existence d'un système autonome, tous ne sont pas convaincus que l'on puisse passer automatiquement de l'autonomie au *sense-making* sans notion intermédiaire. Di Paolo (2005) cherche ainsi à pallier quelques manquements dans la théorie de l'autopoïèse telle que formulée par Varela et Maturana. C'est que l'autopoïèse, selon Di Paolo, est un concept tout ou rien (*all-or-none*): ou bien le système est opérationnellement clos, ou bien il ne l'est pas. Ainsi, même si on indique que le système est précaire, c'est-à-dire qu'il doit lutter pour sa survie, rien n'explique comment le système arrive à éviter les états potentiels menant à ou se dirigeant vers la désintégration. Autrement dit, l'autopoïèse varélienne ne permet qu'une théorie abstraite de la conservation de l'organisation, et non pas une théorie de l'homéostasie active (Di Paolo, 2005).

Comme l'explique Di Paolo (2005), Varela (1997) lui-même sous-entendait dans ses écrits que la signification (*meaning*) pour l'organisme résultait d'un manque, qu'elle avait comme origine les ruptures (*breakdowns*) de l'autopoïèse. Or, il est difficile d'envisager comment rendre compte d'un manque avec une logique binaire vivant/non-vivant. En effet, le manque est un état intermédiaire qui, s'il n'est pas réglé, peut mener à la rupture de certains processus et, à terme, à la désintégration de la fermeture opérationnelle. Le problème est que dans la théorie autopoïétique basique, toute notion de risque ou de manque ne peut être attribuée que par un observateur externe, c'est-à-dire qu'elle est la description de ce qui pourrait arriver si tel ou tel manque n'est pas réglé. Par exemple, la bactérie qui nage en suivant un gradient de sucre cherche à compenser un manque (celui des nutriments). Toutefois, en acceptant l'autopoïèse comme la simple conservation de la fermeture opérationnelle, tous les états externes comprenant du sucrose permettent la conservation, et la bactérie ne devrait pas distinguer les différentes concentrations. Or, dans la pratique, la bactérie va remonter le gradient de sucrose, ce qui suggère qu'elle peut discriminer les différentes concentrations. Cette discrimination implique que la bactérie possède un mécanisme de contrôle interne qui lui permet d'adapter son comportement aux états externes en dépassant la logique binaire conservation/non-conservation. La question est alors de savoir comment intégrer cette discrimination selon le point de vue de l'organisme (sans attribution par un observateur externe déjà doté de capacités intentionnelles). Ce serait la première étape de la stratégie de naturalisation de l'énactivisme axée sur l'intégration des différents types d'intentionnalités. Pour résoudre ce problème, Di Paolo va introduire une nouvelle notion pour expliquer rigoureusement le passage de l'autonomie au *sense-making*: l'*adaptativité*. L'adaptativité est définie comme suit:

A system's capacity, in some circumstances, to regulate its states and its relation to the environment with the result that, if the states are sufficiently close to the boundary of viability,

1. Tendencies are distinguished and acted upon depending on whether the states will approach or recede from the boundary and, as a consequence,

2. Tendencies of the first kind are moved closer to or transformed into tendencies of the second and so future states are prevented from reaching the boundary with an outward velocity. (Di Paolo, 2005, p. 438)

L'adaptativité désigne la possibilité pour l'organisme de changer son organisation interne tout en étant robuste aux perturbations externes (une nécessité considérant que le système doit être thermodynamiquement ouvert). De ce point de vue, le « manque » peut être rigoureusement défini comme une tendance de l'organisme à s'éloigner de la zone de viabilité, cette tendance pouvant être mesurée et évaluée en fonction de la quantité de ressources métaboliques et/ou de changements structuraux nécessaires pour compenser cette tendance. L'organisme autopoïétique adaptatif génère donc non seulement une norme par sa propre activité, mais possède des mécanismes 1) d'autosurveillance pour évaluer où l'organisme se situe par rapport à cette norme et 2) d'autorégulation pour apporter les changements nécessaires pour favoriser la conservation de l'organisme (Di Paolo, 2005). Un exemple répandu de ce genre de mécanisme d'adaptativité dans le monde unicellulaire pourrait être les différents canaux ioniques permettant le passage sélectif de certaines molécules à travers la membrane cellulaire. Les canaux sont à la fois sensibles à certaines molécules tout en permettant des modifications dans le réseau autopoïétique en absorbant ou en rejetant des molécules de la fermeture opérationnelle.

Pour Di Paolo, il est primordial d'articuler ouvertement la notion d'adaptativité pour procéder à la naturalisation rigoureuse du *sense-making*. En effet, l'organisme doit pouvoir avoir accès aux conséquences de ses rencontres au-delà de la logique binaire vie/mort de l'autopoïèse. Par exemple, le concept d'adaptativité permet également d'opérationnaliser des notions comme le risque, le stress, la maladie, etc., des concepts qui ne peuvent être représentés dans la logique binaire.

Le *sense-making* nécessite autant l'autopoïèse (ou l'autonomie) que l'adaptativité, autant une perspective établie par un réseau opérationnellement clos et précaire que des mécanismes de régulation lui permettant d'évaluer tant sa propre condition que le monde externe en fonction de cette perspective normative (Weber & Varela, 2002; Di Paolo, 2005). En effet, des mécanismes de régulation sans une norme générée par l'autonomie reviendraient à une formulation cybernétique du vivant, c'est-à-dire un système de contrôle et de régulation mécanique et mathématique. Comme l'indiquait Jonas (1966), une boucle de rétroaction, prise de manière isolée, ne possède ni but ni téléologie interne autrement que ceux attribués par un observateur externe. C'est à l'intérieur du réseau autopoïétique que la boucle de rétroaction (grâce aux mécanismes permettant sa régulation) peut acquérir une signification pour le réseau, pour l'organisme. Ce n'est pas pour indiquer que la cybernétique est inutile ou qu'elle fait fausse route, mais plutôt qu'elle ne répond pas réellement à la question qui occupe le mémoire, puisque toute intentionnalité présente dans un système purement cybernétique ne peut être attribuée a posteriori que par l'observateur (ou intégrée a priori par un concepteur).

Le travail des scientifiques, pour assurer la naturalisation, revient alors à vérifier empiriquement quels mécanismes peuvent satisfaire la « description de tâche » de l'adaptativité, pour reprendre la formule de

Ramsey (2007). Cette description de tâche comprend, minimalement, un mécanisme d'évaluation ainsi qu'un mécanisme de régulation. Ce que l'ajout de Di Paolo offre à la théorie autopoïétique est la possibilité d'une opérationnalisation du concept de *sense-making*, une construction de la signification basée sur la restructuration du réseau opérationnellement clos via l'adaptativité. Di Paolo (2005) soutient également que cet ajout permet à l'organisme de donner une signification à des éléments externes qui ne sont pas directement impliqués dans la fermeture opérationnelle, par exemple l'odeur de la nourriture ou les traces de pas d'une proie (qui agissent comme des *proxys* ou, pour reprendre le vocabulaire sémiotique, des signes<sup>26</sup>).

Par ailleurs, Sanneke de Haan (2020) argumente que la notion de *sense-making* ne devrait pas être comprise comme une forme de subjectivisme ou de constructivisme. L'organisme ne décide pas ce qui constitue son monde. Plutôt, les valences qui guident le *sense-making* de l'organisme relèvent avant tout de ce que de Haan (2020) nomme une réalité relationnelle. L'idée d'une « réalité relationnelle » soutient que l'organisme ne « construit » pas lui-même la signification (en opposition avec la formulation de Di Paolo, 2005), mais que celle-ci émerge de la relation que l'organisme entretient avec son monde environnant (De Haan, 2020; Thompson, 2007). Pour reprendre le fameux exemple de la cellule, le sucrose n'est pas « objectivement » de la nourriture, pas plus que la cellule « projette » un concept « nourriture » sur le sucrose: le sucrose devient de la nourriture pour la cellule à cause du lien entre, d'une part, les besoins de l'organisme dans la régulation de son autonomie et les caractéristiques du phénotype de la cellule (par exemple la présence de l'enzyme glucose oxydase, capable de briser la molécule de glucose en métabolite) et, d'autre part, les caractéristiques chimiques du sucrose. Autrement dit, le monde (*Umwelt*) est constitué de manière réciproque à l'établissement de l'identité de l'organisme, mais pas de manière délibérée de la part de l'organisme. Néanmoins, on peut comprendre que la stratégie de l'énactivisme pour naturaliser l'intentionnalité biologique consiste à décrire les mécanismes permettant le *sense-making* d'un organisme vivant, et la nécessité de l'activité de *sense-making* découle des propriétés fondamentales du vivant, à savoir l'autonomie supplémentée de l'adaptativité.

### 2.2.3. L'ur-intentionnalité: une alternative radicale au *sense-making* pour la naturalisation de l'intentionnalité biologique?

Si le *sense-making* est prépondérant dans les écrits énoncivistes et que la notion y joue un rôle crucial dans leur stratégie de naturalisation de l'intentionnalité biologique, certains auteurs se réclamant d'un

---

<sup>26</sup> L'énactivisme est parfois comparé avec un autre courant philosophico-scientifique, la *biosémiotique* (Favareau, 2010; Kull, 2011). L'idée derrière la biosémiotique est de prendre le *signe* comme unité centrale comme comprendre le vivant et son comportement. La littérature en philosophie de la biologie marque certaines ressemblances avec l'énactivisme, notamment l'importance accordée à la perspective de l'organisme

énactivisme radical (Hutto & Myin, 2013; 2017; Hutto & Satne, 2015) rejettent toute notion de contenu intentionnel (pour se distinguer des énaïvistes classiques comme Varela, Thompson et Di Paolo). Ils vont carrément rejeter l'idée selon laquelle le *sense-making*, dans sa formulation autopoïétique, est la forme la plus basique d'intentionnalité (biologique) que l'on retrouve dans la nature. Plutôt, ils vont proposer une notion alternative: l'*ur-intentionnalité*. L'*ur-intentionnalité* désigne un type d'intentionnalité sans contenu représentationnel ou propositionnel et pour laquelle la distinction sens-référence ne s'applique pas, mais qui implique tout de même une attitude vers un objet. Grosso modo, l'*ur-intentionnalité* rejoint l'idée d'une intentionnalité biologique telle que décrite dans ce mémoire, mais dans une version très sobre sans appel à la phénoménologie.

Pourquoi proposer une notion alternative? Les énaïvistes radicaux (Hutto & Myin, 2013; 2017) avancent que la notion de *sense-making*, en cherchant à la fois à parler de signification tout en rejetant le représentationnalisme, est ambiguë et difficile à opérationnaliser. Plus particulièrement, Hutto & Myin (2013) vont écrire que « talk of sense-making in describing the responses of simple living systems is misplaced and misleading » (p. 35). Selon eux, le problème principal de la notion de *sense-making* est que cette dernière est beaucoup trop anthropomorphique puisqu'elle implique des notions trop évoluées pour un organisme unicellulaire (Hutto & Myin, 2017). Les énaïvistes radicaux se distinguent de leurs homologues classiques en ce sens que l'influence de la phénoménologie et de la philosophie de la biologie (à la Merleau-Ponty et Jonas) y est beaucoup moins saillante, voire inexistante. Rappelons que l'intuition de Jonas derrière la téléologie intrinsèque du vivant est sa propre expérience vécue:

On the strength of the *immediate testimony of our bodies* we are able to say what no disembodied onlooker would have a cause for saying (...) – the point of life itself: its being self-centered individuality, being for itself (...) » (Jonas, 1966, p. 79, je souligne)

Cela amène Hutto & Myin à reprendre une critique formulée par de Jesus (2016) et Villalobos & Ward (2016), pour qui l'énaïvisme souffre d'un anthropomorphisme hérité de Hans Jonas. Après tout, comme le souligne Froese (2011b), la toute première phrase de *The Embodied Mind* (Varela, Thompson & Rosch, 1991) mentionne l'importance de l'expérience humaine pour le renouveau des sciences cognitives: « This book begins and ends with the conviction that the new sciences of mind need to enlarge their horizon to encompass both *lived human experience* and the possibilities for transformation inherent in human experience. » (Varela *et al.*, 1991, p. lxi, je souligne)

Selon de Jesus (2016), cet anthropomorphisme est problématique dans la mesure où il mettrait à mal les prétentions scientifiques de l'explication du vivant au-delà de l'humain, nuisant ainsi au projet de défendre la thèse de la continuité entre l'esprit et le vivant. Pour éviter ces problèmes (ainsi que la solution de de Jesus qui fait appel à la biosémiotique), Hutto & Myin préfèrent la formulation de Cappuccio & Froese (2014) qui indiquent que le *sense-making* devrait être compris simplement comme

un « pure know-how learned through unprincipled interaction with the world. [In this sense, basic] cognition is primarily a relational form of meaningful engagement » (p. 4). À première vue, il n'est pas si clair que cette définition ne rejoigne pas le sens autopoïétique du *sense-making*, mais Hutto & Myin proposent tout de même de reformuler le problème de l'intentionnalité biologique et de la signification des interactions de l'organisme avec son environnement en termes de réponses intentionnelles (*intentional responses*) sans contenu représentationnel, ce qu'ils nomment l'ur-intentionnalité.

La question cruciale demeure: comment naturaliser l'ur-intentionnalité (et, par le fait même, une forme d'intentionnalité biologique)? Pour Hutto & Satne (2015) et Hutto & Myin (2017), la réponse se trouve dans une théorie déjà abordée dans le mémoire: la téléosémantique. Toutefois, en bons anti-représentationalistes, Hutto, Myin et Satne cherchent justement à dégonfler la téléosémantique de ses ambitions sémantiques (soit d'expliquer naturellement comment un état peut acquérir un contenu représentationnel) pour ne garder que la téléologie (une approche également envisagée par Godfrey-Smith, 2006; voir également Thompson, 2018). Toutefois, comme expliqué précédemment, les principales théories téléosémantiques (Millikan, 1984; Dretske, 1988) échouent à naturaliser l'intentionnalité représentationnelle, notamment en raison de leur incapacité à traiter adéquatement du problème de la disjonction ou encore en raison des difficultés à attribuer un rôle causal aux représentations *en tant que* représentations (Ramsey, 2007). Comment alors Hutto & Myin entendent-ils utiliser ces théories?

Hutto & Myin vont réformer la téléosémantique (ou la RECtifier selon le terme des auteurs<sup>27</sup>) pour lui donner le nom de téléosémiotique (à ne pas confondre avec la biosémiotique<sup>28</sup>). Le but de Hutto et Myin est de fournir une explication du processus d'objectification (l'organisme peut viser des choses en tant qu'objets) dans des termes non représentationnels. Pour ce faire, Hutto & Myin s'attaquent à l'exemple fameux du problème de la disjonction de Fodor, les grenouilles qui déploient leur langue pour n'importe quelle chose ressemblant à une mouche (ex. les *BB flies*, soit les plombs des carabines à plomb). Hutto & Myin vont avancer que l'on peut déterminer l'objet intentionnel du comportement grâce à l'histoire évolutive du développement de ce comportement. Selon Hutto et Myin, il est possible d'affirmer que le comportement de la grenouille a véritablement comme objet la mouche (et donc, qu'il est intentionnellement dirigé vers elle), puisque c'est historiquement ce qui a permis à la grenouille de se nourrir adéquatement. Le comportement a ainsi la fonction d'attraper la mouche, ce qui le rend minimalement intentionnel au sens de l'ur-intentionnalité. Kiverstein & Rietveld (2015) vont nommer

---

<sup>27</sup> REC est un acronyme voulant dire Radically Enactive Cognition, et est le nom du programme de recherche esquissé dans *Evolving Enactivism* de Hutto & Myin (2017). Les auteurs vont ainsi multiplier les jeux de mots avec REC, en parlant de la RECtification comme la réforme radicale de concepts énaclivistes afin de se débarrasser de leurs problèmes. Le jugement quant au bon goût de ces jeux de mots est laissé à la discrétion du lecteur.

<sup>28</sup> Voir De Jesus, 2016 pour une tentative d'intégration de l'énaclivisme et de la biosémiotique.

ce genre de comportement des *targeted response tendencies*, c'est-à-dire des comportements ayant tendance à cibler (ou viser) certains éléments précis de l'environnement. Hutto & Myin conservent ainsi la dimension normative et téléologique de la téléosémantique, mais en délaissant la notion de contenu représentationnel (possédant des conditions d'exactitude). Selon Hutto et Myin, c'est justement cette dimension normative qui permet d'affirmer que l'ur-intentionnalité est belle et bien intentionnelle, et non pas simplement une disposition du genre de celles mises de l'avant par les behavioristes (une critique leur étant fréquemment adressée, voir Roy, 2015 ou Kiverstein & Rietveld, 2015 pour le détail de ces critiques). Rolla & Huffermann (2022) affirment que la notion d'ur-intentionnalité permet d'imaginer un contenu basique, qui ne dépend pas de conditions d'exactitude, mais plutôt de conditions de succès.

### 2.2.5. La naturalisation de l'intentionnalité biologique: un succès?

En bref, malgré certains désaccords des énaïvistes radicaux quant aux thèses énaïvistes, le fil d'Ariane guidant la naturalisation de l'intentionnalité biologique demeure un appel à la normativité des interactions biologiques entre l'organisme et son environnement. Toutefois, il est intéressant de constater que la source de cette normativité diffère selon les théories. En effet, pour les radicaux, la normativité est issue de l'histoire biologique des espèces (à la Millikan) tandis que chez les énaïvistes classiques, c'est la fermeture opérationnelle précaire qui est la source des normes de viabilité que l'organisme établit par lui-même (son autonomie).

Serait-il possible de combiner l'explication téléologique de l'ur-intentionnalité avec la notion de *sense-making* dérivant de l'autonomie des organismes? Thompson (2018) semble contre cette idée. En commentant l'ouvrage de Hutto & Myin (2017), Thompson affirme que « proponents of the enactive approach have long argued that the selectionist theory of teleological biological functions is insufficient for explaining the autonomy of living beings and biological functional norms » (Thompson, 2018). Thompson critique sévèrement Hutto & Myin pour leur omission de la notion d'autonomie, cruciale selon lui pour comprendre le cadre énaïviste. Selon Thompson, sans cette notion d'autonomie, les auteurs soi-disant « radicaux » se rapprochent plutôt des théories écologiques dynamiques de la cognition (comme Chemero, 2009) que de l'énaïvisme tout court.

Malgré ce rejet apparent de Thompson, une piste potentielle d'intégration des deux théories serait la notion *organisationnelle* de fonction biologique (Mossio, Saborido, et Moreno, 2009). Selon cette notion, un trait biologique a une fonction si 1) ce trait participe à l'organisation du système biologique 2) qu'il est produit sous des contraintes créées par le système biologique et que 3) le système biologique est différencié de son environnement. Cette idée est, à bien des égards, similaire à l'idée énaïviste de fonder la normativité biologique grâce à l'autonomie (une fermeture opérationnelle précaire). Là où la

notion organisationnelle devient intéressante, c'est lorsque Mossio et ses collègues avancent qu'elle permet de combiner les approches dispositionnelles de la fonction biologique (Bigelow & Pargetter, 1987) qui peuvent expliquer la normativité de l'organisation actuelle de l'organisme (ses traits biologiques et les processus qui leur permettent de fonctionner) avec les approches étiologiques de la fonction biologique (Millikan, 1989; Godfrey-Smith, 1994) qui peuvent quant à elle expliquer les conditions d'existence (évolutives) de ces traits. Autrement dit, si on réussit à démontrer que la notion organisationnelle de fonction est compatible tant avec l'énactivisme radical qu'avec l'énactivisme, on pourrait possiblement avoir droit à une piste d'intégration d'une dimension téléosémiotique dans l'énactivisme autopoïétique. Toutefois, la réussite de cette entreprise n'est pas garantie, et pour des soucis d'espace, je n'irai pas plus loin ici.

Cela dit, compte tenu de l'état actuel du débat concernant l'intentionnalité dans l'énactivisme, je souscris à l'idée que la naturalisation de l'intentionnalité biologique par le *sense-making* offre un plus grand pouvoir explicatif que celui de l'ur-intentionnalité. La critique principale qu'il est possible d'adresser au concept d'ur-intentionnalité concerne justement son pouvoir explicatif. Si Hutto & Myin indiquent qu'il y a une normativité inhérente à la fonction biologique de l'organisme, il n'est pas si clair en quoi cette normativité participe aux processus d'objectivation qu'ils cherchent tant à naturaliser. Cette critique est notamment soulevée par Steiner (2019), qui estime que les théories énaclivistes sont coupables d'un pan-intentionnalisme (l'idée qu'il y ait de l'intentionnalité presque partout) qui rendrait triviale la notion même d'intentionnalité. Steiner avance que tant l'énactivisme radical que l'énactivisme soutiennent que l'intentionnalité est présente dans tous les organismes vivants. Toutefois, cela ne semble pas être vrai des énaclivistes, la normativité du *sense-making* étant l'affaire d'une opérationnalisation minutieuse d'une téléologie intrinsèque. En formulant le concept d'autonomie et en lui donnant une définition opérationnalisable (une fermeture opérationnelle précaire), la normativité des interactions est explicable par la tendance au maintien de l'intégrité de l'organisme, une activité effectuée par l'organisme lui-même.

Pour toutes ces raisons, le cadre énacliviste radical, s'il offre une alternative intéressante en reprenant le projet de la téléosémantique, ne réussit pas à véritablement naturaliser l'intentionnalité biologique de la même manière que le fait l'énactivisme classique, ou du moins, pas avec le même pouvoir explicatif. Je soutiens donc que le cadre énacliviste classique réussit le mieux à fournir un cadre naturaliste pour expliquer comment quelque chose comme une intentionnalité biologique peut exister dans le monde naturel. La clé de la naturalisation de l'intentionnalité biologique par l'énactivisme se trouve dans l'opérationnalisation d'une normativité issue d'une téléologie intrinsèque où l'organisme devient à la fois sa fin et son moyen. À cet effet, l'agencement entre les concepts d'autonomie, d'adaptativité et de *sense-making* semble fournir la base pour la stratégie d'intégration des trois types d'intentionnalité par l'énactivisme. Pour l'énactivisme, l'intentionnalité biologique prend la forme du *sense-making*,

l'activité de l'organisme donnant un sens à l'environnement pour le transformer en milieu. Ce *sense-making* est la conséquence nécessaire de l'autonomie, la fermeture opérationnelle précaire qui distingue le vivant du non-vivant, et s'exécute entre autres grâce à l'adaptativité qui permet d'effectuer une homéostasie active.

## 2.2.6. L'émergence de l'agentivité: de l'intentionnalité biologique à l'intentionnalité pragmatique

Après avoir expliqué comment l'énactivisme entend naturaliser l'intentionnalité biologique, tout est en place pour passer au deuxième type d'intentionnalité: l'intentionnalité pragmatique. Dans les sections précédentes, j'ai abordé les concepts fondamentaux de l'énactivisme. Parmi ces concepts se trouvaient l'autonomie et le *sense-making*. J'ai décrit le *sense-making* comme étant l'établissement d'une perspective de l'organisme, c'est-à-dire que l'environnement externe, en raison de l'autonomie et des normes de viabilité créées par la fermeture opérationnelle précaire, gagne une signification pour l'organisme et devient ainsi un monde de valences, celles-ci étant positives, négatives ou neutres. J'ai également expliqué que le *sense-making* est une activité, c'est-à-dire que c'était par l'action que l'organisme donnait une signification au monde. Toutefois, cette activité sous-entend également une autre notion, celle de l'*agentivité*.

La notion d'agentivité désigne généralement la manifestation de la capacité d'agir (Schlosser, 2015). Ainsi, un *agent* est un être qui possède cette capacité d'agir. Barandiaran, Di Paolo et Rohde (2009) définissent l'agent (par exemple, les agents biologiques comme ceux qui nous intéressent ici) comme étant des systèmes qui accomplissent quelque chose par eux-mêmes en fonction d'un but ou d'une norme dans un environnement spécifique. Barandiaran, Di Paolo et Rohde identifient trois caractéristiques centrales de l'agentivité: l'individualité, la normativité et l'interaction asymétrique. L'individualité reprend la description faite précédemment de la fermeture opérationnelle: un réseau interdépendant de composantes qui se distingue de son environnement. Di Paolo *et al.* (2017) précisent qu'en vertu de cette définition, un agent est nécessairement un système autonome (opérationnellement clos et précaire). Le concept d'interaction asymétrique, quant à lui, cherche à opérationnaliser l'idée intuitive qu'un agent est en un certain sens la source d'un changement, qui devient une action. Pour les énéactivistes, il y a un couplage structurel entre l'agent et son environnement, mais celui-ci est asymétrique puisque c'est l'agent qui module (modifie des paramètres de) son couplage avec l'environnement et non l'inverse et c'est en ce sens que l'organisme est la source des actions<sup>29</sup>. Enfin, cette modulation est dite normative

---

<sup>29</sup> Barandiaran, Di Paolo & Rohde (2009) offrent une description mathématique du couplage symétrique ainsi que du couplage asymétrique:

$$dS/dt = FQ(S, E) \quad (1)$$

$$dE/dt = GQ(S, E) \quad (2)$$

$$\Delta p = HT(S) \quad p \subset Q, \quad (3)$$

puisqu'elle se produit en vue d'un certain but établi par le maintien de l'autonomie de l'organisme. En résumé, un agent (biologique) est un système autonome qui module de manière normative son interaction avec son environnement.

Au-delà de l'agentivité biologique décrite précédemment, la notion d'agentivité est particulièrement utile pour faire le pont entre le niveau biologique et le niveau sensorimoteur de l'organisme, comme l'explique Di Paolo *et al.* (2017) dans le passage suivant:

These tools [autonomy and sense-making] can help us make sense of how biological agents, organisms that regulate their interaction with their environment, can sometimes also become sensorimotor agents, forms of life that are constituted as self-sustaining, habitual organizations in the structural and functional interrelations between their acts, skills, and dispositions. (Di Paolo, Buhrmann et Barandiaran, 2017, p. 7)

Comme l'indique Di Paolo (2005), « cognition requires a natural centre of activity on the world as well a natural perspective on it » (Di Paolo, 2005, p. 443). L'intérêt de la notion d'agentivité est qu'elle permet de faire le pont entre l'intentionnalité biologique et l'intentionnalité pragmatique en expliquant l'émergence d'une agentivité de plus en plus complexe, de l'activité organismique à l'organisme sensorimoteur. En effet, tel qu'illustré dans le premier chapitre, la catégorie d'intentionnalité pragmatique regroupe des notions comme l'intentionnalité motrice de Merleau-Ponty, laquelle implique un système sensorimoteur, et l'intérêt de la théorie énaïve est qu'elle permet justement d'expliquer le fonctionnement (et l'émergence) de ce système sensorimoteur en utilisant les mêmes ressources théoriques que celles déployées pour expliquer comment le vivant se distingue du non-vivant et comment est-il possible d'affirmer que l'intentionnalité (biologique) apparaît en même temps que le vivant.

Pour adapter la notion d'agent (qui implique l'autonomie) au niveau sensorimoteur, l'ouvrage *Sensorimotor Life* de Di Paolo *et al.* (2017) introduit la notion de contingence sensorimotrice (Beer, 2003; O'Regan & Noë, 2001). Une contingence sensorimotrice désigne la relation (régulière) entre les sensations et les mouvements d'un organisme. Formellement, en acceptant que les sensations et les mouvements moteurs soient deux variables indépendantes, il y a contingence sensorimotrice lorsque la relation entre les variations des variables sera corrélée et régulière, et ce indépendamment de la valeur des variables. Par exemple, chez l'être humain qui possède une vision stéréoscopique et binoculaire, reculer sa tête fait en sorte que le paysage s'agrandit, et s'avancer fait en sorte que le paysage rétrécit.

---

où (1) et (2) représentent le couplage symétrique en prenant comme référence respectivement l'organisme (S) et l'environnement (E). (3) représente le couplage asymétrique entre l'organisme et l'environnement.

Cette relation est maintenue dans des conditions normales indépendamment de la position de la tête ou de l'environnement dans lequel on se trouve.

Di Paolo *et al.* (2017) enrichissent la notion traditionnelle de contingence sensorimotrice (telle qu'utilisée par Noë, 2006, par exemple) en expliquant qu'il existe en réalité quatre sortes de contingences sensorimotrices, de la plus générale à la plus spécifique: l'environnement sensorimoteur, l'habitat sensorimoteur, la coordination sensorimotrice et le schème sensorimoteur. La définition de l'environnement sensorimoteur reprend grosso modo celle de la contingence sensorimotrice, c'est-à-dire qu'il s'agit d'une relation entre le mouvement (la variable indépendante) et les sensations (la variable dépendante). Di Paolo *et al.* (2017) précisent qu'un environnement sensorimoteur donné est spécifique aux agents qui possèdent un corps similaire. Le terme habitat sensorimoteur désigne l'ensemble des trajectoires sensorimotrices (i.e. les mouvements dans l'espace sensorimoteur) qui peuvent être générées par un système couplé avec son environnement (Di Paolo *et al.*, 2017). Concrètement, cela veut dire que l'on prend également en compte la position initiale de l'agent, sa position actuelle ainsi que les états internes de l'agent pour déterminer la relation entre les mouvements et les sensations. Mais toutes les trajectoires qui constituent un habitat sensorimoteur ne sont pas fonctionnelles pour un agent. Le terme coordination sensorimotrice désigne ces trajectoires sensorimotrices qui contribuent fonctionnellement aux buts de l'action d'un agent (Di Paolo *et al.*, 2017). Di Paolo *et al.* (2017) donnent cet exemple: pour déterminer la douceur d'une éponge, on presse celle-ci entre ces doigts, et la douceur est déterminée par une corrélation entre la pression appliquée et la résistance ressentie. Plus il sera facile de peser, plus l'agent déterminera que l'éponge est « douce ». Cette contingence sensorimotrice contribue fonctionnellement à la tâche de catégorisation de l'agent et elle est donc, selon la typologie de Di Paolo *et al.* (2017), une coordination sensorimotrice. Enfin, la notion de schème sensorimoteur désigne un ensemble de patrons de coordinations sensorimotrices souvent réalisés (*enacted*) ensembles (en même temps, en séquence ou dans un cycle) parce que, au fil des essais de l'agent en vue de la réalisation d'un but précis (par ex.: boire), des coordinations sensorimotrices se sont auto-organisées par assimilations et accommodations successives en une organisation (un schème) qui permet à l'agent d'accomplir son but de manière optimale. La notion de schème sensorimoteur comporte une composante normative, qui est déterminée par l'autonomie de l'agent: un schème peut-être plus ou moins efficace (ou « *successful* ») pour accomplir une action donnée.

Certains schèmes sensorimoteurs peuvent être interdépendants au point où un réseau précaire de schèmes peut s'individualiser en une capacité sensorimotrice plus complexe. Piaget (1936, 1947) parlait du phénomène d'équilibration d'un schème moteur, qui comportait une tension entre l'assimilation (l'intégration d'un élément externe dans un schème sensorimoteur) et l'accommodation (la modulation du schème sensorimoteur en réponse à l'environnement extérieur). Pour illustrer ces phénomènes, Di Paolo *et al.* (2017) reprennent un exemple très simple de Piaget (1936): l'alimentation d'un bébé par

biberon. Pour le bébé, l'acte de s'alimenter est constitué de trois coordinations sensorimotrices: la succion, la déglutition et la respiration. La succion implique à la fois que le bébé s'adapte au biberon, mais que le biberon soit également assimilé dans la coordination sensorimotrice de la bouche du bébé (autrement dit, le bébé ne peut pas sucer dans le vide et l'action a un objet). Si la succion est réussie, du lait se retrouve maintenant dans la bouche, et le lait sera assimilé dans la deuxième coordination sensorimotrice: la déglutition. Enfin, une troisième coordination sensorimotrice doit interagir avec les deux précédents: la respiration. En effet, pour que les deux coordinations précédentes puissent avoir lieu, il faut que le bébé retienne son souffle pour éviter que la trachée ne s'ouvre en même temps que l'œsophage. Ainsi, même une action aussi simple que l'alimentation d'un bébé nécessite l'organisation de plusieurs coordinations sensorimotrices en un schème sensorimoteur. Cette organisation en schème, fermée sur elle-même (fermeture opérationnelle), s'individue en ce sens que l'action de s'alimenter dépend de plusieurs composantes (les trois coordinations sensorimotrices), et que l'ensemble est précaire: l'échec d'une de ces trois coordinations peut mener à la disruption du schème principal.

Par ailleurs, un schème sensorimoteur qui est réalisé de manière répétée peut devenir une habitude, et cette habitude est précaire en ce sens qu'elle nécessite le bon maintien des différentes structures dont elle dépend, par exemple les muscles et le système nerveux (Di Paolo *et al.*, 2017). Ainsi, une habitude n'est pas un réflexe conditionné, mais une structure qui évolue dynamiquement et normativement en fonction des buts et objectifs de l'organisme. Certains schèmes peuvent céder le pas à d'autres (comme lorsque l'enfant passe du biberon au verre comme manière de d'abreuver), permettant une adaptation aux conditions réelles d'existence et de maintien de l'organisme sensorimoteur.

Pour Di Paolo, Buhrmann et Barandiaran, il est possible de parler d'agentivité sensorimotrice puisque même au niveau sensorimoteur, on peut retrouver les critères d'individuation, d'interaction asymétrique et de normativité. En effet, dans un schème sensorimoteur, les différentes coordinations sensorimotrices sont organisées, sélectionnées, développées et évaluées en fonction d'un but ou d'un objectif. Par exemple, prendre un objet dans ses mains nécessite l'organisation de multiples coordinations sensorimotrices pour s'approcher de l'objet, le serrer avec la bonne force, l'approcher de soi, etc. Souvent, il est possible d'évaluer les schèmes sensorimoteurs en fonction d'une norme biologique (Barandiaran et Moreno, 2008), c'est-à-dire que certains schèmes sont plus utiles que d'autres pour la survie. Toutefois, une certaine normativité sensorimotrice émerge et dépasse le cadre strictement biologique. Par exemple, il est possible d'échouer à la réalisation de l'action (par exemple, un bébé pourrait échouer l'action de se nourrir du sein de sa mère en aspirant trop de lait et devant en recracher une partie), sans que cet échec n'affecte directement le fonctionnement biologique de l'organisme. On dira alors de tels schèmes sensorimoteurs qu'ils sont mieux (ou moins bien) réussis, plus (ou moins) efficaces, voire plus (ou moins) élégants ou harmonieux. D'autres habitudes peuvent également être créées et maintenues, sans avoir de réel impact sur l'autonomie biologique de l'organisme. Les schèmes

sensorimoteurs peuvent ainsi gagner une « autonomie » qui leur est propre et qui, si elle dépend en partie de l'autonomie biologique, ne s'y réduit pas puisque de nouvelles normes peuvent apparaître au niveau des schèmes sensorimoteurs eux-mêmes.

L'enjeu consiste maintenant à déterminer si le passage de l'autonomie biologique à l'autonomie sensorimotrice nous permet d'assurer le passage de l'intentionnalité biologique à l'intentionnalité pragmatique. Je soutiens que oui en argumentant que la description de Di Paolo *et al.* (2017) est intentionnelle (au sens pragmatique) à deux niveaux: au niveau des schèmes sensorimoteurs eux-mêmes (qui prennent quelque chose comme objet) ainsi qu'au niveau de l'agent sensorimoteur.

D'une part, la notion de schème sensorimoteur décrite par l'énactivisme est toujours située et cette ouverture est dirigée. En effet, le schème est toujours considéré dans l'environnement dans lequel il s'établit et s'individue. De la même manière que la notion de *sense-making* dépend justement du milieu ambiant pour avoir un sens, nous avons vu que le schème sensorimoteur dépend également des éléments externes pour s'individualiser. Un schème sensorimoteur se doit donc d'être ouvert au monde ambiant pour pouvoir se former convenablement. Cette ouverture du schème sensorimoteur est également dirigée spécifiquement. En effet, les schèmes sensorimoteurs peuvent s'individualiser parce qu'ils réussissent à assimiler certains éléments extérieurs (comme la succion avec la tétine du biberon). Plus encore, cette assimilation a toujours un certain degré de spécificité, c'est-à-dire que le schème sensorimoteur va cibler certains éléments physiques plutôt que d'autres. Par exemple, dans le schème sensorimoteur de l'alimentation par biberon, la coordination de la succion cible la tétine, tandis que la coordination de la déglutition cible le lait, et la respiration cible l'air. Cette spécificité peut avoir différents degrés de généralités, c'est-à-dire que la coordination sensorimotrice de la succion peut cibler une tétine, mais également n'importe quel objet présentant les mêmes caractéristiques physiques, et ce même si un schème sensorimoteur pourrait créer une habitude avec un objet spécifique (par exemple ce biberon) qui nécessiterait une certaine accommodation si l'objet était remplacé par un autre objet (par exemple, le passage du sein de la mère au biberon en caoutchouc). De plus l'objet est toujours ciblé sous un certain aspect. Par exemple, le lait est l'objet de la déglutition parce qu'il est liquide, et non pas parce qu'il est blanc. En ce sens, la déglutition cible principalement la viscosité des objets assimilés dans cette coordination sensorimotrice. Ainsi, le schème de la déglutition développé pour téter le lait aurait besoin d'une bien moins grande accommodation pour du jus d'orange que, par exemple, pour un yogourt ou encore de la nourriture solide. « Pour » le schème sensorimoteur, le lait et le jus d'orange sont quasi identiques sous l'aspect de la viscosité, tandis que le lait et le yogourt sont assez différents.

La logique est la même au niveau de l'agent sensorimoteur. En effet, comme l'autonomie sensorimotrice est l'affaire d'un réseau précaire de schèmes sensorimoteurs, il n'est pas rare que certains schèmes sensorimoteurs ciblent d'autres schèmes sensorimoteurs, ou encore que les schèmes soient enchâssés les

uns dans les autres de manière hiérarchique. Par exemple, jouer au hockey nécessite la coordination de plusieurs schèmes sensorimoteurs interdépendants. En effet, patiner nécessite plusieurs schèmes comme accélérer, freiner, tourner, etc. Ce schème de patiner est également associé à d'autres schèmes comme celui de tenir le bâton, de manier la rondelle avec le bâton, de faire une passe, etc. Les schèmes sont interdépendants, et par exemple faire une passe cible non seulement la rondelle ainsi que le coéquipier, mais également le schème du patinage qui permet de se placer exactement au bon angle pour faire la passe et ne pas perdre la rondelle. Tous ces schèmes ciblent spécifiquement des éléments externes.

D'autre part, les schèmes sensorimoteurs ont une dimension normative, et ce à plusieurs niveaux. En effet, chaque schème peut réussir ou échouer à différents degrés. Comme décrit précédemment, la réussite d'un schème peut être normée en vertu de fonctions biologiques (comme l'alimentation), mais ces schèmes peuvent également avoir une normativité d'un niveau supérieur et indépendant des normes biologiques. Il y a une norme interne au schème sensorimoteur qui ne dépend pas de fonctions biologiques directement. Par exemple, pour que le schème d'alimentation du bébé fonctionne comme schème, il faut que l'enfant extraie telle quantité approximative de lait du biberon (moins il avalera de l'air, plus il s'étouffera), il faut qu'il déglutisse avec telle régularité, il faut qu'il aspire telle quantité d'air. Telles sont les normes intrinsèques de ce schème. Cette description rejoint la notion de contenu basique de Rolla et Huffermann (2022), c'est-à-dire que les schèmes sensorimoteurs ont des conditions de succès. D'autres schèmes auront d'autres normes, d'autres conditions de succès. Qui plus est, la normativité peut également survenir au niveau de l'organisme entier, c'est-à-dire qu'un schème sensorimoteur qui est évalué positivement peut être renforcé en une habitude jusqu'à ce que, pour une raison ou une autre, le même schème ne soit plus évalué de la même façon et devient soudainement fautif. Par exemple, un sportif ayant une certaine technique de course qui était suffisante pour gagner une course à un bas niveau de compétition devra modifier son enjambée pour pouvoir gagner à un plus haut niveau de compétition. Autrement dit, les schèmes sensorimoteurs ont toujours une dimension normative, et celle-ci est parfois irréductible aux normes biologiques. Cette dimension normative permet de rendre compte d'une notion non représentationnelle de contenu, le contenu basique (Rolla et Huffermann, 2022).

J'avance que cette description des schèmes sensorimoteurs est utile dans le projet de naturalisation de l'intentionnalité pour deux raisons. D'une part, la théorie établie dans *Sensorimotor life* (2017) permet de rendre compte des caractéristiques de l'intentionnalité pragmatique telles que définies dans le premier chapitre. En effet, j'écrivais que l'intentionnalité pragmatique était l'affaire d'une action dirigée vers le monde externe, et que la signification de l'action était comme, en suivant la Phénoménologie de la perception de Merleau-Ponty, « une saisie du résultat assurée par le corps lui-même comme puissance motrice, un 'projet moteur' (Bewegensentwurf), une 'intentionnalité motrice' (...) » (p. 141). La notion de schème sensorimoteur tel que décrit par l'énactivisme permet de retrouver cette définition. Par exemple, la notion d'assimilation indique que les schèmes sensorimoteurs ont souvent besoin d'intégrer

les éléments externes pour avoir un sens. Téter n'a un sens que s'il y a quelque chose à téter. Déglutir n'a un sens que s'il y a quelque chose à avaler (ne serait-ce que la présence de salive). Plus encore, ces schémas ne pourraient pas se développer sans la présence des éléments externes. Autrement dit, les éléments externes, au terme de l'assimilation, viennent à faire partie du schème sensorimoteur. Par ailleurs, les schèmes sensorimoteurs plus complexes peuvent témoigner d'une intentionnalité différente de celles des sous-schémas ou des coordinations sensorimotrices la composant. Par exemple, il serait possible de dire que tous les schèmes sensorimoteurs nécessaires pour cuisiner un repas ciblent des éléments spécifiques à la conception du repas, mais aucun de ces sous-schémas, individuellement, n'est dirigé intentionnellement vers le repas en tant que tel. On retrouve la structure directionnelle de l'intentionnalité sans avoir besoin d'inclure une représentation d'un but ou de l'objet ciblé par le schème sensorimoteur. En ce sens, les notions de schème sensorimoteur et d'agentivité sensorimotrice permettent la naturalisation de ce que j'ai nommé l'intentionnalité pragmatique.

De plus, la naturalisation de l'intentionnalité pragmatique par l'énactivisme réutilise les notions qui servaient à naturaliser l'intentionnalité biologique. En effet, la notion d'agentivité sensorimotrice reprend les critères établis par Barandiaran, Di Paolo et Rohde (2009) de même que les discussions de Di Paolo (2005; 2009) sur l'agentivité biologique (l'individualité, la normativité et l'interaction asymétrique), elle-même basée sur la notion d'autonomie (fermeture opérationnelle précaire) pour l'appliquer au niveau sensorimoteur. Le cadre conceptuel énaïviste qui s'en dégage est à la fois assez complet pour intégrer l'intentionnalité biologique et pragmatique, mais assez souple pour expliquer les spécificités de chaque niveau. En effet, si l'intentionnalité pragmatique découle de la même logique directionnelle que l'intentionnalité biologique, elle ne saurait s'y réduire. Les normes sensorimotrices peuvent même parfois entrer en tension avec les normes biologiques, lors des habitudes dommageables pour la santé, par exemple.

En bref, au sein de l'énactivisme, ce sont les concepts d'autonomie et d'agentivité qui permettent de faire le pont entre l'intentionnalité biologique et l'intentionnalité pragmatique. En effet, les deux niveaux nécessitent une individuation, mais surtout une interaction asymétrique normative. Autrement dit, il y a un point d'origine de l'action, et cette action doit intégrer spécifiquement des objets externes afin de s'individualiser et de se maintenir. Le schème sensorimoteur permet ainsi d'offrir une version naturalisée de l'intentionnalité pragmatique.

2.3. Le *Skilled Intentionality Framework*: une alternative pour l'intégration des différents types d'intentionnalité?

Jusqu'à présent, j'ai présenté ce que j'ai nommé la stratégie d'intégration de l'intentionnalité de l'énactivisme. L'énactivisme explique d'abord l'intentionnalité biologique du vivant grâce aux notions

d'autonomie et de *sense-making*, puis il étend ces notions au niveau sensorimoteur pour expliquer l'intentionnalité pragmatique. Toutefois, l'énactivisme n'est pas la seule théorie qui suit une stratégie d'intégration des différents types d'intentionnalité.

En plus de l'énactivisme, une autre théorie semble suivre cette stratégie: le *Skilled Intentionality Framework* (SIF) élaboré par Bruineberg et Rietveld (2014) ainsi que Van Dijk et Kiverstein (2017). Le SIF propose d'intégrer différentes notions issues de la psychologie écologique, de la phénoménologie et de la neurobiologie pour construire un cadre cohérent et complet. Le cœur de ce cadre théorique comprend la notion d'affordance (Gibson, 1977), la notion de prise maximale de Merleau-Ponty (1949) et enfin la notion d'énergie libre (variationnelle) de Friston (2005). Une affordance (de l'anglais *to afford*, « permettre ») est une possibilité d'action offerte par l'environnement à un certain organisme en fonction de ses capacités sensori-motrices (par exemple, une chaise présente l'affordance de s'asseoir pour un être humain sans handicap). Je retiendrai pour le mémoire l'interprétation selon laquelle l'affordance est une propriété relationnelle<sup>30</sup> du système environnement-organisme (Chemero, 2009), c'est-à-dire qu'elle n'existe pas réellement « dans » l'environnement (ou « dans » l'esprit, par ailleurs)<sup>31</sup>.

Selon Lobo *et al.* (2018), l'une des inspirations de Gibson pour la notion d'affordance remonte au concept de schéma corporel (à ne pas confondre avec l'image corporelle<sup>32</sup>) chez Merleau-Ponty (1945). La psychologie écologique, qui rejette à la fois l'intellectualisme du cognitivisme et le physicalisme du béhaviorisme, tire ses inspirations notamment du pragmatisme de W. James, mais également de la phénoménologie merleau-pontienne<sup>33</sup> (Lobo *et al.*, 2018; Heft, 2001). En parlant du corps, Merleau-

---

<sup>30</sup> J'ai fait ce choix entre autres pour simplifier la présentation, mais aucun enjeu nous concernant ici ne dépend de ce choix.

<sup>31</sup> Dans *The Ecological approach to visual perception* (1979), le texte séminal de Gibson, celui-ci indique que les affordances « are both physical and psychical, yet neither » (p. 121), une position que Lobo *et al.* (2018) rattachent au « monisme neutre » tiré du pragmatisme de W. James. Lobo *et al.* (2018) font également la distinction entre deux interprétations de l'affordance: l'affordance comme caractéristique de l'environnement qui est complémentaire à l'organisme (Turvey, 1992) et l'affordance comme caractéristique du système organisme-environnement (Chemero, 2009).

<sup>32</sup> Il est important de distinguer le schéma corporel de l'*image corporelle* (Gallagher & Cole, 1995). Gallagher & Cole décrivent ainsi le schéma corporel:

(...) a system of motor capacities, abilities, and habits that enable movement and the maintenance of posture. The body schema is not a perception, a belief, or an attitude. Rather, it is a system of motor and postural functions that operate below the level of self-referential intentionality, although such functions can enter into and support intentional activity.(Gallagher & Cole, 1995, p. 371)

À l'inverse, l'image corporelle désigne un ensemble de croyances et d'attitudes ayant comme contenu intentionnel le corps et est donc à rapprocher de l'intentionnalité représentationnelle plutôt que de l'intentionnalité pragmatique, même si Gallagher & Cole reconnaissent qu'une interaction est possible entre le schéma corporel et l'image corporelle, et plus généralement entre l'intentionnalité pragmatique (*self-referential intentionality*) et l'intentionnalité représentationnelle.

<sup>33</sup> Lobo *et al.* (2018) notent également l'influence du béhaviorisme et de la théorie de la Gestalt.

Ponty indique que « c'est en tant qu'il est polarisé par ses tâches, qu'il existe vers elles, qu'il se ramasse sur lui-même pour atteindre son but, et le 'schéma corporel' est finalement une manière d'exprimer que mon corps est au monde » (1945, p. 130). Lobo *et al.* (2018) détaillent la manière dont la notion d'affordance et son émergence issue des contraintes réciproques entre l'organisme et son environnement sont analogues à la manière dont Merleau-Ponty décrit l'émergence de « milliers de signes » qui guident l'action de l'organisme tout en étant partiellement générés par l'action de l'organisme même:

The environment (...) guides and constraints behavior in (the Merleau-Pontian) approach through signs, and those signs are also partially generated by the active capacities of agents. This is clearly in line with the Gibsonian view by which the combination of action and environmental elements gives rise to the emergence of ecological information and affordances. Those signs are similar to affordances, because they are meaningful in an embodied way and guide the agent's behavior. (Lobo *et al.*, 2018, p. 4)

Cet extrait nous permet de saisir le lien entre l'intentionnalité motrice merleau-pontienne et l'affordance. Chez Merleau-Ponty, le corps est toujours dirigé vers des tâches (des actions), il est au monde, c'est-à-dire qu'il est dirigé vers lui. L'affordance, comme possibilité d'action offerte par l'environnement, sous-entend également cette même directionnalité du corps vers son environnement, qui offre des possibilités d'action en fonction des capacités sensori-motrices du corps (de l'organisme). Les affordances sont significatives comme les signes de Merleau-Ponty sans qu'aucune intentionnalité représentationnelle ne soit nécessaire comme intermédiaire entre le monde perçu et l'action: les affordances sont directement perçues par les systèmes sensorimoteurs, non pas comme une succession de représentations sensorielles et motrices (selon une combinaison input-output), mais tout de même de manière intentionnelle. En ce sens, la notion d'affordance (Gibson, 1979; Rietveld & Kiverstein, 2014) peut se rapprocher d'une forme d'intentionnalité pragmatique<sup>34</sup>.

Erik Rietveld (2012) pousse plus loin l'association entre l'intentionnalité motrice de Merleau-Ponty et les affordances. Plutôt que de suivre ce qu'il nomme « l'interprétation orthodoxe » de l'intentionnalité motrice où l'action est dirigée vers des objets, comme une main qui se tend vers une tasse (Thompson, 2007; Kelly, 2002), Rietveld tente de redéfinir l'intentionnalité motrice comme étant la réactivité de l'organisme à un champ d'affordances (Rietveld, 2012). Un champ d'affordance désigne l'ensemble des sollicitations (des affordances particulièrement pertinentes et saillantes) disponibles à un moment donné pour un organisme (Bruineberg & Rietveld, 2014; Rietveld & Kiverstein, 2014; Van Dijk & Kiverstein, 2017). Rietveld reprend l'analyse de Dreyfus (2002) où l'intentionnalité motrice merleau-pontienne est conçue comme un mouvement du corps vers une prise optimale (*optimal grip*) sur le monde, ce qui sous-

---

<sup>34</sup> Dreyfus (2002), en commentant la critique merleau-pontienne du concept de représentation, fait un rapprochement similaire entre la notion d'affordance de Gibson et ce qu'il nomme une « intelligence sans représentation ».

entend une normativité<sup>35</sup> dans l'intentionnalité motrice. Le corps n'est pas simplement dirigé vers le monde, mais cherche constamment à améliorer son emprise sur celui-ci. Dans le vocabulaire de l'écologie psychologique, cela signifie que le corps se positionne en fonction de la pertinence des affordances qui le sollicite dans le champ d'affordance: plus une affordance est pertinente, plus celle-ci va solliciter l'organisme, et plus l'organisme aura tendance à agir pour répondre à cette affordance, se dirigeant ainsi vers une prise optimale sur le monde. Toutefois, cela n'implique pas pour autant que le corps réponde toujours de la même façon au champ d'affordance, en se dirigeant toujours vers l'affordance la plus saillante: plusieurs affordances peuvent entrer en compétition, et peuvent créer une tension dans l'atteinte de l'optimum (Dreyfus, 2007).

Enfin, le SIF intègre également une dimension biologique à son explication de l'intentionnalité. En effet, Bruineberg et Rietveld (2014) expliquent qu'il est possible d'expliquer l'auto-organisation minimale des êtres vivants par le Principe d'énergie libre (variationnelle). Le Principe d'énergie libre (Friston, 2005; Friston, 2006; Friston, 2010) stipule qu'un organisme, pour survivre, tend à minimiser le désordre (ou la désorganisation) de ses propres états afin de résister à l'augmentation de l'entropie à laquelle doit faire face tout système vivant. Pour ce faire, il tend à minimiser l'énergie libre (variationnelle), c'est-à-dire l'écart entre les états dans lequel l'organisme se trouve présentement et les états dans lesquels l'organisme a tendance à se trouver typiquement. La logique est similaire à celle de l'énactivisme, c'est-à-dire que l'organisme cherche à se maintenir dans une zone de viabilité pour conserver son auto-organisation. Par ailleurs, Kirchhoff (2018) soutient que l'énactivisme et le Principe d'énergie libre sont compatibles à de nombreux égards, notamment concernant la thèse de la continuité entre l'esprit et le vivant.

Si le SIF (et, dans une mesure plus large, le Principe d'énergie libre) a un potentiel intéressant pour comprendre l'intégration des différents types d'intentionnalité, j'argumente que l'énactivisme est plus avancé dans l'organisation de son cadre théorique dans l'état actuel des choses. Cela ne veut pas dire que l'énactivisme sera nécessairement la voie du futur (ou même que son entreprise fonctionnera totalement), mais pour le moment, les concepts énéactives semblent mieux articulés entre eux que les concepts du SIF, qui viennent de cadres théoriques différents et qui nécessitent encore un travail d'intégration entre eux. Par ailleurs, il n'est pas impossible que les défenseurs des deux cadres théoriques (l'énactivisme et le SIF) réalisent que ces théories sont assez compatibles pour être intégrées dans une seule grande théorie, qui expliquerait l'intentionnalité selon la stratégie de l'intégration que j'ai

---

<sup>35</sup> Cela est sans compter la normativité *culturelle* qui peut interagir avec les affordances, par exemple dans l'utilisation des outils, qui possèdent des affordances spécifiquement culturelles au-delà de la possibilité physique d'interaction. Autrement dit, l'intentionnalité motrice peut interagir avec des formes d'intentionnalité représentationnelle (éléments culturels, consignes, symboles, etc.). Voir les notions d'*affordance culturelle* dans Ramstead, Veissière et Kirmayer (2016) et Rietveld & Kiverstein (2014).

présentée dans le mémoire. Une ébauche de cette grande intégration énoncivisme-Principe d'énergie libre a notamment été formulée par Wiese et Friston (2021), qui proposent justement une continuité entre ce qu'ils nomment l'intentionnalité autopoïétique (l'intentionnalité biologique et pragmatique selon les termes du mémoire) et la représentationnalité (l'intentionnalité représentationnelle). Toutefois, l'examen minutieux de la possibilité de cette entreprise d'intégration dépasse le cadre du mémoire. Je me contenterai donc de dire que l'énoncivisme semble avoir une certaine avance dans sa stratégie d'intégration des différents types d'intentionnalité.

#### 2.4. Conclusion : une naturalisation partielle par l'intégration des intentionnalités biologique et pragmatique

Dans ce chapitre, nous avons vu que la stratégie traditionnelle pour naturaliser l'intentionnalité était une stratégie de réduction, c'est-à-dire que le problème était perçu comme étant souvent celui de l'explication biologique de l'intentionnalité représentationnelle. Ainsi, les différentes théories cherchaient à expliquer l'existence de représentations internes. Je me suis attardé sur les théories téléosémantiques, pour lesquelles la solution au problème de naturalisation de l'intentionnalité se trouve dans la téléologie, c'est-à-dire que les structures internes sont des représentations si elles acquièrent la fonction de représenter (des éléments externes). Le défi de naturalisation de l'intentionnalité représentationnelle revient alors à expliquer l'acquisition de la fonction représentationnelle par des structures internes. À ce propos, j'ai présenté rapidement l'explication ontogénétique de Dretske ainsi que l'explication phylogénétique de Millikan. Toutefois, ces théories ont été vertement critiquées à plusieurs égards et leur potentiel d'explication a été remis en question.

J'ai alors présenté l'énoncivisme comme solution potentielle au problème de naturalisation de l'intentionnalité. La stratégie des énoncivistes est de naturaliser non pas l'intentionnalité représentationnelle directement, mais plutôt de suggérer que l'intentionnalité (biologique) est une propriété fondamentale du vivant sur laquelle s'échafaude les autres types d'intentionnalités plus complexes. J'ai présenté les concepts fondamentaux d'autonomie et de *sense-making*. J'ai par la suite expliqué que le *sense-making* était la forme que prenait l'intentionnalité biologique selon l'énoncivisme. Cette intentionnalité biologique de base est l'affaire d'une recherche de l'autoconservation, réalisant une téléologie intrinsèque qui transforme l'environnement ambiant en milieu avec des valences. Par la suite, j'ai expliqué que l'autonomie s'appliquait également au niveau sensorimoteur et qu'une normativité émerge également au niveau des schèmes sensorimoteurs. Cette normativité permet de rendre compte de l'intentionnalité pragmatique, c'est-à-dire que les schèmes ont toujours une direction et un objet, que ce soit un élément physique ou encore un autre schème sensorimoteur.

La stratégie de l'énactivisme est donc bel et bien celle d'une intégration de différents types d'intentionnalité, en ce sens que l'intentionnalité biologique est comprise comme étant l'intentionnalité fondamentale qui permet d'échafauder l'intentionnalité pragmatique qui voit une normativité distincte (i.e., des normes distinctes) émerger. En effet, certaines normes du niveau sensorimoteur peuvent découler des normes biologiques, mais d'autres normes peuvent carrément émerger et entrer en contradiction avec les normes biologiques. Ainsi, si les normes sensorimotrices tirent leur origine du caractère vivant de l'organisme, elles ne sauraient s'y réduire. Tel est le sens de l'intégration des différents types d'intentionnalité par l'énactivisme.

Toutefois, un défi se pointe à l'horizon de l'intégration: l'intentionnalité représentationnelle. En effet, même si je crois avoir démontré que l'énactivisme fait un travail adéquat pour naturaliser l'intentionnalité biologique et l'intentionnalité pragmatique, il reste encore à montrer que ce cadre théorique permet également de rendre compte de l'intentionnalité représentationnelle (la cible la plus importante, diront certains). À cette fin, le troisième et dernier chapitre s'intéressera à la manière dont l'énactivisme entend naturaliser l'intentionnalité représentationnelle. Je présenterai rapidement l'objection du *scaling-up* (l'incapacité apparente d'expliquer les phénomènes cognitifs de « haut niveau » comme le langage et les normes socioculturelles), puis j'esquisserai les grandes lignes du modèle du corps linguistique de Di Paolo *et al.* (2018) pour déterminer si l'énactivisme est de taille pour surmonter le défi du *scaling-up*, naturaliser l'intentionnalité représentationnelle et ainsi compléter la stratégie d'intégration des différents types d'intentionnalité.

## CHAPITRE III

### NATURALISER L'INTENTIONNALITÉ REPRÉSENTATIONNELLE : MISSION IMPOSSIBLE POUR L'ÉNACTIVISME?

Jusqu'à présent, j'ai tenté de démontrer quelle était la stratégie de l'énactivisme pour traiter le problème de la naturalisation de l'intentionnalité. La première étape était de défendre l'existence de trois grands types d'intentionnalité: biologique, pragmatique et représentationnelle. Après avoir présenté rapidement quelques tentatives de naturalisation de l'intentionnalité représentationnelle par les théories téléosémantiques de Dretske et Millikan, et leur échec, j'ai résumé le cadre conceptuel de l'énactivisme (construit, pour ce qui nous concerne, autour des notions d'autonomie et de *sense-making*), puis j'ai montré que l'énactivisme pouvait naturaliser l'intentionnalité biologique grâce à ces concepts, et qu'il pouvait également naturaliser l'intentionnalité pragmatique en étendant ceux-ci au niveau sensorimoteur. La prochaine étape pour compléter la présentation de la stratégie énacliviste d'intégration de l'intentionnalité est de vérifier si cette stratégie est à même de naturaliser l'intentionnalité représentationnelle. L'énactivisme est-il à la hauteur du défi?

#### 3.1. Le problème du *scaling-up* de l'énactivisme

Malgré le gain en intérêt de l'énactivisme – et de son rejet de la notion de « représentation » - cette approche n'est pas exempte de critiques. En effet, même en acceptant que l'énactivisme est à même d'expliquer les phénomènes de « bas niveau » (comme la perception et l'action), certains auteurs vont questionner sa capacité d'expliquer la cognition dite de « haut niveau », par exemple le langage, les normes socioculturelles, la réflexion, l'imagination, etc. Gallagher nomme cette objection le *scaling-up problem* (Gallagher, 2017). Plusieurs variantes de l'objection existent (*cognitive gap*, problème de la cognition virtuelle, etc.), mais elles dérivent principalement de la formulation de Clark et Toribio (1994): le problème de l'énactivisme pour expliquer la cognition dite « *representation-hungry* ».

Grosso modo, le problème de la cognition « *representation-hungry* » réfère à la difficulté apparente des approches anti-représentationalistes de rendre compte de certains phénomènes cognitifs, notamment les phénomènes impliquant une relation intentionnelle à des choses absentes ou imaginaires. Clark et Toribio spécifient qu'un phénomène est « *representation-hungry* » s'il respecte l'une des deux conditions suivantes: 1) le phénomène implique de raisonner à propos de situations absentes, inexistantes ou contrefactuelles; 2) le phénomène implique d'être attentif à des paramètres dont les manifestations physiques sont complexes et difficilement prédictibles par des régularités nomologiques, par exemple des paramètres inobservables et abstraits qui ne sont pas accessibles à la perception sensori-

motrice (Clark & Toribio, 1994). On peut également penser à ce que Hockett (1960) nomme le « déplacement de la référence » en linguistique, c'est-à-dire le fait que le langage humain peut référer à des choses qui dépassent l'ici et le maintenant. La cognition non linguistique peut également faire intervenir des objets qui dépassent l'ici et le maintenant, notamment dans les cas de l'imagination et des souvenirs. Dans ces situations, la notion de représentation (comme état qui tient lieu d'un autre) est bien utile. En tant que « *stand-in* » ou « *proxy* » d'une chose absente, la représentation permet justement de combler le vide et d'expliquer comment un organisme peut interagir avec des choses avec lesquelles il n'est pas en contact direct, voire ne peut l'être, celles-ci n'existant pas.

Le problème du *scaling-up* de l'énactivisme se manifeste tant dans le domaine linguistique que non-linguistique. D'une part, il survient lorsque l'on cherche à expliquer par une explication énaïvistique (fondée sur le couplage sensorimoteur entre l'organisme et l'environnement) de phénomènes qui font intervenir des entités absentes, inexistantes ou imaginaires, par exemple la remémoration d'un événement passé. Si l'énactivisme est adéquat pour expliquer l'intentionnalité biologique et pragmatique en termes d'interactions avec son environnement immédiat, comment peut-il expliquer le couplage d'un organisme avec des entités qui n'existent pas ou avec lesquels l'organisme n'a pas de contact direct? Une façon de répondre au problème du *scaling-up* de l'énactivisme serait de prendre en compte l'existence de représentations portant un contenu représentationnel vers ces entités absentes, inexistantes ou imaginaires, et cela tant au niveau du langage que de celui d'autres manifestations culturelles. Même en acceptant que l'esprit soit, à son fondement, non représentationnel, l'énactivisme peut-il expliquer l'émergence de représentations (externes) et de nos capacités de manipulation de celles-ci?

Le problème du *scaling-up* permet de mettre en lumière la différence entre les stratégies téléosémantiques et énaïvistes face au problème de la naturalisation de l'intentionnalité. D'un côté, la stratégie téléosémantique (de Dretske et Millikan) présume l'existence des représentations mentales, puis cherche à les expliquer dans un vocabulaire biologique non représentationnel. On pourrait dire que c'est une approche descendante (*top-down*), des phénomènes les plus complexes (l'intentionnalité représentationnelle) jusqu'aux plus « simples » (la biologie des organismes). De l'autre, l'énactivisme fait le pari que l'intentionnalité se trouve sur un spectre continu entre les niveaux biologique, sensorimoteur et socioculturel et que l'explication passera par une intégration du niveau le plus simple jusqu'au plus complexe. Le défi des énaïvistes consiste dès lors à démontrer que leur appareil conceptuel peut rendre compte non seulement du niveau organique des êtres vivants, mais également des comportements de plus en plus complexes qu'ils développent, jusqu'aux manifestations socioculturelles.

Ce dernier chapitre pose ainsi une question claire: peut-on naturaliser l'intentionnalité représentationnelle à l'aide des notions développées dans les chapitres précédents et régler ainsi le problème du *scaling-up* de l'énactivisme? Pour répondre à cette question, je propose de m'intéresser au modèle du *corps linguistique* (*linguistic body*) développé par Di Paolo, Cuffari et de Jaegher dans leur ouvrage *Linguistic Bodies* (2018), issu notamment des travaux de Cuffari *et al.* (2015) ainsi que de De Jaegher (2007). Ce modèle présente une manière dont l'énactivisme entend naturaliser les capacités langagières des êtres humains. J'argumente que la stratégie de Di Paolo et de ses collègues se rapproche des théories néo-pragmatistes de l'intentionnalité représentationnelle, pour lesquelles le contenu représentationnel (et éventuellement, propositionnel) tire son origine des normes socioculturelles. Ce modèle du corps linguistique n'est évidemment pas la seule tentative existante pour répondre au problème du *scaling-up* de l'énactivisme. On peut notamment citer les travaux de Sanneke de Haan (2020), Shaun Gallagher (2017), Michelle Maiese (2016; 2022) ainsi que ceux de Kiverstein et Rietveld (2018) que nous avons discutés plus tôt. Toutefois, je me concentrerai sur le modèle du corps linguistique puisque je crois que c'est le modèle qui, jusqu'à présent, permet d'envisager le plus efficacement l'intégration entre les intentionnalités biologiques, pragmatiques et représentationnelles. En effet, comme nous verrons dans les sections suivantes, le modèle du corps linguistique se construit sur les concepts énaïvistes. J'argumente que ce modèle permet (en partie) d'expliquer le passage des intentionnalités biologiques et pragmatiques à l'intentionnalité représentationnelle, laissant entrevoir la possibilité de la complétion du projet de naturalisation de l'intentionnalité tel que présenté comme problématique du mémoire.

Le reste du chapitre sera organisé de la manière suivante. Dans un premier temps, à la section 3.2, je présenterai les grandes lignes des théories néo-pragmatistes du contenu en mettant l'accent sur trois thèses: le pragmatisme, la socialité et la normativité. Par la suite, à la section 3.3, je présenterai le modèle du corps linguistique de Di Paolo *et al.* (2018) en trois parties: le *sense-making* participatif, le *linguaging* et enfin la grammaire énaïviste. Pour finir, à la section 3.4, je discuterai de quelques enjeux en lien avec le modèle du corps linguistique, notamment la notion de contenu utilisée, les critiques du néo-pragmatisme et enfin la possibilité d'une intégration avec l'énactivisme radical pour ajouter une dimension diachronique au modèle.

### 3.2. Une conception néo-pragmatiste de l'intentionnalité représentationnelle

Comment intégrer l'intentionnalité représentationnelle dans le cadre énaïviste pour compléter la stratégie d'intégration des types d'intentionnalité? Dans cette section, je propose qu'une avenue prometteuse pour l'énactivisme repose sur le fait d'accepter une conception néo-pragmatiste (Haugeland, 1990) de l'intentionnalité représentationnelle. La tâche de l'énactivisme sera par la suite de

montrer que leurs concepts permettent de rendre compte de l'intentionnalité représentationnelle à la manière des néo-pragmatistes.

La thèse principale des théories néo-pragmatistes de l'intentionnalité est que le contenu représentationnel et propositionnel est normatif, et que cette normativité dépend ultimement des pratiques sociales (Cash, 2009; Green, 2021; Kripke, 1982; Tracy, 2020). Concrètement, l'idée est que s'il existe un contenu représentationnel, il émerge des interactions sociales et culturelles plutôt que de représentations mentales internes. Le néo-pragmatisme inverse donc la conception cognitiviste de l'intentionnalité représentationnelle (Haugeland, 1990): ce n'est pas l'intentionnalité des représentations mentales qui est originelle et qui permet de dériver l'intentionnalité du langage (et des normes socioculturelles); c'est l'intentionnalité de la culture et du langage qui serait originelle et l'intentionnalité représentationnelle « privée » qui en serait dérivée. Des variations de cette idée se retrouvent également chez Brandom (1994), Sellars (1953), et ces idées s'inspirent, dans une certaine mesure, de la philosophie de Wittgenstein<sup>36</sup>, notamment son travail dans *Philosophical Investigations* (2010/1953).

Pour comprendre l'intentionnalité représentationnelle, il faut donc selon le néo-pragmatisme déterminer comment s'établissent les normes sociales. Haugeland (1990) avance que c'est par le conformisme, c'est-à-dire la propension des êtres humains à faire comme leurs pairs (l'imitation) combinée à leur propension à supprimer les variations du comportement (la censure). Par exemple, un enfant pourrait imiter ses parents, et les parents pourraient le gronder si l'enfant agit autrement. Selon Haugeland, le conformisme provoque des changements dans les dispositions comportementales des participants, ce qui l'amène à affirmer que sans conformisme, il n'y aurait pas d'explication néo-pragmatiste de l'intentionnalité:

Conformism - imitativeness and censoriousness together - is a causal interactive process that results in more or less permanent changes in the behavioral dispositions of its participants. Without such a process, there can be no community, no social practices, and no neo-pragmatist account of intentionality. (Haugeland, 1990, p. 405)

Les normes sociales émergent lorsque les dispositions comportementales des membres d'une communauté s'agrègent et convergent en un point commun, l'ensemble des normes sociales de la communauté établissant une forme de vie (Haugeland, 1990; voir également Wittgenstein, 2010/1953). Les normes sociales qui composent cette forme de vie sont interdépendantes. La littérature néo-pragmatiste illustre souvent cette idée au moyen des outils: chaque outil a un usage normatif, mais

---

<sup>36</sup> En suivant la métaphore de baseball de Haugeland (1990), on peut comprendre que celui-ci n'inclut pas entièrement Wittgenstein dans le néo-pragmatisme, mais qu'il le place plutôt à mi-chemin entre le néo-pragmatisme et le néo-béhaviorisme. La pertinence de cette classification historique est cependant secondaire dans le cadre de ce mémoire.

l'ensemble des outils forme un réseau normatif. Par exemple, le marteau est fait pour frapper des clous, le tournevis pour serrer des vis, etc., les deux étant différentes manières de joindre des matériaux ou des objets de manière relativement stable et permanente.

Selon Haugeland (1990), les néo-pragmatistes considèrent que le langage lui-même est un ensemble d'outils et que son usage principal est d'influencer les « circonstances normatives », c'est-à-dire modifier le cadre normatif dans lequel se déroule l'interaction sociale (la conversation) et préciser « *what's proper* » (Haugeland, 1990, p. 412), ce qui est approprié dans ce cadre normatif nouvellement établi. Ces outils prennent la forme d'actes de langage (Searle, 1969; voir aussi Austin, 1975/1962), c'est-à-dire que le langage permet d'accomplir une action par son énonciation même (par exemple: promettre, questionner, ordonner, etc.). Comme pour les autres outils, les actes de langage forment un réseau interdépendant, c'est-à-dire que les possibilités d'action offertes par certains énoncés sont déterminées par d'autres énoncés. Par ailleurs, fondamentalement, les actes de langage sont composés de symboles articulés sémantiquement (Haugeland, 1990). Toutefois, comme dit précédemment, les tokens linguistiques « are external, not internal, to the systems that use them; and the holistic pattern in which they occur includes far more than just their relations to one another indeed, it includes in principle all the interdependent relationships instituted by the way of life of which they are a part » (Haugeland, 1990, p. 412). Autrement dit, non seulement les tokens linguistiques dépendent des autres tokens linguistiques (et non pas de représentations mentales), mais dépendent également de relations normatives qui dépassent le cadre strictement linguistique.

Quel est le lien entre les normes sociales et le contenu propositionnel des énoncés? Deux thèses (principales) existent à ce propos (Glüer, Wikfross et Ganapini, 2009). Il y a ceux qui soutiennent que la normativité est engendrée par le contenu propositionnel (marquée CE pour *content engendered normativity*) et ceux qui soutiennent que le contenu est déterminé par la normativité (marquée CD pour *content determining normativity*). La version CE soutient ainsi que l'existence du contenu propositionnel a des conséquences normatives, notamment qu'un énoncé portant un certain contenu ne devrait être utilisé que lorsque ses conditions de satisfaction sont présentes. Par exemple, je ne devrais dire « Il pleut dehors. » que s'il pleut véritablement. La version CD est plus forte et soutient que ce sont les normes sociales qui fondent l'existence du contenu propositionnel, c'est-à-dire que les énoncés portent un contenu en vertu de l'existence de certaines normes sociales. Une version populaire de cette thèse se définit de la manière suivante: le contenu d'un énoncé est attribué par les autres membres de la communauté, qui sont en droit (*are entitled*) de s'attendre à certains engagements (*commitments*) de la part de la personne qui émet l'énoncé (Brandt, 1994; Tracy, 2019). Par exemple, si une personne énonce « Il pleut dehors. », les autres ont le droit de s'attendre à ce qu'il pleuve dehors, et cela va également les inciter à agir de telle ou telle manière. En retour, la personne qui émet l'énoncé s'engage à certaines choses, notamment à ce qu'il pleuve dehors et à agir en conséquence. Il est intéressant de

noter que la thèse CD implique la thèse CE; si le contenu est déterminé par des normes, il s'ensuit que son existence a des conséquences normatives (Tison, 2022). En citant Brandom (1994) qui s'inspire lui-même des travaux de Wittgenstein (Tracey, 2019), il est possible de résumer ainsi les thèses d'une théorie néo-pragmatiste:

[1] the insistence on the normative character of language and intentionality, [2] the pragmatist commitment to understanding these norms in terms of practices rather than exclusively in terms of rules, and [3] the recognition of the essentially social character of such norms. (Brandom, 1994, p. 55, nous avons ajouté la numérotation)

J'avance que l'énactivisme peut expliquer l'intentionnalité représentationnelle en reprenant ces trois thèses du néo-pragmatisme, à savoir la normativité, le pragmatisme, et la socialité, et qu'il peut faire ceci grâce à un cadre théorique développé par Di Paolo, Cuffari et de Jaegher (2018): le modèle du corps linguistique<sup>37</sup>.

Ce modèle énonctiviste du langage cherche à expliquer celui-ci en reprenant les concepts énonctivistes développés dans les sections précédentes, et notamment les deux concepts intentionnels que nous avons présentés: l'intentionnalité biologique et l'intentionnalité pragmatique ou sensorimotrice. D'une part, je soutiens que le modèle du corps linguistique est pragmatique, cette idée étant l'un des thèmes centraux de l'énactivisme, notamment en ce qui a trait au *sense-making* sensorimoteur. Ainsi, je ne reviendrai pas en détail sur cette dimension pragmatique de l'énactivisme, qui je crois semble assez évidente à ce point-ci du mémoire. D'autre part, je montrerai que le modèle du corps linguistique est social: j'expliquerai comment l'énactivisme entend expliquer la cognition sociale grâce à un nouveau concept, le *sense-making* participatif (De Jaegher et Di Paolo, 2007). Le *sense-making* participatif émerge lorsque deux participants interagissent (volontairement ou involontairement), générant un nouveau type de *sense-making* qui dépasse le niveau individuel du *sense-making* sensorimoteur. Enfin, pour expliquer la normativité spécifique au langage, je présenterai le modèle du corps linguistique de Di Paolo *et al.* (2018), qui cherche à montrer comment les normes proprement linguistiques émergent des interactions sociales comprises comme *sense-making* participatif. Le langage sera défini comme étant un *sense-making* réflexif, lequel permet de réguler non seulement les normes d'une interaction sociale, mais également de réguler la régulation de ces normes elle-même, permettant par le fait même une méta-régulation.

La prochaine section sera donc dédiée à la démonstration des dimensions sociales et normatives du modèle du corps linguistique de Di Paolo *et al.* (2018). Premièrement, à la section 3.3.1, je présenterai

---

<sup>37</sup> Dans leur ouvrage *Evolving Enactivism: Basic Minds Meet Content*, Daniel Hutto et Erik Myin (2017, p. 124), énoncent une proposition similaire, mais sans la développer en détails. Au mieux, ils rédigent une ébauche. Voir également la section 3.4.3.

le concept du *sense-making* participatif. Ensuite, à la section 3.3.2, je présenterai les étapes conceptuelles permettant de passer du *sense-making* participatif au langage selon le modèle de Di Paolo *et al.* (2018). Enfin, à la section 3.3.3, j'évaluerai critiqueusement le modèle du corps linguistique pour voir s'il répond aux attentes présentes, c'est-à-dire s'il permet d'intégrer l'intentionnalité représentationnelle au sein du cadre éenactiviste.

### 3.3. Le *scaling-up* de l'éenactivisme autopoïétique: le modèle du corps linguistique

Le modèle du corps linguistique (*linguistic body*) est un modèle dialectique développé par Di Paolo et ses collègues visant à expliquer comment la normativité sociale émerge des *sense-making* basique et sensorimoteur. Étant donné la thèse néo-pragmatiste énoncée à la section précédente, le modèle du corps linguistique nous permet d'envisager comment l'intentionnalité représentationnelle peut s'intégrer aux intentionnalités biologiques et pragmatiques présentées respectivement aux sections 2.2.2 et 2.2.5. La stratégie de Di Paolo et de ses collègues pour développer le modèle du corps linguistique, présentée dans l'ouvrage *Linguistic bodies* (2018), est de montrer d'abord comment le *sense-making* sensorimoteur peut devenir ce qu'ils appellent le *sense-making* participatif, c'est-à-dire la forme de *sense-making* présente lorsque deux personnes interagissent de manière sensorimotrice (De Jaegher et Di Paolo, 2007), puis comment ce *sense-making* social s'organise en dialogues, générant ainsi un nouveau type de normativité.

#### 3.3.1. Le *sense-making* participatif

La première étape pour pouvoir expliquer l'intentionnalité représentationnelle consiste à expliquer les interactions sociales grâce aux concepts éenactivistes. Pour ce faire, une nouvelle notion est introduite: le *sense-making* participatif (*participatory sense-making*, de Jaegher et di Paolo, 2007). De Jaegher et Di Paolo définissent cette notion comme étant « the coordination of intentional activity in interaction, whereby individual sense-making processes are affected and new domains of social sensemaking can be generated that were not available to each individual on her own » (de Jaegher et di Paolo, 2007, p. 497).

L'idée derrière cette notion est de pouvoir rendre compte d'une nouvelle dimension de la cognition (la dimension sociale) qui se distingue des niveaux précédents, mais sans changer radicalement la nature des interactions entre l'organisme et son environnement, maintenant composé d'autres organismes avec lesquels il interagira. D'une certaine manière, l'idée de l'éenactivisme est d'expliquer l'émergence d'un milieu social de manière analogue à l'émergence d'un milieu naturel pour l'organisme, dans la lignée de la thèse de la continuité entre le vivant et l'esprit (Thompson, 2007). Comme le souligne de Haan (2020), le milieu ambiant des êtres humains est également un milieu socioculturel. Di Paolo et de Jaegher

vont donc tenter de développer un concept d'interaction entre les individus qui prendra la forme d'un *sense-making* participatif.

Di Paolo et de Jaegher conçoivent l'interaction (sociale) entre plusieurs individus comme plus qu'une simple « coïncidence spatio-temporelle de deux agents qui s'influencent mutuellement » (Di Paolo & de Jaegher, 2007, p. 492). Di Paolo et de Jaegher vont s'intéresser à l'interaction elle-même et vont avancer que celle-ci possède sa propre structure interne. Selon eux, une interaction sociale est un *pattern* dynamique de coordination (une covariation non accidentelle, mais pas nécessairement délibérée ou volontaire) entre deux agents (ou plus) qui est l'objet d'une double influence : d'une part, la coordination des agents influence le déroulement de l'interaction et, d'autre part, le *pattern* dynamique de l'interaction elle-même influence la propension des agents à continuer ou à terminer la rencontre. Autrement dit, les agents maintiennent l'interaction, et l'interaction elle-même influence les agents et les transforme en interacteurs d'une manière qui tend à maintenir l'interaction.

Di Paolo et de Jaegher notent que, de ces interactions entre les agents, émerge un système précaire doté de fermeture opérationnelle, c'est-à-dire un système autonome: « In accordance with the core ideas of enaction, the above description is nothing less than that of an emergent and autonomous process. » (De Jaegher et Di Paolo, 2007, p. 492). Un exemple de De Jaegher et Di Paolo implique deux personnes se croisant dans un corridor étroit et se rangeant en même temps sur le même côté du corridor pour tenter de laisser passer l'autre, mais qui ainsi se bloquent mutuellement par inadvertance. La symétrie des mouvements entre les deux individus augmente les chances que ceux-ci tentent en même temps de se déplacer de l'autre côté du corridor pour tenter de laisser passer l'autre, puisque les possibilités d'action sont limitées par la symétrie elle-même et par l'étroitesse du corridor. Autrement dit, l'interaction elle-même (la symétrie des mouvements des personnes se bloquant le chemin) influence les actions des participants, qui en viennent à maintenir l'interaction (malgré eux). Ainsi, selon De Jaegher et Di Paolo, une interaction véritablement sociale déploie une structure opérationnellement fermée (comme vu précédemment dans la présentation du concept d'autonomie), dans la mesure où les agents en tant qu'interacteurs forment une boucle maintenant la structure de l'interaction. De Jaegher et Di Paolo vont jusqu'à ajouter ceci : « Viewing interactions from this perspective is akin to understanding the growth of an adaptive system. » (De Jaegher et Di Paolo, 2007, p. 496). Ce point vise à indiquer que certaines actions ont tendance à maintenir l'interaction (continuer le mouvement de symétrie) ou à la briser (prendre un pas de recul ou s'arrêter et laisser passer l'autre), comme dans le cas d'un système autonome capable de s'adapter. Autrement dit, il y a une certaine régulation de l'interaction sociale par les actions des participants (même si cette régulation ne donne pas nécessairement les résultats escomptés par ceux-ci!).

Cuffari, Di Paolo et de Jaegher (2015) vont parler d'une normativité interactive pour décrire les actions affectant l'interaction elle-même. Cette normativité interactive dépasse la normativité biologique ou sensorimotrice issue du *sense-making* individuel, et peut même entrer en conflit avec celle-ci (comme lorsque les agents tentent de mettre fin à l'interaction sociale et que celle-ci continue malgré eux). Ils soulignent que pour qu'une interaction demeure véritablement sociale, l'autonomie des participants se doit d'être préservée (même si elle peut être augmentée ou diminuée). Retirer complètement l'autonomie – au sens énaïviste - d'un participant reviendrait à le transformer en « outil » pour le ou les autres participants. (De Jaegher & Di Paolo, 2007) Si la régulation de l'interaction sociale se fait par des actions, et que les actions sont la manière de réguler le *sense-making* des organismes, il en résulte que les participants à une interaction sociale peuvent coordonner leur *sense-making* au sein de cette interaction sociale. C'est ce que De Jaegher et Di Paolo (2007) nomment *sense-making* participatif.

Les points importants à retenir peuvent se résumer à 1) la reconnaissance du caractère « socialement constitué » du *sense-making* humain et 2) la manière dont le concept de *sense-making* se voit combiné à celui d'interaction sociale pour engendrer le *sense-making* participatif, c'est-à-dire la coordination des *sense-making* individuels. L'important ici consiste donc à comprendre que les énaïvistes tentent de conserver une continuité entre l'organisation du vivant et celle de la socialité, et que cette continuité passe par le concept de *sense-making*.

### 3.3.2. Le modèle du corps linguistique: un *sense-making* participatif sous forme de dialogue

Si la notion de *sense-making* participatif est très intéressante pour ajouter une dimension sociale à l'énaïvisme, elle n'est pas suffisante pour rendre compte de la normativité propre au niveau socioculturel, notamment celle propre au langage et, suivant les thèses néo-pragmatistes que nous venons de discuter, à l'intentionnalité représentationnelle. En particulier, rappelons-nous de la thèse CD portant sur la normativité du contenu intentionnel, c'est-à-dire que ce sont les normes qui déterminent le contenu intentionnel. Peut-on expliquer cette thèse dans des termes énaïvistes? J'avance que c'est ce que permet (du moins en partie) le modèle du corps linguistique de Di Paolo *et al.* (2018). Comme je le montrerai dans cette section et la suivante, le modèle a le potentiel d'expliquer le langage dans des termes énaïvistes et, ainsi, expliquer comment l'énaïvisme peut rendre compte d'un phénomène d'intentionnalité représentationnelle. Ensuite, à la section 3.3.4, je discuterai de certaines limites du modèle du corps linguistique.

Di Paolo et ses collègues vont définir le langage comme étant du *linguaging*, « a form of social agency involving a double regulation of self and interaction that integrates the tensions inherent in dialogical organization and participation genres » (Cuffari *et al.*, 2015, p. 1092). En gros, le *linguaging* est une forme particulière de *sense-making* participatif qui implique une régulation de soi-même et de

l'interaction sociale sous la forme d'un dialogue. Le terme « *linguaging* » est repris de Maturana (2002), qui utilise cette notion pour décrire la manière humaine de vivre. Cette idée est poursuivie par Di Paolo *et al.* (2018, voir aussi Anderson, 2014), qui vont argumenter que le langage ne devrait pas principalement être compris comme un ensemble abstrait de symboles et de règles, mais plutôt comme une réalité concrète qui influence et est influencée par les autres formes de *sense-making*, notamment le *sense-making* biologique et le *sense-making* sensorimoteur (les symboles et les règles, comme nous le verrons à la section 3.3.3, émergent de l'activité de *linguaging*). Di Paolo, Cuffari et De Jaegher parlent de corps linguistiques à dessein: selon eux, ce qui est traditionnellement appelé « langage » est avant tout composé d'actions accomplies par des corps, en adéquation avec la thèse pragmatique de Brandom (1994) décrite dans la section précédente. Nous avons également vu dans la dernière section que la coordination des *sense-making* individuels permet la génération du *sense-making* participatif, rendant compte de la thèse sociale de Brandom (1994). Le *linguaging* est donc une affaire d'actes sociaux. Il reste maintenant à définir la normativité spécifique au *linguaging* afin de rejoindre la thèse « normative » de Brandom (1994). Ainsi, l'objectif de cette section sera de montrer comment le modèle du corps linguistique explique comment une normativité proprement « linguistique » peut émerger du *sense-making* participatif<sup>38</sup>. Pour ce faire, trois notions nous seront particulièrement utiles: le dialogue, la récursivité et les genres participatifs.

D'un côté, le *sense-making* participatif peut devenir un dialogue en raison de la tension entre différents participants d'une interaction sociale, chacun pouvant assumer les rôles d'agent régulateur et d'agent régulé (Di Paolo *et al.*, 2018). Dans une interaction sociale, l'agent régulateur est celui qui utilise un acte pour influencer les actions possibles de l'autre agent, lequel assume dès lors le rôle d'agent régulé. Cette tension découle du fait que les actes sociaux sont (la plupart du temps) partiels et co-définis, c'est-à-dire qu'ils en nécessitent d'autres pour être complets (par exemple, l'acte de donner dépend de l'acte de recevoir pour être complet). Ces actes partiels co-définis offrent un répertoire de réponses potentielles aux agents dans leur négociation entre leurs normes individuelles et les normes sociales et permettent de gérer les éventuelles ruptures de l'interaction sociale de multiples façons. Toutefois, les actes sociaux n'ont pas tous le même pouvoir d'influence sur l'interaction sociale, et certains actes sont particulièrement efficaces. Par exemple, pour moduler la distance entre deux agents (par exemple si un agent s'approche trop), un bras tendu avec la paume levée possède un fort pouvoir déontique, c'est-à-dire que l'autre agent est fortement encouragé à reculer pour respecter la distance établie par le bras tendu. Ce pouvoir déontique découle de la nature même du geste, qui bloque physiquement le passage.

---

<sup>38</sup> Dans *Linguistic Bodies* (2018), Di Paolo et ses collègues introduisent un long modèle dialectique pour expliquer le passage du concept de *sense-making* participatif au *linguaging*. À des fins de présentation, j'ai décidé d'éliminer la dimension dialectique du modèle de Di Paolo *et al.* (2018) pour me concentrer sur les notions essentielles, évitant les étapes intermédiaires.

L'aspect dialogique d'une interaction sociale émerge notamment du fait que chaque participant de cette interaction sociale peut assumer autant les rôles d'agent régulateur et d'agent régulé. Cela présage l'apparition d'une dynamique d'alternance entre les régulateurs et les régulés, qui peuvent tour à tour changer dynamiquement de rôle tout au long de l'interaction sociale. La possibilité de cette alternance des tours permet aux agents d'identifier les *utterances*<sup>39</sup> du dialogue. Une *utterance* consiste en une série d'actes sociaux qui compose un tour et qui peut être identifiée ainsi par les participants (Bakhtin, 1986; Di Paolo *et al.*, 2018; Haugeland, 1990). Di Paolo *et al.* (2018) précisent que leur utilisation du terme « *utterance* » est très générale et dépasse la parole verbale ou gestuelle. Ils illustrent leur propos par l'exemple d'un maître artisan qui interrompt son apprenti pour lui montrer une séquence d'action à reproduire correctement. Toute la séquence d'actes où le maître régule son apprenti (en captant son attention dans l'interaction sociale) compte comme une *utterance*, puisque la démonstration a une fonction régulatrice (indiquer à l'apprenti comme faire quelque chose). Les tours (c'est-à-dire le début et la fin des *utterances*) sont déterminés par un changement de la direction de l'influence, i.e. le régulateur devient le régulé (et vice-versa). C'est ainsi que le *sense-making* participatif prend la forme d'un dialogue, un échange entre des agents régulateurs et des agents régulés qui produisent à tour de rôle des *utterances* dans le but d'influencer l'interaction sociale.

De l'autre côté, Di Paolo, Cuffari et de Jaegher mettent l'accent sur le fait que les dialogues sont très souvent récursifs, c'est-à-dire que les *utterances* influencent souvent d'autres *utterances*. Par exemple, un concept important dans les processus de production et d'interprétation est celui d'*utterance* reportée. Les *utterances* reportées sont des *utterances* « that echo, reflect, refract, or somehow make use of other *utterances*, the producer's own or those of others » (Di Paolo *et al.*, 2018, p. 187). Cette forme d'*utterance* « documents, brings into the open, the producer's interpretation of the utterances it repeats or reflects » (Di Paolo *et al.*, 2018, p. 187). Par exemple, quelqu'un pourrait répéter l'*utterance* d'un producteur, mais plus lentement, pour mettre l'accent sur certaines parties de l'*utterance*, et ainsi modifier les possibilités d'interprétations pour réguler les interactions sociales. Imaginons deux personnes transportant un objet lourd et cherchant un endroit pour le poser. La première personne pourrait produire une *utterance* pour proposer un endroit (par exemple en pointant du regard), et la deuxième personne, en désaccord, pourra pointer au même endroit avec son regard, mais en levant les sourcils pour amener la première personne à réinterpréter sa proposition pour possiblement lui faire changer d'idée. Les *utterances* reportées permettent de modifier les possibilités d'interprétation des *utterances*, ce qui devient une manière de modifier le déroulement de l'interaction sociale.

---

<sup>39</sup> J'ai choisi de conserver le terme anglais *utterance* plutôt que de traduire par *énoncé* ou *énonciation* pour éviter la confusion avec le terme courant « énonciation » qui a tendance à exclure les gestes non-linguistiques. L'utilisation du terme *utterance* en français permet donc de comprendre que l'on a affaire à une notion théorique, de manière similaire au terme *affordance* en psychologie écologique.

Par ailleurs, cette récursivité du *sense-making* dialogique permet l'émergence d'un *sense-making* réflexif (Di Paolo *et al.*, 2018). Ce *sense-making* réflexif prend la forme d'un dialogue personnel, c'est-à-dire qu'un agent produit et interprète ses propres *utterances* dans le but de se réguler soi-même dans le cadre (potentiel) d'une interaction sociale. Par exemple, un enseignant peut remettre en question la clarté de son explication, en s'interprétant comme un étudiant qui interpréterait un professeur. Cela permet au professeur de s'ajuster pour permettre de meilleures interactions sociales. Il s'agit donc ici d'un contrôle de soi socialement déterminé (*social self-control*), où l'on se régule soi-même comme un autre, et on est régulé par soi-même comme par un autre (Di Paolo *et al.*, 2018).

Enfin, les dialogues permettent l'émergence de ce que Di Paolo, Cuffari et De Jaegher nomment des genres participatifs (Di Paolo *et al.*, 2018). Di Paolo, Cuffari et De Jaegher s'inspirent ici du concept de *speech genre* de Bahktin (1986) pour définir les genres participatifs comme des cadres normatifs qui (pré)structurent l'organisation des *utterances* en termes de durée, de style et d'expressivité. Autrement dit, un genre participatif est justement le cadre normatif spécifique à une interaction dialogique. Le genre structure les conditions normatives de la production et de l'interprétation des *utterances*. En effet, l'alternance à tour de rôle (entre agents régulateurs et agents régulés) génère une tension au niveau des *utterances* elles-mêmes entre la production et l'interprétation d'une *utterance*. (Di Paolo *et al.*, 2018) Par exemple, une *utterance* peut être « ouverte » en ce sens qu'elle accepte plusieurs réponses possibles, et c'est le rôle de l'interprétation de clarifier la fonction d'une *utterance* et l'intention de la personne l'ayant produite et de proposer une réponse appropriée. Réciproquement, le producteur de l'*utterance* doit tenir compte des possibilités d'interprétation (de réponses) dans le contexte de l'interaction sociale pour pouvoir jouer son rôle de régulateur.

La production et l'interprétation de l'*utterance* imposent des contraintes inverses: la production de l'*utterance* est facilitée si celle-ci est courte et partielle (plutôt qu'holistique et détaillée), tandis qu'au contraire l'interprétation de l'*utterance* est facilitée si cette dernière est longue et détaillée. Par exemple, un petit hochement de tête est très facile à produire, mais peut être difficile à interpréter selon les circonstances. Inversement, une longue démonstration d'une technique (comme celle du maître précédemment) peut être difficile à accomplir, mais facilite grandement l'interprétation (par l'apprenti, par exemple). Selon le contexte, différentes contraintes auront priorité. Dans le contexte de l'enseignement magistral d'un professeur, la clarté et le détail des *utterances* auront priorité sur la facilité de production. Lorsque les étudiants sortent de la classe pour aller se rejoindre dans un bar, le bruit ambiant peut faire en sorte que la priorité sera donnée aux *utterances* courtes et faciles à interpréter. Ainsi, selon les différents contextes, différentes normes encadrant la production et les interprétations peuvent émerger. Ce sont ces différents cadres normatifs que Di Paolo *et al.* (2018) nomment genre participatif. Plus généralement, la récursivité fait en sorte la régulation des genres participatifs devient une méta-régulation, c'est-à-dire que les *utterances* (comme moyen de régulation de l'interaction

sociale) vont réguler les paramètres normatifs qui vont contraindre les possibilités de régulation à l'intérieur de l'interaction sociale. Concrètement, cela signifie que les *utterances* peuvent servir non seulement à établir le genre participatif de l'interaction sociale, mais également à le modifier, influençant ainsi les *utterances* futures du discours

En bref, pour Di Paolo et ses collègues, le *linguaging* prend la forme d'un *sense-making* dialogique (souvent récursif) entre des producteurs et des interprètes d'*utterances* (des actes sociaux) qui co-régulent les interactions selon des cadres normatifs variés, les genres participatifs. A-t-on réussi à rendre compte de la thèse normative de Brandom (1994)? Je soutiens qu'à ce stade-ci, le modèle énonciviste fournit la base conceptuelle qui permet de présager l'émergence d'une normativité proprement linguistique. Toutefois, il reste encore un peu de travail à faire pour montrer l'aspect véritablement linguistique du modèle. Ce travail nous concernera dans la section suivante.

### 3.3.3. La grammaire énoncive: l'explication énonciviste de la syntaxe et des symboles

Dans la section précédente, j'ai présenté rapidement le modèle du corps linguistique de Di Paolo *et al.* (2018). Ce qu'ils nomment le *linguaging* est présenté comme étant un type de *sense-making* participatif qui prend la forme d'un dialogue entre plusieurs participants, et ce dialogue sert à la méta-régulation de leur interaction tout au long de leur échange. Toutefois, Di Paolo, Cuffari et De Jaegher sont conscients que cette formule peut désorienter certains lecteurs qui ne voient pas en quoi leur modèle du « *linguaging* » permet d'expliquer véritablement le langage. Ils reconnaissent carrément que « we have held language—as we typically, unreflectively know it and call it—at arm's length throughout the first and second parts of this book » (Di Paolo *et al.*, 2018, p. 279). Di Paolo, Cuffari et De Jaegher proposent alors une nouvelle manière de concevoir le langage (encore là distinct du *linguaging*): la *grammaire énoncive*. La grammaire énoncive est « fondée théoriquement sur la continuité entre l'esprit et le vivant » (Di Paolo *et al.*, 2018, p. 281) et cherche à expliquer comment la grammaire émerge grâce à l'agentivité des corps linguistiques.

D'une part, Di Paolo, Cuffari et De Jaegher vont argumenter que la notion de syntaxe peut être restituée par leur modèle comme étant une forme plus aboutie de méta-régulation des interactions sociales. L'argument rejoint l'idée d'une « grammaire émergente » (Hopper, 1987; Linell, 2009). Selon cette idée, les *utterances* (qui sont maintenant des énoncés linguistiques à proprement parler) ne sont pas construites en respectant les règles d'une syntaxe déjà formulée a priori dans l'esprit, mais c'est plutôt la syntaxe qui émerge lorsque les routines de réutilisation des *utterances* sont régularisées et deviennent de plus en plus conventionnelles. Cela est cohérent avec l'approche développée jusqu'à présent par le modèle dialectique de Di Paolo *et al.* (2018). Comme l'écrivent les auteurs, « grammar gets sedimented as reusable solutions to coordination problems in dialogic practices » (Di Paolo *et al.*, 2018, p. 288).

À titre d'exemple, dans la grammaire classique, la notion de proposition (grammaticale) désigne une unité d'interaction qui est complète en soi et prenant souvent la forme d'un prédicat associé à un syntagme nominal (ex. « Le chien est rouge. »). Ces unités sont utiles, car elles permettent aux interprètes de suivre les actions sociales qu'elles décrivent (Thompson & Coupler-Kuhlen, 2005), par exemple « Billy a volé le vélo de Joseph. ». Ces propositions sont souvent marquées par une ponctuation, divisant ainsi le discours en unités dont il est possible de comprendre le sens individuel. Selon le modèle dialectique de Di Paolo *et al.* (2018), ces ponctuations peuvent être comprises comme étant des moments à haut potentiel de régulation, c'est-à-dire qu'ils ont fortement tendance à inciter un certain type d'action de la part des membres du dialogue (souvent un changement dans la prise de tour). Ainsi, la ponctuation est avant tout un acte social qui sert à la méta-régulation de l'interaction en indiquant aux participants la fin d'une *utterance* et le début d'une autre. Cette fin facilite le changement de tour et le déplacement de l'influence au sein d'une interaction en diminuant les risques de rupture.

Plus généralement, la syntaxe émerge dans la mesure où elle facilite la méta-régulation du dialogue lors d'une interaction sociale, mais celle-ci peut dépasser le cadre strictement « linguistique ». En effet, Di Paolo, Cuffari et De Jaegher avancent qu'une grande panoplie d'activités sensori-motrices complexes (comme tricoter, préparer une recette, réparer un toit, etc.) « are hard to conceive or exercise without the action organization that linguistic bodies learn from grammatically structured practices » (Di Paolo *et al.*, 2018, p. 292). L'émergence des normes syntaxiques et leur apprentissage permettent d'apprendre encore plus finement l'organisation des schèmes sensorimoteurs, puisque la syntaxe permet des enchâssements, de la planification ou encore des variations dans les répétitions d'un même acte, tous des éléments qui sont utiles aux schèmes sensorimoteurs (Root, 1975). Bref, la syntaxe énonciviste ne prend pas la forme d'une représentation interne de règles grammaticales, mais émerge plutôt comme la complexification et la spécialisation des normes dans le *sense-making* participatif.

D'autre part, si l'énoncivisme refuse la notion de représentation mentale, il est difficile de soutenir que le langage (et d'autres activités culturelles<sup>40</sup>) n'ont pas recours à des représentations (externes). Cela, même les anti-représentationnistes les plus convaincus (par exemple Hutto & Myin, 2017) peuvent l'accorder: « some cognition is content-involving and (some) organisms become capable of content-involving cognition by mastering special sociocultural practices » (Hutto et Myin, 2017, p. xix; notons que les auteurs utilisent le terme « contenu » au sens de « contenu représentationnel »). Dans ce cas, peut-on expliquer l'émergence de l'utilisation de ce type de représentations externes grâce au modèle énonciviste? Di Paolo et ses collègues ne parlent pas spécifiquement de représentations mais vont argumenter qu'il est possible de restituer la notion de symbole grâce à leur modèle dialectique. S'il en

---

<sup>40</sup> Voir par exemple la description détaillée que fait Giere (1988) de l'usage socialement et culturellement normé de représentations schématiques en sciences.

existe plusieurs définitions, le concept de symbole renvoie souvent à un signe dont la signification est établie par une convention. Deacon (2011), d'inspiration sémiotique, classe les symboles en deux grandes catégories: les symboles non linguistiques, c'est-à-dire des artefacts dont la signification est investie culturellement, et les symboles linguistiques, c'est-à-dire des traces écrites ou orales qui peuvent être combinées à d'autres traces selon des règles explicites (une syntaxe). Dans les deux cas, ce qui semble distinguer le symbole des autres formes de signes est l'aspect apparemment arbitraire du lien entre le signifiant et le signifié. Comment restituer ces caractéristiques au sein du modèle énonciatif?

Déjà, Di Paolo *et al.* (2018) appellent à une conversion du concept de « symbole » pour le remplacer par l'idée de *symbolisation*. Selon Di Paolo, Cuffari et De Jaegher, ce que l'on nomme « symbole » découle des mêmes phénomènes qui permettent d'expliquer les autres formes de *sense-making*, à savoir une interaction dynamique entre le corps et l'environnement (ce dernier pouvant être constitué d'autres corps, c'est-à-dire des corps d'autres organismes, notamment dans le cas du *language*). De plus, les symboles sont avant tout matériels avant d'être idéaux (ou formels) puisque les symboles sont des *utterances* qui découlent du même modèle dialectique que d'autres actes sociaux, modèle qui est axé sur l'interaction entre des corps (de plus en plus linguistiques).

L'un des phénomènes permettant d'expliquer l'émergence de symboles de plus en plus arbitraires est celui de la réorientation de soi-même au travers des pratiques sociales sédimentées. Lorsqu'une *utterance* est utilisée hors de son contexte habituel, elle est investie d'un potentiel symbolique, c'est-à-dire qu'elle a le potentiel de créer momentanément ce que Di Paolo, Cuffari et De Jaegher nomment un micro-contexte qui redirige l'attention des participants vers des *utterances* ou des objets en particulier. Di Paolo et ses collègues s'inspirent à ce sujet des travaux de Rivière et Español (2003) qui proposent le concept de symbole énonciatif. Andrén (2017) propose un exemple pour comprendre le concept de symbole énonciatif: celui d'un jeune enfant pointant le briquet de ses parents et mimant le geste pour signifier qu'il désire voir du feu. Bien que l'enfant ne le fasse pas consciemment, il se produit un décalage entre l'effet initial du geste (produire du feu) et le nouvel effet. Comme le geste est fait dans le vide, il n'est pas efficace pour produire du feu. Plutôt, ce geste a des implications virtuelles, c'est-à-dire qu'il attire l'attention des parents vers la potentialité d'allumer le feu grâce au briquet. Survient donc une séparation entre le geste et son contexte d'usage usuel, et c'est cette séparation qui forme la base de la sédimentation de « symboles » tels que nous les connaissons. En effet, en considérant que l'effet des gestes (maintenant symboliques) peut être décalé de leur effet pragmatique immédiat, cela ouvre la porte à de nouvelles méthodes de régulation sociale, plus complexes et plus nuancées. À cet égard, les *utterances* reportées facilitent la séparation entre l'utilisation pragmatique et l'utilisation de plus en plus symbolique de ce qui deviendra les « symboles ». Les symboles peuvent par la suite réguler et contraindre les dialogues (et les dialogues personnels) permettant de re-présenter (en rendant présent de nouveau) des implications virtuelles. Di Paolo *et al.* (2018) l'expliquent ainsi:

Referring, then, is an emerging outcome of sense-making processes of linguistic bodies becoming together. The interplay of living use and given histories shows that in addition to grammaticalizing, linguistic bodies also symbolize and interact with symbols as « products » of this activity. Symbols re-present in a world-involving way for linguistic bodies who are sensitized to interactive moves and affordances through their codefined and complementary know-how. (Di Paolo *et al.*, 2018, p. 295)

C'est une inversion de l'approche de naturalisation décrite au premier chapitre. Rappelons-nous que le projet consistait en une double réduction: l'explication des symboles et représentations linguistiques par l'existence de représentations mentales, puis une explication naturaliste (par exemple biologique ou neurologique) de ces représentations mentales. Di Paolo et ses collègues affirment au contraire que ce que l'on appelle symbole émerge de l'interaction des corps linguistiques.

Plus généralement, les symboles énatifs sont vus comme des forces qui contraignent les actions par leur matérialité physique, une idée inspirée de la Théorie de l'engagement matériel (Malafouris, 2013). Comme le dit Malafouris: « Material signs are not simply message carriers in some pre-ordered social universe. Material signs are the actual physical forces that shape the social and cognitive universe (...) » (Malafouris, 2013, p. 97). Cette vision des symboles comme contrainte des interactions dynamiques trouve de l'inspiration également au sein des travaux de Pattee et Rączaszek-Leonardi (2012), ainsi que plus généralement au sein de l'œuvre d'Howard Pattee (Pattee & Kull, 2009; Pattee, 2012). À ce propos, Di Paolo, Cuffari et De Jaegher vont affirmer que l'idée de symbole comme contrainte « leads Rączaszek-Leonardi to a perspective on languaging close to our own, as well as an ultimately non representational approach to explaining representational acts » (Di Paolo *et al.*, 2018, p. 295). Toutefois, il resterait à déterminer à quel point les deux approches sont véritablement compatibles, mais ce travail dépasse le cadre de ce mémoire.

Évidemment, Di Paolo et ses collègues reconnaissent que leur modèle est plus une ébauche qu'une théorie scientifique en bonne et due forme. Ils expliquent que la grammaire énative devrait être comprise comme « a general concept that further work could link with more specific aspects of grammar as studied by linguists. » (Di Paolo *et al.*, 2018, p. 281) Ainsi, la théorie du corps linguistique ne permet pas dans son état actuel d'expliquer des phénomènes linguistiques qui sont étudiés par des linguistes. On peut néanmoins apprécier l'entreprise de Di Paolo, Cuffari et De Jaegher, qui veulent paver la voie pour une nouvelle conception (énactive) du langage, et c'est ce travail de défrichage qui est accompli dans *Linguistic Bodies* (2018).

En bref, une manière d'expliquer l'intentionnalité représentationnelle dans un cadre énativiste passe 1) par l'adoption d'une conception néo-pragmatique de l'intentionnalité représentationnelle et 2) par le modèle du corps linguistique comme possibilité d'implémentation énativiste de cette conception. Ce modèle propose que l'intentionnalité représentationnelle tire son origine de la socialité, qui prend la

forme d'un *sense-making* participatif au sein de l'énactivisme. Les participants apprennent à réguler non seulement les interactions sociales et les autres participants, mais ils apprennent également à réguler la régulation elle-même. Le *linguaging* est justement une forme de *sense-making* qui permet cette méta-régulation de l'interaction sociale. Enfin, Di Paolo, Cuffari et De Jaegher proposent que la syntaxe et les symboles peuvent être expliquées comme étant des moyens de faciliter la modulation des interactions entre plusieurs participants et entre plusieurs groupes. C'est comme cela que Di Paolo *et al.* (2018) proposent d'expliquer le langage humain dans des termes énonctiviste. Le tout est accompli sans pour autant prétendre que le modèle est complet ou qu'il remplace les travaux empiriques de la linguistique. En effet, les auteurs reconnaissent que leur modèle est en quelque sorte une esquisse et qu'à ce propos, il ne peut pas expliquer les phénomènes spécifiques qui intéressent les linguistes. Néanmoins, ils soutiennent que leur modèle est une première étape pour intégrer le langage au sein de la continuité entre l'esprit et le vivant.

### 3.4. Discussion du modèle du corps linguistique

Dans la section précédente, nous avons rapidement exposé la solution que propose l'énactivisme pour répondre au défi du *scaling-up*, soit le modèle du corps linguistique, qui décrit le comportement langagier humain comme étant une forme de *sense-making* dialogique émergeant de la tension entre différents participants d'une interaction sociale. Dans cette section, je m'intéresserai à trois enjeux liés au modèle du corps linguistique développé par Di Paolo *et al.* (2018). D'abord, je chercherai à déterminer comment comprendre la notion de contenu intentionnel au sein du modèle du corps linguistique. Ensuite, j'exposerai le modèle énonctiviste à certaines critiques des modèles néo-pragmatistes pour comprendre comment l'énactivisme permet de répondre à ces critiques. Enfin, je discuterai de la proposition de Rolla & Huffermann (2022) à l'effet d'intégrer le modèle téléosémiotique de Hutto et Myin (2017) de l'énactivisme radical au modèle du corps linguistique pour y ajouter une dimension diachronique.

#### 3.4.1. Quelle notion de contenu intentionnel pour le modèle du corps linguistique?

Les discussions autour du problème du *scaling-up* de l'énactivisme permettent de faire ressortir la tension qui existe autour de la notion de contenu intentionnel au sein de l'énactivisme en général. Rappelons que dans l'introduction, j'ai introduit la terminologie de Rolla & Huffermann (2022) qui distinguait le contenu basique (déterminé par des conditions de succès), le contenu représentationnel (déterminé par des conditions d'exactitudes) et le contenu propositionnel (déterminé par des conditions de vérité). Je propose d'interpréter la position de Di Paolo *et al.* (2018) ainsi: Le contenu intentionnel d'une *utterance* est un contenu basique dont les conditions de succès dépendent de l'effet de l'*utterance*

sur la régulation (normative) de l'interaction sociale dans laquelle elle se situe. Analysons un passage de *Linguistic Bodies* pour mieux comprendre leur position:

All sense-making is inherently world-relating and world-involving; moreover, embodied, social acts have real effects for sense-makers that can take forms more or less recognizable as answering a question of content. 'Aboutness' is not limited to propositions. It lives in embodied intentionality of sensorimotor and intersubjective bodies going about their business, in acts of guiding, modifying, or even projecting others' acts (...) (Di Paolo *et al.*, 2018, p. 201).

Déjà, ils affirment que la notion d'*aboutness* (le caractère 'à propos de' caractéristique de l'intentionnalité) n'est pas limitée aux phrases ou représentations exprimant des propositions. Cette thèse n'est pas très surprenante compte tenu des chapitres précédents, où j'ai tenté de montrer que la directionnalité n'était pas que l'affaire de représentations exprimant des propositions, comme le pensaient les partisans du projet original de naturalisation de l'intentionnalité, mais qu'elle se manifestait aux niveaux biologiques et sensorimoteurs. Jusqu'ici donc, rien de nouveau. Toutefois, là où cela devient intéressant, c'est lorsque les Di Paolo *et al.* (2018) affirment que la notion de « contenu » peut être comprise comme étant l'effet de certaines *utterances*. Rappelons que les *utterances* servent à réguler la normativité d'une interaction sociale. Pour reprendre un exemple donné précédemment, lever le bras pour instituer une distance physique avec un autre participant lors d'une interaction sociale est une *utterance*, et son contenu est l'effet de réguler le déroulement de l'interaction en modifiant ses paramètres normatifs (la distance adéquate pour le déroulement de l'interaction), ce qui affecte également les possibilités de réponse de l'autre participant. Le contenu intentionnel de cette *utterance* est lié à son effet sur le déroulement de l'interaction, soit que la personne à qui l'acte est adressé ajuste sa distance adéquatement de manière à maintenir l'interaction (bref le contenu est lié à la maintenance de l'organisation précaire de la situation d'interaction: si la personne s'avance trop, il y aura conflit; si elle s'éloigne trop, il sera impossible pour les participants de se voir ou de s'entendre). J'avance que ce contenu est un contenu basique (au sens de Rolla & Huffermann, 2022) qui dépend de conditions de succès (la personne ajuste sa distance en répondant au bras comme souhaité). En suivant Di Paolo, Cuffari et De Jaegher, le contenu intentionnel d'une *utterance* n'est pas déterminé par des conditions d'exactitude (ou de vérité), mais dépend plutôt de conditions de succès quant à la régulation des interactions sociales, ce qui en fait un contenu basique. Toutefois, ce contenu basique est plus sophistiqué que le contenu basique biologique ou le contenu basique pragmatique, en ce sens que les conditions de succès sont déterminées par des normes sociales, ce qui ajoute une complexité dans la détermination des conditions de succès.

À cet égard, ce qui manque au modèle du corps linguistique pour véritablement naturaliser l'intentionnalité représentationnelle, c'est d'expliquer comment les conditions de succès qui déterminent

le contenu d'une *utterance* (visant la régulation d'une interaction sociale) peuvent devenir des conditions d'exactitude, puis enfin des conditions de vérité. Comment procéder?

Je soutiens qu'une partie de la réponse se trouve au sein des inspirations néo-pragmatistes du modèle du corps linguistique. D'abord, je veux montrer que le modèle du corps linguistique rejoint bel et bien la thèse CD du néo-pragmatisme, c'est-à-dire que le contenu intentionnel des *utterances* est déterminé par des normes sociales. Rappelons-nous qu'à la section 3.2., j'ai suggéré que le modèle du corps linguistique s'inspirait des idées néo-pragmatistes (Haugeland, 1990) selon lesquelles l'intentionnalité originelle est issue des interactions sociales, et que l'intentionnalité privée (comme celle des représentations mentales) en est dérivée. J'y ai notamment présenté deux thèses alternatives quant au lien entre la normativité sociale et le contenu propositionnel: 1) la thèse CE qui veut que le contenu propositionnel ait des conséquences normatives 2) la thèse CD qui veut que le contenu propositionnel soit déterminé par les normes sociales (Gluër, Wikfross et Ganapini, 2009). J'avais ajouté que la thèse CD implique la thèse CE (Tison, 2022). De quelle manière le modèle du corps linguistique rejoint-il la thèse CD (et, du même coup, par implication, la thèse CE)?

Commençons par la thèse CD. Une manière simple de répondre à cette question est de rappeler que le *linguaging* est avant tout une forme de *sense-making* participatif, qui lui-même consiste en la négociation entre les normes individuelles (biologiques et sensorimotrices) et sociales (propres à la forme particulière de l'interaction sociale en cours). Une *utterance* est un acte qui vise à réguler la normativité de l'interaction sociale, et si on accepte que le contenu basique d'une *utterance* consiste en ses effets, on peut comprendre que l'effet de l'acte d'*utterance* dépend de la normativité des interactions sociales qu'il vise à réguler. Par exemple, le bras tendu qui incite l'autre personne à reculer modifie les conditions normatives de l'interaction sociale (i.e., à réguler la distance entre les participants), et c'est en vertu de l'existence de ces conditions normatives (une distance socialement appropriée entre les participants) que cet acte social peut les réguler ou les modifier. Si ces normes n'existaient pas, il n'y aurait rien à réguler. Le fondement du contenu basique de toutes les *utterances* (qui sont des actes sociaux) est ainsi l'existence d'une normativité propre aux interactions sociales qui ne se réduit pas aux normes individuelles de survie ou d'action sensorimotrice des organismes. Quant à la thèse CE, comme les actes ont des effets sur la régulation de la normativité de l'interaction, il est possible de dire que les *utterances* ont des conséquences normatives puisqu'elles affectent les possibilités d'actions des autres participants (en modifiant les autres actes sociaux). Par exemple, mon bras tendu modifie les conditions normatives de l'interaction sociale, et l'autre ne devrait pas chercher à outrepasser mon bras tendu, puisque cela risque de causer une rupture de l'interaction sociale. On respecte ainsi la cohérence de l'idée selon laquelle la thèse CD implique la thèse CE.

Évidemment, les théories néo-pragmatistes ne sont pas exemptes de critiques, et il est important de savoir si ces critiques peuvent affecter la crédibilité du modèle du corps linguistique. Pour cette section, je me concentrerai sur Cash (2009) qui considère deux objections présentes dans la littérature<sup>41</sup>: 1) les normes sociales seraient impropres à fonder le contenu, puisqu'elles dépendraient d'un désir de se conformer; 2) elles ne permettent pas d'expliquer l'efficacité causale du contenu.

La première objection faite aux théories néo-pragmatistes est que le conformisme n'est pas une bonne fondation pour les théories néo-pragmatistes, puisque le conformisme est basé sur des désirs contingents qui pourraient exister ou ne pas exister (Hattiangadi, 2006). Si le désir disparaissait, le contenu disparaîtrait également. La réponse de Cash (2009) est assez simple. Selon lui, il existe plusieurs éléments de preuves suggérant que le conformisme est un mécanisme répandu au sein de l'espèce humaine et que celui-ci ne dépend pas des désirs contingents de certains individus ou de certaines communautés (Kochanska, Forman, Aksan, & Dunbar, 2005; voir également Wright, 1982; 1983 pour les mécanismes à l'œuvre chez les jeunes enfants). Dans le cadre du modèle du corps linguistique, on pourrait argumenter que les agents ne cherchent pas à se conformer parce qu'ils ont un désir de se conformer, mais plutôt parce que cela permet le maintien des interactions sociales. Rappelons que le *sense-making* participatif (et par extension, le *linguaging*) est une négociation entre les normes individuelles et sociales. Les participants ne se conforment pas de manière aveugle. Néanmoins, les participants sont toujours en interaction dans les cadres normatifs des interactions sociales (quitte à chercher à les modifier), et c'est ce qui permet de fonder le contenu des *utterances* comme les effets sur les conditions normatives des interactions sociales.

La deuxième objection est que le néo-pragmatisme n'est pas capable d'expliquer le rôle causal du contenu puisque ce dernier dépendrait de normes externes à l'individu (Boghossian, 2003). Cette objection découle d'une conception du problème de la naturalisation de l'intentionnalité similaire à celle de Dretske (1988; 1990) et de Ramsey (2007) que nous avons décrite dans la section 2.1. Si la réponse de Cash (2009) ne nous intéresse pas tellement dans le contexte du mémoire - il cherche à montrer qu'il est possible d'expliquer par le néo-pragmatisme l'existence de représentations neuronales - j'avance qu'il est relativement aisé d'imaginer une réponse énonciviste à cette objection. En effet, si on suit les énoncivistes, ce qu'on appelle le contenu intentionnel est lié au produit de l'activité du *sense-making* (les effets sur l'interaction sociale) plutôt qu'à sa cause. Il en résulte que la question de l'efficacité causale du contenu ne se pose pas vraiment. En rejetant la notion de représentation mentale (et la conception cartésienne de l'esprit qui voit les représentations mentales comme causes du comportement; cf van

---

<sup>41</sup> Dans son article, Cash (2009) soulève deux autres objections, soit 3) le néo-pragmatiste est anti-naturaliste et 4) le néo-pragmatisme sur-intellectualise le fait de posséder des états intentionnels. Par souci de concision, j'ai décidé de laisser ces objections de côté.

Gelder 1995), les énoncivistes évitent ainsi le défi de la description de tâche de Ramsey et n'ont pas à expliquer l'effet causal du contenu intentionnel, du moins pas au sens de Ramsey et Dretske.

Après avoir montré la similitude entre le néo-pragmatisme et le modèle du corps linguistique, il est permis de poser la question: est-ce que le modèle du corps linguistique permet d'expliquer adéquatement le passage des conditions de succès aux conditions d'exactitude et aux conditions de vérité? Dans son état actuel, il faut admettre que la réponse est non. Toutefois, quelques pistes de solutions permettent de penser que l'énoncivisme pourrait, un jour, compléter le travail. Pour ce faire, le modèle du corps linguistique devrait élaborer la manière dont les conditions normatives du déroulement des interactions sociales peuvent donner lieu à des conditions d'exactitude et, éventuellement, des conditions de vérité dans l'utilisation des *utterances*. À ce propos, deux pistes pourraient être suivies. D'une part, le modèle pourrait élaborer son explication des symboles énoncivifs pour expliquer comment une *utterance* reportée peut être utilisée correctement pour décrire une situation présente ou absente. L'effet de l'*utterance* serait, ici, une compréhension de la situation qui peut être jugée exacte ou non (conditions d'exactitude), leur permettant d'agir adéquatement grâce à cette compréhension. D'autre part, le modèle énonciviste pourrait élaborer son explication et sa description de sa grammaire énoncive pour pouvoir rendre compte de notions comme celles de prédication ou de négation. Cette explication, s'inspirant des théories néo-pragmatistes permettrait d'envisager une explication des conditions de vérité des *utterances* dans le modèle énonciviste. Toutefois, en acceptant que le modèle du corps linguistique ne soit pour l'instant qu'une ébauche, il est permis de pardonner ces limites de la théorie de Di Paolo *et al.* (2018), tout comme il est permis d'envisager la complétion du projet.

### 3.4.2. Une dimension diachronique grâce à l'énoncivisme radical?

Tel qu'expliqué précédemment à la section 3.4.1, le modèle du corps linguistique n'est pas un modèle diachronique (c'est-à-dire qui est ontogénétique ou évolutionnaire sur le plan biologique ou culturel) du développement progressif du langage d'un point de vue anthropologique (même si certaines formulations de Di Paolo et de ses collègues peuvent laisser croire le contraire...). Est-il possible d'ajouter une dimension diachronique au modèle de Di Paolo *et al.* (2018) pour le rendre plus complet?

À ce propos, Rolla & Huffermann (2022) proposent de combiner le modèle du corps linguistique avec l'approche téléosémiotique de l'énoncivisme radical (Hutto & Myin, 2013; 2017; Hutto & Satne, 2015). Rolla & Huffermann (2022) argumentent que malgré les différences théoriques entre l'énoncivisme radical et l'énoncivisme classique, il est possible de combiner les approches si on accepte la définition de contenu basique qui implique des conditions de succès. Cette proposition est similaire à celle du mémoire qui affirme que l'intentionnalité biologique est fondamentalement normative. Qui plus est,

Hutto & Myin (2017) se réclament explicitement du néo-pragmatisme, ce qui facilite la combinaison avec le modèle du corps linguistique.

Hutto & Satne (2015) et Hutto & Myin (2017) proposent une explication diachronique du développement du langage basée sur l'échafaudage socioculturel. Cet échafaudage socioculturel peut être initié par des mécanismes d'imitation et d'apprentissage social. Dans les mots de Hutto & Myin (2017), « the capacities in question can be understood in biological terms as mechanisms through which basic minds are set up to be set up by other minds and to be set off by certain things. » (Hutto & Muin, 2017, p. 139). Autrement dit, certains mécanismes pourraient être sélectionnés pour faciliter l'attention envers les congénères. L'émergence d'une norme sociale pour réguler les interactions sociales pourrait dès lors être décrite selon le modèle du corps linguistique de Di Paolo *et al.* (2018). Pour les formes plus spécifiques d'intentionnalité représentationnelle, Hutto & Myin indiquent qu'elles peuvent émerger lorsque des pratiques sociales ont développé des systèmes symboliques publics en même temps que des niches socioculturelles particulières.

À cet égard, il est possible de citer Satne (2016) qui propose une relecture énonciviste radicale de Tomasello (2014). Tomasello propose en effet deux étapes pour rendre compte de la culture chez les humains: 1) le développement de l'intentionnalité conjointe, i.e. l'intentionnalité sociale à échelle locale et 2) le développement de l'intentionnalité collective, i.e. la collaboration à large échelle impliquant la création d'un fonds culturel commun (Tomasello, 2014). Satne (2016) propose trois critères essentiels pour le développement d'une intentionnalité collective: 1) la conformité sociale, 2) une coordination qui permet aux individus de se conformer au comportement du groupe et 3) la possibilité d'identifier l'approbation ou la désapprobation du groupe. Ces critères reprennent, d'une certaine manière, les mécanismes de conformisme décrits par Haugeland (1990). Selon Satne, ces critères ne nécessitent pas de suivre des règles, mais plutôt de pouvoir s'ajuster (« *to tune itself* ») au comportement des autres.

Selon Rolla & Huffermann (2022), cette caractérisation des éléments nécessaires à l'évolution d'une intentionnalité collective (et plus généralement d'une intentionnalité représentationnelle) est compatible avec le modèle du corps linguistique de Di Paolo *et al.* (2018). Le premier élément (le conformisme) peut se traduire par la régulation de l'interaction sociale pour permettre son maintien dans le temps. Le deuxième élément (la coordination sociale) peut émerger de la méta-régulation de plusieurs participants d'une interaction sociale. Enfin, le troisième élément (l'approbation) peut se décliner de plusieurs manières, notamment grâce aux actes d'imbrication qui permettent de fortement réguler le comportement d'autrui. Rolla & Huffermann (2022) soutiennent que le modèle du corps linguistique offre un modèle synchronique, tandis que le modèle énonciviste radical offre une explication diachronique.

Si l'idée d'ajouter une dimension diachronique au modèle du corps linguistique est intéressante pour rendre le modèle plus complet, je ne suis pas convaincu que l'explication de Hutto, Satne et Myin soit la meilleure candidate pour y arriver. Au mieux, leur modèle n'est qu'une esquisse d'un véritable modèle évolutif du langage et à cet égard, les énaquistes radicaux ne sont pas plus avancés que Di Paolo et ses collègues. Pour accomplir un véritable travail diachronique, il faudrait chercher à combiner véritablement les modèles énaquistes synchroniques comme celui de Di Paolo avec des théories diachroniques du développement du langage, ce qui ne semble pas fait jusqu'à maintenant.

Évidemment, construire un véritable modèle diachronique énaquiste du langage dépasse largement le cadre de ce mémoire. Toutefois, l'ébauche de ce modèle diachronique du langage est relativement prometteuse, et cela est de bon augure pour défendre la thèse que l'énaquisme permet de naturaliser (tant du point de vue synchronique que diachronique) l'intentionnalité représentationnelle.

### 3.5. Conclusion : une ébauche pour la naturalisation de l'intentionnalité représentationnelle

Nous avons vu que l'énaquisme, comme approche anti-représentationnaliste, a été accusé de souffrir d'une incapacité à expliquer les capacités cognitives dites « *representation-hungry* » (Clark & Toribio, 1994). Cette objection a été surnommée « le problème du *scaling-up* » (Gallagher, 2017). Le problème consiste principalement à expliquer les capacités cognitives qui font intervenir des choses absentes ou imaginaires (des phénomènes où la notion de représentation devient très utile). Une des manières proposées par l'énaquisme est le modèle du corps linguistique, qui vise à expliquer le langage humain comme une capacité de méta-régulation du *sense-making* participatif, c'est-à-dire le *sense-making* qui émerge des interactions sociales. Ces pratiques de méta-régulation peuvent, par exemple, spécifier les normes du discours sur les objets abstraits, absents ou inexistantes). J'ai tenté de montrer que ce modèle reprenait la stratégie néo-pragmatiste d'explication de l'intentionnalité représentationnelle, c'est-à-dire que l'intentionnalité originelle est issue des interactions sociales (et des normes socioculturelles) pour par la suite être internalisées au plan individuel, renversant ainsi la stratégie de naturalisation du cognitivisme.

Toutefois, les critiques abordées permettent de saisir que ce modèle est une ébauche intéressante, mais qu'il ne fait pas office de théorie complète pour l'instant. J'ai évalué (en suivant Rolla & Huffermann 2022) la possibilité de combiner l'explication synchronique du modèle du corps linguistique avec l'explication téléosémiotique de Hutto & Myin (2017). Si les deux modèles sont chacun encore au stade d'ébauche, il n'en demeure pas moins que le potentiel d'explication est présent. Qui sait, l'avenir nous réservera peut-être une explication complètement énaquiste de l'intentionnalité linguistique et d'autres phénomènes socioculturels.

## CONCLUSION

À travers ce mémoire, j'ai cherché à comprendre comment l'énactivisme entendait-il naturaliser l'intentionnalité, une tâche très difficile où de nombreuses théories ont échoué. La thèse principale de ce mémoire était que l'énactivisme entend naturaliser l'intentionnalité en (re)définissant le problème comme celui d'une intégration des intentionnalités biologique, pragmatique et représentationnelle.

La première tâche de ce mémoire était donc de montrer l'existence dans la littérature de ces différents concepts d'intentionnalité, ce qui a été accompli au sein du premier chapitre. Traditionnellement, la notion d'intentionnalité est comprise comme étant le caractère directionnel de certains états: ceux-ci sont dirigés vers quelque chose. Aussi, l'intentionnalité est considérée comme étant l'une des propriétés fondamentales de l'esprit humain (l'autre étant la conscience phénoménale, volontairement laissée de côté non seulement en raison de la difficulté de la question, mais aussi parce qu'elle est presque systématiquement mise de côté par la littérature sur la naturalisation du contenu en sciences cognitives). À travers le premier chapitre, j'ai tenté de montrer qu'il n'y avait pas qu'un seul concept d'intentionnalité dans la littérature, mais qu'il en existait bien plusieurs. Pour éviter de rédiger le catalogue de l'ensemble des concepts sur le marché (une tâche dépassant grandement l'échelle d'un mémoire de maîtrise), j'ai proposé une catégorisation tripartite: 1) l'intentionnalité représentationnelle; 2) l'intentionnalité pragmatique; 3) l'intentionnalité biologique.

L'intentionnalité représentationnelle conçoit la directionnalité en termes de représentations, c'est-à-dire que les représentations représentent toujours quelque chose, soit leur contenu. Dans le contexte de l'esprit humain, l'intentionnalité est expliquée en postulant l'existence de représentations mentales. Selon cette conception, naturaliser l'intentionnalité revient donc à réduire les représentations mentales à des descriptions non représentationnelles, le plus souvent des descriptions neurologiques ou physiologiques. Si cette conception a été historiquement très populaire, le premier chapitre avait également comme but de montrer l'existence de conceptions non représentationnelles de l'intentionnalité. C'est ainsi que j'ai présenté la deuxième catégorie d'intentionnalité: l'intentionnalité pragmatique. L'idée derrière la notion d'intentionnalité pragmatique est que l'intentionnalité de l'esprit peut être comprise comme étant la directionnalité des actions du corps, qui visent des éléments (des objets, des situations, des conditions, etc.) de l'environnement. J'ai présenté succinctement les concepts d'intentionnalité motrice chez Merleau-Ponty et d'intentionnalité opérante chez Gallagher comme étant des exemples de notions pouvant être placées dans la catégorie de l'intentionnalité pragmatique. Enfin, j'ai présenté un troisième type d'intentionnalité, encore plus fondamentale que les deux autres: l'intentionnalité biologique. L'intentionnalité biologique a été décrite comme étant la directionnalité fondamentale du vivant envers son environnement. En effet, en s'inspirant des travaux de Jonas et de

Schrödinger, j'ai tenté de montrer que même au niveau organique fondamental il était possible de retrouver une directionnalité minimale qui prend la forme d'une ouverture sélective vers l'environnement. La nécessité de cette ouverture sélective était explicable par la nature des êtres organiques, qui sont éloignés de l'équilibre thermodynamique et qui doivent continuellement lutter contre l'augmentation de leur entropie.

Après que le premier chapitre ait présenté les différents types d'intentionnalités, le deuxième chapitre avait comme objectif de s'attaquer à la question principale du mémoire, à savoir comment l'énactivisme entendait-il naturaliser l'intentionnalité? Pour ce faire, j'ai d'abord présenté des tentatives passées de naturalisation de l'intentionnalité selon la stratégie de réduction évoquée dans le premier chapitre pour en montrer les limites (mais aussi parce que certaines de ses stratégies seront reprises par certains énaquistes). Ainsi, j'ai abordé les stratégies téléosémantiques de Dretske et Millikan, deux des théories les plus populaires pour tenter de résoudre le problème de la naturalisation de l'intentionnalité (représentationnelle). Les deux théories cherchaient à expliquer les représentations mentales des êtres vivants en termes de fonctions biologiques. Ainsi, selon ces théories, une structure interne capable de covarier avec une condition du monde externe devient une représentation mentale lorsque cette structure acquiert la fonction de représenter cette condition. Le problème revient alors à déterminer comment ces structures acquièrent cette fonction. Dretske propose une solution ontogénétique à ce problème, tandis que Millikan propose une solution phylogénétique. Toutefois, malgré l'attrait et la popularité de ces théories, j'ai montré que celles-ci faisaient face à des critiques qui mettaient en doute la perspective de les voir régler le problème de la naturalisation de l'intentionnalité.

C'est ici qu'est entré en scène l'énactivisme. L'énactivisme est une théorie anti-représentationnaliste de la cognition qui propose de concevoir l'esprit comme n'étant pas simplement l'affaire du cerveau, mais plutôt le résultat d'une interaction entre le cerveau, le corps et l'environnement. J'ai présenté les concepts fondamentaux de l'énactivisme, notamment l'autonomie et le *sense-making*. L'autonomie est la propriété de systèmes ayant deux caractéristiques: 1) ils sont opérationnellement fermés (ils forment une unité distincte de dont les composantes et processus internes sont interdépendantes) et 2) ils sont précaires, c'est-à-dire que la rupture d'un processus ou la destruction d'une composante peut entraîner la disparition du système. L'autonomie implique l'établissement de normes biologiques qui permettent à l'organisme de persister. J'ai expliqué que le concept d'autonomie (en l'agencant avec le concept d'adaptativité) impliquait le concept de *sense-making*, l'activité de « faire-sens » de l'organisme. Le *sense-making* de l'organisme est l'énaction (la création ou *bringing forth* par le biais d'une action) d'une évaluation du milieu ambiant, ce dernier gagnant ainsi une valeur en fonction des normes établies par l'autonomie de l'organisme.

Après avoir présenté le cadre théorique de l'énactivisme, j'ai entamé l'explication de la stratégie d'intégration des types d'intentionnalité. La première étape de cette stratégie est de naturaliser l'intentionnalité biologique. J'ai montré que cette étape est réalisée relativement facilement par l'énactivisme, c'est-à-dire que le *sense-making* (fondé sur les concepts d'autonomie et d'adaptativité) permet de rendre compte de l'intentionnalité biologique telle que décrite dans le premier chapitre. En effet, le *sense-making* est une activité d'évaluation de l'environnement qui passe par l'action de l'organisme. L'activité de l'organisme (le *sense-making*) est toujours dirigée vers le monde selon les valences déterminées par les normes de l'autonomie. Le concept d'agentivité (l'autonomie et l'asymétrie interactionnelle) s'est avéré utile pour formaliser la notion d'action de l'organisme sur son environnement.

J'ai par la suite montré que l'énactivisme cherchait à étendre ses concepts d'autonomie et de *sense-making* au niveau sensorimoteur pour naturaliser l'intentionnalité pragmatique. Pour ce faire, j'ai introduit les différentes notions de contingences sensorimotrices telles que décrites par Di Paolo *et al.* (2017) et notamment la notion de schème sensorimoteur. J'ai tenté de montrer que non seulement il y avait une autonomie au niveau des schèmes sensorimoteurs, mais que les schèmes témoignent d'une intentionnalité pragmatique en ce sens que ceux-ci visent toujours quelque chose, que ce soit des éléments extérieurs ou encore d'autres schèmes sensorimoteurs. Ainsi, j'ai tenté de montrer que l'énactivisme pouvait naturaliser l'intentionnalité pragmatique en montrant que l'intentionnalité pragmatique pouvait être expliquée par les mêmes concepts que l'intentionnalité biologique. Le deuxième chapitre se terminait ainsi avec la conclusion que l'énactivisme était au moins capable d'intégrer l'intentionnalité biologique et l'intentionnalité pragmatique dans son cadre théorique pour les naturaliser.

Enfin, le troisième et dernier chapitre abordait la question plus difficile de la naturalisation de l'intentionnalité représentationnelle par l'énactivisme. J'ai ouvert le chapitre sur un problème qui revient de manière récurrente chez les critiques de l'énactivisme: le problème du *scaling-up*. J'ai présenté le problème tel que formulé par Clark et Toribio (1994), c'est-à-dire que l'énactivisme serait inadéquat pour traiter des phénomènes qui impliquent une cognition *representation-hungry*, c'est-à-dire des actes cognitifs qui impliquent de raisonner avec des éléments imaginaires, absents ou abstraits, des phénomènes qui rendent le concept de « représentation mentale » très utile pour l'explication.

Pour réaliser la naturalisation de l'intentionnalité représentationnelle par l'énactivisme, j'ai présenté une solution qui, à mon avis, présente un bon potentiel: le modèle du corps linguistique de Di Paolo, Cuffari et de Jaegher (2018). Le modèle propose d'expliquer le langage à l'aide des concepts énaclivistes comme le *sense-making* et l'autonomie. J'ai argumenté qu'il était possible de rapprocher ce modèle énacliviste des stratégies néo-pragmatistes de l'explication de l'intentionnalité représentationnelle. Ces stratégies

défendent que l'intentionnalité représentationnelle tire son origine des normes sociales, notamment grâce à des mécanismes de conformisme. Plus généralement, j'ai mis en valeur trois thèses du néo-pragmatisme: l'action, la socialité et la normativité. Mon but était alors de montrer que le modèle énonciviste du corps linguistique pouvait rendre compte de ces trois thèses.

D'une part, il était facile de montrer que le modèle est pragmatique, puisqu'il se construit sur les bases du *sense-making* sensorimoteur. D'autre part, j'ai tenté de montrer que le modèle était social en expliquant la notion de *sense-making* participatif. Le *sense-making* participatif désigne l'activité des participants qui coordonnent leurs *sense-making* sensorimoteurs de telle manière à ce qu'une interaction sociale autonome (opérationnellement close et précaire) dotée d'une normativité propre émerge de cette coordination. Le *sense-making* participatif consiste alors en une négociation entre les normes individuelles et sociales.

Enfin, le plus difficile était de montrer que le modèle du corps linguistique permet un nouveau type de normativité, propre aux interactions linguistiques. Pour ce faire, j'ai présenté rapidement les éléments principaux du modèle. Nous avons vu qu'une interaction sociale devient progressivement « linguistique » lorsque la dynamique prend la forme d'un dialogue, c'est-à-dire que les participants s'échangent les rôles d'agent régulateur et d'agent régulé. Cet échange dialogique permet d'identifier des *utterances*, c'est-à-dire des séquences d'actions qui visent à modifier le cadre normatif dans lequel se déroule l'interaction sociale. J'ai également présenté l'argument de Di Paolo et de ses collègues à l'effet que l'*utterance* possédait un contenu basique dont les conditions de succès dépendent de l'effet de l'*utterance* sur la régulation normative de l'interaction sociale. Rappelons-nous de la personne qui tend son bras pour inciter l'autre à reculer. Ici, le « contenu » de l'*utterance* est de pouvoir réguler les conditions normatives de l'échange (la distance), et les conditions de succès du geste impliquent que la personne respecte le bras de distance. Dans ce contexte, les « représentations » (externes seulement) émergent sous forme de symboles énoncivistes, c'est-à-dire que certains gestes voient leur effet normatif modifié et décalé par rapport à leur effet pragmatique. Par exemple, l'enfant qui mime le geste du briquet dans le vide (le pouce qui glisse vers le bas) pour inciter ses parents à lui montrer une flamme du briquet. Le geste est décalé par rapport à son effet pragmatique normal (allumer un feu) et gagne un nouvel effet: celui d'inciter les autres à lui montrer du feu. Enfin, une grammaire énonciviste émerge comme étant la sophistication des normes des échanges dialogiques, c'est-à-dire que les normes qui encadrent les dialogues deviennent de plus en plus complexes et partagées parmi les participants pour faciliter les échanges sur des sujets de plus en plus variés et nuancés.

J'ai enfin conclu le chapitre en présentant une courte discussion de la pertinence et les faiblesses possibles du modèle. D'un côté, j'ai montré que le modèle pouvait résister à certaines objections traditionnellement adressées au néo-pragmatisme. D'autre part, j'ai reconnu que le modèle était encore

à un stade d'ébauche et que ce dernier mériterait plus de travail pour procéder à une véritable explication du langage.

Sachant tout ce qui a été écrit précédemment, est-il possible de conclure que l'énactivisme permet de naturaliser l'intentionnalité? Je crois que la stratégie de l'énactivisme a un bon potentiel pour compléter l'intégration des intentionnalités biologique, pragmatique et représentationnelle, et ainsi réaliser la naturalisation de l'intentionnalité telle que présagée dans l'introduction. En effet, en montrant qu'on peut dériver le *linguaging* du *sense-making* participatif, lui-même dérivable du *sense-making* sensorimoteur, et lui-même dérivable du *sense-making* biologique, il est possible d'affirmer que l'énactivisme accomplit une véritable intégration des différents types d'intentionnalité. J'ai argumenté dans le deuxième chapitre que le *sense-making* biologique était une forme d'intentionnalité biologique, et que le *sense-making* sensorimoteur était une forme d'intentionnalité pragmatique. J'ai argumenté dans le troisième chapitre que le modèle du corps linguistique permet à Di Paolo, Cuffari et De Jaeger d'intégrer (en partie) les phénomènes d'intentionnalité représentationnelle dans le cadre énonciviste, notamment grâce à leur explication des symboles et de la syntaxe. Cependant, s'il est possible d'espérer que l'énactivisme permette l'intégration des différents types d'intentionnalité, force est d'admettre que du travail supplémentaire devra être accompli pour pouvoir naturaliser tous les phénomènes d'intentionnalité représentationnelle, et ainsi compléter le projet de la naturalisation de l'intentionnalité.

## RÉFÉRENCES

- Anderson, M. L. (2014). *After phrenology: Neural reuse and the interactive brain*. MIT Press.
- Andr n, M. (2017). Children’s expressive handling of objects in a shared world. Dans C. Meyer, J. Streeck, & J. S. Jordan (dir.), *Intercorporeality: Emerging Socialities in Interaction* (p. 105–141). Oxford University Press.
- Austin, J. L. (1975). *How to do things with words*. Oxford University Press.
- Bakhtin, M. M. (1986). *Speech genres and other late essays*. University of Texas Press.
- Barandiaran, X. E., Di Paolo, E., & Rohde, M. (2009). Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive Behavior*, 17(5), 367-386.
- Barandiaran, X., & Moreno, A. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior*, 16(5), 325-344.
- Bateson, G. (1979). *Mind and nature: A necessary unity*. E.P. Dutton
- Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4), 209-243.
- Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. John Wiley & Sons.
- Bigelow, J. & Pargetter, R. (1987). Functions. *Journal of Philosophy*, 84, 181-96.
- Brandom, R. (1994). *Making it explicit: reasoning, representing and discursive commitment*. Harvard University Press.
- Brentano, F. (1874). *Psychology from an empirical standpoint*. Routledge.
- Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*, 8, 1-14.
- Burge, T. (2010). *Origins of objectivity*. Oxford University Press.
- Byers, A. M. (1999). Communication and material culture: Pleistocene tools as action cues. *Cambridge Archaeological Journal*, 9(1), 23–41.
- Cappucio, M. & Froese, T. (2014). *Enactive cognition at the edge of sense-making: making sense of non-sense*. Springer.
- Cash, M. (2008). The normativity problem: Evolution and naturalized semantics. *The Journal of Mind and Behavior*, 29(1), 99-137.
- Cash, M. (2009). Normativity is the mother of intention: Wittgenstein, normative practices and neurological representations. *New Ideas in Psychology*, 27(2), 133-147.
- Chalmers, D. J. (1995). The puzzle of conscious experience. *Scientific American*, 273(6), 80-86.
- Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.

- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78, 67–90.
- Clark, A. (2001). Reasons, robots and the extended mind. *Mind & Language*, 16(2), 121-145.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181-204.
- Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19.
- Clark, A., & Toribio, J. (1994). Doing without representing? *Synthese*, 101, 401-431.
- Cuffari, E. C., Di Paolo, E., & De Jaegher, H. (2015). From participatory sense-making to language: there and back again. *Phenomenology and the Cognitive Sciences*, 14, 1089-1125.
- Damasio, A. (2011). *Self comes to mind*. Pantheon.
- Davidson, D. (1987). Knowing One's Own Mind. *American Philosophical Association*, 60(3), 441-458.
- De Haan, S. (2020). *Enactive psychiatry*. Cambridge University Press.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the Cognitive Sciences*, 6(4), 485-507.
- De Jesus, P. (2016). From enactive phenomenology to biosemiotic enactivism. *Adaptive Behavior*, 24(2), 130-146.
- Deacon, T. W. (2011). The symbol concept. Dans M. Tallerman, & K. Gibson (dir.). *The Oxford handbook of language evolution* (p. 393-405). Oxford University Press.
- Degenaar, J., & Myin, E. (2014). Representation-hunger reconsidered. *Synthese*, 191(15), 3639-3648.
- Dennett, D. C. (1969). *Content and consciousness*. Routledge.
- Dennett, D. C. (1987). *The intentional stance*. MIT Press.
- Di Paolo, E. (2009). Extended life. *Topoi*, 28, 9-21.
- Di Paolo, E. A. (2005). Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4), 429-452.
- Di Paolo, E. A., Cuffari, E. C., & De Jaegher, H. (2018). *Linguistic bodies: The continuity between life and language*. MIT Press.
- Di Paolo, E., & Thompson, E. (2014). The enactive approach. Dans L. Shapiro (dir.), *The Routledge handbook of embodied cognition* (p. 86-96). Routledge.
- Di Paolo, E., Buhrmann, T., & Barandiaran, X. (2017). *Sensorimotor life: An enactive proposal*. Oxford University Press.
- Dretske, F. (1986). Misrepresentation. Dans R. Bogdan (dir.), *Belief: Form, content, and function* (p. 17-36). Oxford University Press.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. MIT Press.

- Dretske, F. (1988). *Explaining behavior: Reasons in a world of causes*. MIT Press.
- Dretske, F. (1995). *Naturalizing the mind*. MIT Press.
- Dretske, F. (2000). *Perception, knowledge and belief: selected essays*. Cambridge University Press.
- Dreyfus, H. L. (2002). Intelligence without representation—Merleau-Ponty's critique of mental representation: The relevance of phenomenology to scientific explanation. *Phenomenology and the Cognitive Sciences*, 1(4), 367-383.
- Dreyfus, H. L. (2007). The return of the myth of the mental. *Inquiry*, 50(4), 352-365.
- Egan, F. (2014). How to think about mental content. *Philosophical Studies*, 170(1), 115–135.
- Fisette, D., & Poirier, P. (2000). *Philosophie de l'esprit: État des lieux*. Vrin.
- Fitch, W. T. (2008). Nano-intentionality: A defense of intrinsic intentionality. *Biology and Philosophy*, 23, 157–177.
- Fodor, J. (1975). *The language of thought*. Harvard University Press.
- Fodor, J. (1981). *Representations*. MIT Press.
- Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. MIT Press.
- Fodor, J. (1989). Review Essay: Remnants of Meaning by Stephen Schiffer. *Philosophy and Phenomenological Research*, 50(2), 409-423.
- Fodor, J. (1990). *A theory of content and other essays*. MIT Press.
- Fodor, J. (2008). *LOT 2: The language of thought revisited*. Oxford University Press.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological sciences*, 360(1456), 815-836.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127-138.
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of physiology-Paris*, 100(1-3), 70-87.
- Froese, T. (2011). Breathing new life into cognitive science. *Avant Pismo Awangardy Filozoficzno-Naukowej*, 1, 113-129.
- Gallagher, S. (2017). *Enactivist interventions: Rethinking the mind*. Oxford University Press.
- Gallagher, S., & Cole, J. (1995). Body image and body schema in a deafferented subject. *The Journal of Mind and Behavior*, 16(4), 369-389.
- Gastelum, M. (2020). Bringing forth languages: Enacting humanity. Review of linguistic bodies: The continuity between life and language by Ezequiel A. Di Paolo, Elena Clare Cuffari and Hanne De Jaegher. *Constructivist Foundations*, 15(2), 194-198.
- Gibson, J. J. (1977). The theory of affordances. *Hilldale, USA*, 1(2), 67-82.
- Gibson, J. J. (1979). *The ecological approach to visual perception: classic edition*. Psychology Press.

- Giere, R. N. (1988). *Explaining science: A cognitive approach*. University of Chicago Press.
- Glüer, K., Wikforss, A., & Ganapini, M. (2009). *The Normativity of Meaning and Content*. The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/fall2023/entries/meaning-normativity/>.
- Godfrey-Smith, P. (1994). A modern history theory of functions. *Noûs*, 28(3), 344-362.
- Godfrey-Smith, P. (1996). Precis of Complexity and the function of mind in nature. *Adaptive Behavior*, 4(3-4), 453-465.
- Godfrey-Smith, P. (2006). Mental representation, naturalism, and teleosemantics. Dans D. Papineau & G. Macdonald (dir.), *Teleosemantics* (p. 42-68). Oxford: Oxford University Press.
- Godfrey-Smith, P. (2013). *Philosophy of biology*. Princeton University Press.
- Green, D. (2021). Rules of belief and the normativity of intentional content. *Acta Analytica*, 36(2), 159-169.
- Habermas, J. (1984). *The Theory of Communicative Action*. Beacon Press.
- Hattiangadi, A. (2006). Is meaning normative? *Mind & Language*, 21(2), 220-240.
- Haugeland, J. (1990). The intentionality all-stars. *Philosophical Perspectives*, 4, 383-427.
- Heft, H. (2001). *Ecological psychology in context: James Gibson, Roger Barker, and the legacy of William James's radical empiricism*. Psychology Press.
- Heidegger, M. (1927/1986). *Être et Temps* (trad. Vézin, F.). Gallimard.
- Hockett, C. F., & Hockett, C. D. (1960). The origin of speech. *Scientific American*, 203(3), 88-97.
- Hopper, P. (1987). Emergent grammar. *Annual Meeting of the Berkeley Linguistics Society*, 13, 139-157.
- Horgan, T., & Tienson, J. (2002). The intentionality of phenomenology and the phenomenology of intentionality. Dans D. Chalmers (dir.), *Philosophy of Mind: Classical and Contemporary Readings* (p. 520-533). Oxford University Press.
- Hutto, D. D., & Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. MIT Press.
- Hutto, D. D., & Myin, E. (2017). *Evolving enactivism*. MIT Press.
- Hutto, D. D., & Satne, G. (2015). The natural origins of content. *Philosophia*, 43(3), 521-536.
- Jacob, P. (2003). *Intentionality*. The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/spr2023/entries/intentionality>
- Jacob, P., & Jeannerod, M. (2003). *Ways of seeing. The scope and limits of visual cognition*. Oxford University Press.
- James, M. M. (2020). Bringing forth within: enhabiting at the intersection between enaction and ecological psychology. *Frontiers in Psychology*, 11(1348), 1-15.

- Jonas, H., & Jonas, E. (1966). *The phenomenon of life: Toward a philosophical biology*. Northwestern University Press.
- Kelly, S. D. (2002). Merleau-Ponty on the body: The logic of motor intentionality. *Ratio*, 15(4), 376–391.
- Kirchhoff, M. D. (2018). Autopoiesis, free energy, and the life–mind continuity thesis. *Synthese*, 195(6), 2519-2540.
- Kirchhoff, M. D., & Froese, T. (2017). Where there is life there is mind: In support of a strong life–mind continuity thesis. *Entropy*, 19(169), 1-18.
- Kiverstein, J. D., & Rietveld, E. (2018). Reconceiving representation-hungry cognition: an ecological-enactive proposal. *Adaptive Behavior*, 26(4), 147-163.
- Kiverstein, J., & Rietveld, E. (2015). The primacy of skilled intentionality: on Hutto & Satne’s the natural origins of content. *Philosophia*, 43, 701-721.
- Kochanska, G., Forman, D. R., Aksan, N., & Dunbar, S. B. (2005). Pathways to conscience: Early mother–child mutually responsive orientation and children's moral emotion, conduct, and cognition. *Journal of Child Psychology and Psychiatry*, 46(1), 19-34.
- Kriegel, U. (2013). *Phenomenal intentionality*. Oxford University Press.
- Kripke, S. A. (1982). *Wittgenstein on rules and private language: An elementary exposition*. Harvard University Press.
- Linell, P. (2009). *Rethinking language, mind, and world dialogically*. IAP.
- Lobo, L., Heras-Escribano, M., & Travieso, D. (2018). The history and philosophy of ecological psychology. *Frontiers in Psychology*, 9(2228), 1-15.
- Maiese, M. (2016). *Embodied selves and divided minds*. Oxford University Press.
- Maiese, M. (2022). *Autonomy, enactivism, and mental disorder: A philosophical account*. Taylor & Francis.
- Malafouris, L. (2013). *How things shape the mind*. MIT Press.
- Matthews, R. (2007). *The measure of mind*. Oxford University Press.
- Maturana, H. (2002). Autopoiesis, structural coupling and cognition: a history of these and other notions in the biology of cognition. *Cybernetics & Human Knowing*, 9(3-4), 5-34.
- Maturana, H. & Varela, F. (1980). Autopoiesis and Cognition. The Realization of the Living. *Boston Studies in the Philosophy of Science*, 42, 1-141.
- McGinn, C. (1989). Can we solve the Mind-Body problem? *Mind*, 98(391), 349-366.
- Mead, G. H. (1934). *Mind, self, and society*. University of Chicago Press
- Merleau-Ponty, M. (1942). *La structure du comportement*. Presses Universitaires de France.
- Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Presses Universitaires de France.

- Millikan, R. G. (1984). *Language, thought, and other biological categories: New foundations for realism*. MIT Press.
- Millikan, R. G. (1989). Biosemantics. *Journal of Philosophy*, 86, 281–297.
- Miyahara, K. (2011). Neo-pragmatic intentionality and enactive perception: A compromise between extended and enactive minds. *Phenomenology and the Cognitive Sciences*, 10(4), 499-519.
- Mossio, M., Saborido, C., & Moreno, A. (2009). An organizational account of biological functions. *The British Journal for the Philosophy of Science*, 60(4), 813-841.
- Neander, K. (1995). Misrepresenting & malfunctioning. *Philosophical Studies*, 79(2), 109-141.
- Neander, K. (1996). Swampman meets swampcow. *Mind & Language*, 11(1), 118-129.
- Neander, K. (2016). *A mark of the mental*. Oxford University Press.
- Nelson, M. (2000). *Propositional Attitude Reports*. The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/spr2023/entries/prop-attitude-reports/>
- Newell, A. (1980). Physical symbol systems. *Cognitive science*, 4(2), 135-183.
- Noë, A. (2004). *Action in perception*. MIT Press.
- O'regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5), 939-973.
- O'Brien, G., & Opie, J. (2004). Notes toward a structuralist theory of mental representation. Dans H. Clapin, P. Staines, & P. Slezak (dir.), *Representation in mind: Approaches to mental representation* (p. 1–20). Elsevier.
- Papineau, D. (2022). Swampman, teleosemantics and kind essences. *Synthese*, 200(509), 1-15.
- Pattee, H. H. (2012). Evolving self-reference: matter, symbols, and semantic closure. Dans H. H. Pattee & J. Rączaszek-Leonardi (dir.), *Laws, Language and Life: Howard Pattee's classic papers on the physics of symbols* (p. 211-226). Springer.
- Pattee, H. H., & Kull, K. (2009). A biosemiotic conversation: Between physics and semiotics. *Sign Systems Studies*, 37(1/2), 311-331.
- Pattee, H. H., Rączaszek-Leonardi, J. (2012). Language as a system of replicable constraints. Dans H. H. Pattee & J. Rączaszek-Leonardi (dir.), *Laws, Language and Life: Howard Pattee's classic papers on the physics of symbols* (p. 295-328). Springer.
- Peirce, C. S. (1974). *Collected papers of Charles Sanders Peirce*. Harvard University Press.
- Piaget, J. (1936). *La naissance de l'intelligence chez l'enfant*. Neuchâtel, Paris: Delachaux et Niestlé.
- Piaget, J. (1947). *The psychology of intelligence*. London: Routledge.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515-526
- Quine, W. V. (1969). *Ontological relativity and other essays*. Columbia University Press.
- Ramsey, W. M. (2007). *Representation reconsidered*. Cambridge University Press.

- Ramstead, M. J. D., Badcock, P. B., & Friston, K. J. (2018). Answering Schrödinger's question: A free-energy formulation. *Physics of life reviews*, 24, 1-16.
- Ramstead, M. J., Kirchhoff, M. D., Constant, A., & Friston, K. J. (2021). Multiscale integration: beyond internalism and externalism. *Synthese*, 198(1), 41-70.
- Ramstead, M. J., Veissière, S. P., & Kirmayer, L. J. (2016). Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. *Frontiers in psychology*, 7(1090), 1-15.
- Rietveld, E. (2012). Bodily intentionality and social affordances in context. Dans F. Paglieri (dir.), *Consciousness in interaction: The role of the natural and social context in shaping consciousness* (p. 207-226). John Benjamins.
- Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological psychology*, 26(4), 325-352.
- Ritchie, J. (2014). *Understanding naturalism*. Routledge.
- Rivière, Á., & Español, S. (2003). La suspensión como mecanismo de creación semiótica. *Estudios de Psicología*, 24, 261–275.
- Rolla, G., & Huffermann, J. (2022). Converging enactivisms: radical enactivism meets linguistic bodies. *Adaptive Behavior*, 30(4), 345-359.
- Root, M. D. (1975). Language, rules, and complex behavior. Dans K. Gunderson (dir.), *Language, Mind, and Knowledge* (p. 321–343). University of Minnesota Press.
- Rosenberg, A. (2013). How Jerry Fodor slid down the slippery slope to anti-Darwinism, and how we can avoid the same fate. *European Journal for Philosophy of Science*, 3(1), 1-17.
- Satne, G. (2016). Social approaches to intentionality. Dans J. Kiverstein, (dir.), *The Routledge handbook of philosophy of the social mind* (p. 544-560). Routledge.
- Schlicht, T. (2018). Does separating intentionality from mental representation imply radical enactivism? *Frontiers in Psychology*, 9(1497), 1-14.
- Schlicht, T., & Starzak, T. (2021). Prospects of enactivist approaches to intentionality and cognition. *Synthese*, 198(1), 89-113.
- Schlosser, M. (2015). *Agency*. The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/win2019/entries/agency/>
- Schrödinger, E. (1944). *What is life?: With mind and matter and autobiographical sketches*. Cambridge University Press.
- Searle, J. R., & Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- Sellars, W. (1953). Inference and meaning. *Mind*, 62(247), 313-338.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423.

- Shea, N. (2014). Exploitable isomorphism and structural representation. *Proceedings of the Aristotelian Society*, 114(2), 123-144.
- Short, T. L. (2007). *Peirce's theory of signs*. Cambridge University Press.
- Sinigaglia, C. (2008). Enactive understanding and motor intentionality. Dans F. Morganti, A. Carassa, & G. Riva (dir.), *Enacting intersubjectivity: A cognitive and social perspective on the study of interactions* (p. 17–32). IOS Press.
- Slaby, J., Stephan, A., & Walter, H. (2011). *Affektive Intentionalität*. Mentis.
- Sprevak, M. (2011). Review of representation reconsidered by William Ramsey. *British Journal for the Philosophy of Science*, 62, 669-675.
- Sprevak, M. (2013). Fictionalism about neural representations. *The Monist*, 96, 539–560.
- Stampe, D. W. (1977). Toward a causal theory of linguistic representation. *Midwest Studies in Philosophy*, 2(1), 42-63.
- Steiner, P. (2019). Content, Mental Representation and Intentionality: Challenging the Revolutionary Character of Radical Enactivism. *Croatian Journal of Philosophy*, 19(55), 153-174.
- Stich, S. (1990). *The fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*. MIT Press.
- Tarski, A. (1931). Sur les ensembles définissables de nombres réels. I. *Fundamenta Mathematicae*, 17, 210–239.
- Thompson, E. (2004). Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*, 3(4), 381-398.
- Thompson, E. (2010/2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Harvard University Press.
- Thompson, E. (2018). Review of Daniel D. Hutto and Erik Myin, *Evolving Enactivism: Basic Minds Meet Content*, *Notre Dame Philosophical Reviews*. <https://ndpr.nd.edu/reviews/evolving-enactivism-basic-minds-meet-content/>
- Thompson, S. A., & Couper-Kuhlen, E. (2005). The clause as a locus of grammar and interaction. *Discourse Studies*, 7(4-5), 481-505.
- Tison, R. (2022). The fanciest sort of intentionality: Active inference, mindshaping and linguistic content. *Philosophical Psychology*, 35, 1-41.
- Tomasello, M. (2009). *The cultural origins of human cognition*. Harvard University Press.
- Tomasello, M. (2014). The ultra-social animal. *European journal of social psychology*, 44(3), 187-194.
- Tracy, E. V. (2020). Unfollowed Rules and the Normativity of Content. *Analytic Philosophy*, 61(4), 323-344.
- Turvey, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, 4(3), 173-187.

- Van Dijk, L., & Rietveld, E. (2017). Foregrounding sociomaterial practice in our understanding of affordances: The skilled intentionality framework. *Frontiers in Psychology*, 7.
- van Gelder, T. (1995), The Distinction Between Mind and Cognition, dans Y. Hounq & J. Ho (dir.), *Mind and Cognition: Collected Papers from 1993 International Symposium on Mind & Cognition, Taipei, Institute of European and American Studies*. Academia Sinica.
- Varela, F. J. (1997). Patterns of life: Intertwining identity and cognition. *Brain and Cognition*, 34(1), 72-87.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind, revised edition: Cognitive science and human experience*. MIT Press.
- Veissière, S. P., Constant, A., Ramstead, M. J., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, 43, 1-75.
- Villalobos, M., & Ward, D. (2016). Lived Experience and Cognitive Science Reappraising Enactivism's Jonasian Turn. *Constructivist Foundations*, 11(2), 204-212.
- Ward, D., Silverman, D., & Villalobos, M. (2017). Introduction: The varieties of enactivism. *Topoi*, 36, 365-375.
- Weber, A., & Varela, F. J. (2002). Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1(2), 97-125.
- Wiese, W., & Friston, K. J. (2021). Examining the continuity between life and mind: Is there a continuity between autopoietic intentionality and representationality? *Philosophies*, 6(1), 18.
- Wittgenstein, L. (2010/1953). *Philosophical investigations*. John Wiley & Sons.
- Wright, D. (1982). Piaget's theory of practical morality. *British Journal of Psychology*, 73, 279–283.
- Wright, D. (1983). Piaget's theory of moral development. Dans S. Modgil, & C. Modgil (dir.), *Jean Piaget: Consensus and controversy* (p. 207–217). Praeger Publishers.
- Zahavi, D. (2018). *Phenomenology the basics*. Routledge.