

Le Moi comme un objet intentionnel. Une sémantique de 'Je' sans engagement ontologique.

Manuel Rebuschi

Article paru dans *Dialogue*, 51 (02), juin 2012, pp. 187-210

Résumé :

Les attitudes *de se* sont généralement considérées comme constituant une classe particulière d'attitudes *de re*. Une analyse alternative est proposée, qui s'appuie sur la notion d'attitude *de objecto* et qui évite un engagement ontologique envers le sujet. La proposition élabore l'idée de Hintikka d'une logique épistémique dite de seconde génération, qui introduit un marqueur syntaxique permettant d'exprimer des relations d'indépendance entre certaines constantes logiques. De cette sémantique résulte une conception du Moi, dénotation de 'Je', comme un objet simplement intentionnel.

Abstract:

De se attitudes were generally regarded as being a proper part of *de re* attitudes. In this paper an alternative analysis is proposed. It is based on the notion of *de objecto* attitude, that avoids ontological commitment to the self. The proposal develops the idea of Hintikka's second generation epistemic logic, which introduces a syntactic marker expressing independence relations between certain logical constants. According to this semantic account, the Self, as denotation of 'I', is conceived of as a mere intentional object.

1. Introduction

La plupart des traitements sémantiques des locutions en première personne reposent sur un préjugé ontologique inacceptable : elles présupposent que le pronom personnel « je » désigne un individu dans le monde. Cela n'est pas inacceptable en soi, mais inacceptable de la part de théories *sémantiques*, dont la prétention est d'analyser la signification linguistique, et dont la vocation n'est pas de dire *ce qu'il y a* dans le monde, qui est du ressort de l'ontologie et de la métaphysique.

L'objectif de cet article est de proposer une analyse sémantique des locutions en première personne qui soit dégagée de présupposés ontologiques. Le point n'est pas linguistique (rendre compte de la signification des usages de « je ») mais philosophique, et il peut être plus ou moins ambitieux. Le *programme minimum* consiste à clarifier la structure conceptuelle des pensées *égologiques* (j'emprunte le terme à Stéphane Chauvier, 2001), c'est-à-dire à mettre au jour certaines des contraintes qui pèsent sur manière dont nous pensons en première personne. Le *programme maximum* vise à éclairer la nature du Moi, autrement dit à tirer des leçons métaphysiques de l'analyse logique du langage. Il est cependant difficile de trancher par avance quelle sera la portée de l'analyse sémantique, qui plus est quand l'objet de l'analyse est le Moi. Il semble que l'évaluation du résultat (programme minimum ou maximum) dépende en grande partie de positions philosophiques adoptées par ailleurs.

Le plan de l'article est le suivant.

Je vais tout d'abord (Section 2) exposer en quoi consiste le préjugé ontologique fâcheux partagé par la plupart des analyses sémantiques de « je ». J'indiquerai quelques-uns des réquisits fondamentaux d'une analyse sémantique de « je », tels qu'ils ressortent des travaux contemporains en sémantique linguistique, mais aussi en philosophie du langage et de l'esprit. Ces réquisits vont encadrer le cœur de l'article, où je présenterai ma propre analyse de « je », fondée sur la notion d'attitude *de objecto* développée dans Rebuschi & Tulenheimo, 2011.

Je présenterai ensuite l'arrière-plan théorique et les concepts formels mobilisés par les attitudes *de objecto*. Nous allons croiser Hintikka, à deux reprises. Il y a tout d'abord la logique épistémique et doxastique, initiée par Hintikka il y a un demi-siècle (Hintikka, 1962), qui permet d'utiliser l'appareillage de la logique modale pour formaliser les notions de connaissance et de croyance (Section 3); il y a ensuite la logique IF (*independence-friendly*, ou *faite pour l'indépendance*) qui est une extension de la logique du premier ordre développée par Hintikka et Sandu depuis la fin des années 1980 (cf. Hintikka & Sandu, 1997), et dont une application à la logique épistémique et doxastique donne lieu à ce que Hintikka a qualifié, au début des années 2000, de *logique épistémique de seconde génération* (Hintikka, 2003) (Section 4). C'est dans ce contexte théorique que sont construites les attitudes *de objecto*.

Je présenterai (Section 5) l'application qui peut être faite des attitudes *de objecto* pour rendre compte des locutions en première personne. Je montrerai de quelles manières cette analyse sémantique, quoique neutre, peut être adaptée à différentes conceptions ontologiques du Moi. Dans la conclusion, je reviendrai sur la portée philosophique de ma proposition (programme minimum versus maximum), et sur ce que l'ascension sémantique permet d'inférer sur une question spécifique comme celle de l'analyse du Moi.

2. Le problème des analyses standard des locutions en première personne

Nous parlons pour exprimer nos croyances, nos connaissances, nos désirs, nos craintes. Autrement dit, nous exprimons linguistiquement nos *attitudes mentales*. Par exemple, si Mathilde déclare : « Lacan est un psychanalyste renommé », et si certaines conditions de sincérité sont remplies, elle exprime sa croyance, voire sa connaissance que Lacan est un psychanalyste renommé ; si Mathilde crie, à l'impératif : « Jacques, viens ici ! », elle exprime alors son désir que Jacques vienne à l'endroit où elle se trouve. Les croyances, connaissances, désirs, etc. sont différents types d'attitudes mentales. Le terme d'*attitude mentale* que j'emploie ici est neutre : il ne présuppose rien aux plans métaphysique — c'est-à-dire quant à la nature desdites attitudes mentales — ou ontologique — quant à leur existence authentique. Certains considèrent que les attitudes sont des propriétés authentiquement possédées par les sujets, et qu'il s'agit de décrire adéquatement; d'autres que ce sont des relations entre les sujets et des propositions, ou entre des sujets et d'autres types d'objets, qu'il s'agit également de décrire; d'autres enfin considèrent que les attitudes n'existent pas, mais que nous faisons comme si elles existaient en les attribuant aux autres, aux bébés, aux robots, et à nous-

mêmes¹, ou plus sobrement, que ce sont des concepts mentaux que nous utilisons pour interpréter les comportements.

Je vais me pencher sur le fonctionnement des *locutions intentionnelles*, c'est-à-dire des phrases comme « Mathilde croit que Lacan est un psychanalyste renommé » ou « Mathilde désire que Jacques vienne à l'endroit où elle se trouve » : des phrases qui ont la forme grammaticale :

S *VA* que *P*

S étant le sujet grammatical (comme « Mathilde »), *VA* un verbe d'attitude (« croit »), et *P* une expression décrivant le contenu de l'attitude (« Lacan est un psychanalyste renommé »), ou une structure plus complexe. L'énonciation d'une locution intentionnelle est à distinguer de l'expression directe d'une attitude (comme quand Mathilde dit « Lacan est un psychanalyste renommé ») : c'est une *attribution* d'attitude. Pour l'instant je ne veux pas trancher sur le caractère, descriptif ou autre, de ces attributions. Nous attribuons des attitudes mentales aux autres, à nous-mêmes (Mathilde pourrait s'auto-attribuer une croyance en disant : « Je crois que Lacan est un psychanalyste renommé »; ce faisant, elle exprimerait une croyance de second-ordre²), et parfois à nos animaux de compagnie (« Il veut sortir »), à nos ordinateurs (« Il veut prendre ma reine »), ou à des inexistants (« James Bond veut attraper le méchant avec la mâchoire métallique »). C'est à l'aide de ces attributions que nous comprenons ou tentons de comprendre les autres, nous-mêmes, les êtres plus ou moins intelligents qui nous environnent, les personnages des films que nous regardons, etc.³

Sans trancher sur leur nature ni même sur leur réalité, nous pouvons dégager quelques-unes des caractéristiques des attitudes que nous attribuons par nos locutions intentionnelles.

Tout d'abord on se restreint généralement aux cas d'états mentaux statiques, parfois durables et implicites, plutôt que de processus mentaux. Quand nous attribuons à Mathilde la croyance que Lacan est un psychanalyste renommé, nous lui attribuons une qualité statique, permanente et le plus souvent implicite, à la différence de situations où nous lui attribuerions une action mentale, comme dans : « Mathilde calcule la note au restaurant », ou « Mathilde étudie l'impact de la vitesse des neutrinos sur la théorie de la relativité ».

D'autre part, les attitudes ont (généralement) un *contenu*. La croyance attribuée à Mathilde a pour contenu *quelque chose* qui est exprimé par « Lacan est un psychanalyste renommé ». On considère souvent que ce *quelque chose* qui constitue le contenu est une *proposition* (d'où l'appellation contrôlée d'*attitude propositionnelle*, donnée aux attitudes dont le contenu est une proposition), ou un *état de choses*, mais il y a d'autres conceptions (suivant lesquelles le

¹ C'est *grosso modo* la conception défendue par Dennett dans *The Intentional Stance* (1987) et que je partage. Je la reprendrai plus loin, dans la section 5.

² Même si la lecture dominante consistera à lui attribuer l'usage d'un performatif pour exprimer sa croyance de premier ordre. Mon propos ici n'est cependant pas d'interpréter l'assertion de Mathilde, mais de donner un exemple d'expression possible d'une attitude de second ordre.

³ On parle ici d'attributions authentiques, pas de pseudo-attributions comme quand nous disons de notre voiture : « Ce matin, elle ne veut pas démarrer ».

contenu est une phrase en langue naturelle, ou une phrase d'un langage mental, etc.) et je vais esquiver le débat.

Les attributions d'attitudes ont ainsi (généralement) une *forme relationnelle* : les locutions intentionnelles ont (généralement) la forme d'un verbe qui relie un *sujet* (grammatical) et un *contenu* (grammaticalement : une subordonnée dans les exemples envisagés jusqu'ici). Il faut résister à la tentation de prendre cette forme grammaticale pour argent comptant et de considérer que si les locutions intentionnelles sont (généralement) relationnelles, c'est parce qu'elles décrivent d'authentiques relations entre un Sujet qui serait une entité réelle, et un contenu qui serait une entité également réelle.

La forme relationnelle des locutions intentionnelles, entre un sujet et un contenu, est souvent vérifiée, mais pas toujours. Il y a tout d'abord les cas d'attitudes sans contenu apparent, dont l'attribution consiste en un prédicat (psychologique) monadique (comme « Mathilde a mal »). Je ne vais pas m'attarder sur ces cas mais sur d'autres types d'attitudes, dont le contenu est manifestement scindé en au moins deux parties : un objet, et une qualité éventuellement complexe.

Le contraste entre attitudes au contenu simple et attitudes au contenu scindé a été étudié par la Scolastique, et il a nourri de nombreuses disputes dans les années 1960 et 70 sur la possibilité d'une logique épistémique et doxastique, et plus généralement sur la possibilité d'une logique modale quantifiée. Il s'agit du contraste entre attitudes *de dicto* et attitude *de re*. Pour reprendre l'un des exemples fameux de Quine (1956), on peut contraster les deux locutions suivantes :

- *de dicto* : Ralph croit qu'il y a un voisin qui est espion.
- *de re* : Ralph croit de l'un des voisins qu'il est espion.

Les deux attitudes n'ont pas la même structure. Dans le premier cas, le contenu de la croyance de Ralph est une phrase, un *dictum* complet, que l'on peut extraire et exprimer tel quel (« il y a un voisin qui est espion »), qui peut être vrai ou faux, etc. Dans le second cas, le contenu est scindé en un individu ou plus généralement une chose (*res*) dans le monde (un des voisins particuliers de Ralph), et une qualité (être un espion) attribuée par Ralph à cet individu; le *dictum* « il est espion » comporte un pronom anaphorique, « il », qui le rend incomplet : il faut ici que la *res* vienne saturer le contenu de l'attitude.

A la différence d'une attitude *de dicto* qui est générale, une attitude *de re* est focalisée sur une *res* particulière, ce qui en rend le contenu plus déterminé. Ainsi, on considérera le plus souvent que l'attitude *de re* de Ralph contient une détermination que n'a pas l'attitude *de dicto* correspondante, si bien que l'attitude *de re* implique l'attitude *de dicto*, tandis que la réciproque est fautive. Les attitudes *de re* ont été (et sont encore) contestées. Elles sont souvent interprétées comme incluant une relation directe entre le sujet et un objet dans le monde (la *res*). A ces attitudes dites relationnelles, impliquant des pensées singulières, on oppose les

attitudes dites notionnelles, impliquant des pensées générales, qui peuvent être purement internes au sujet⁴.

Avec ce contraste à l'esprit, venons-en aux énonciations en première personne, et à la manière dont on en rend compte de façon standard.

Si Mathilde dit « *Je suis une psychanalyste renommée* », elle exprime par là une attitude, probablement une croyance, que nous pourrions lui attribuer par la locution suivante : « Mathilde croit qu'*elle est une psychanalyste renommée* ». Or il semble bien que la pensée de Mathilde est alors focalisée sur un individu particulier (à savoir elle-même), et que l'on pourrait paraphraser l'attribution en explicitant cette détermination : « Mathilde croit *d'elle-même* qu'elle est une psychanalyste renommée ». Les locutions en première personne seraient ainsi, fondamentalement, l'expression d'attitudes *de re* plutôt que *de dicto*. Voilà précisément le préjugé dominant auquel je voudrais m'attaquer. Les locutions en première personne ont des traits qui les rapprochent des attitudes *de re* plutôt que des attitudes *de dicto*, cela ne fait aucun doute et j'y reviendrai. Mais les assimiler à des attitudes *de re* constitue un vrai problème.

Il a été noté par Wittgenstein dans les années 1930 (1964) puis par de nombreux philosophes dans les années 1960 (Perry, Castañeda, Kaplan pour n'en citer que quelques-uns) que les locutions en première personne ont des propriétés spécifiques qui conduisent à considérer qu'elles expriment des attitudes mentales d'une catégorie à part, les attitudes *de se*. Il est en effet impossible de paraphraser en en conservant le sens une locution en première personne, que ce soit à l'aide d'un nom propre ou d'une description définie. Quand Mathilde dit A : « *Je suis une psychanalyste renommée* », ce n'est ni équivalent à B : « *Mathilde est une psychanalyste renommée* », ni à C : « *La personne qui énonce cette phrase est une psychanalyste renommée* ». Pour chacune de ces deux paraphrases prétendues, on peut en effet envisager des situations plus ou moins farfelues où Mathilde accepterait de dire B sans dire A, ou de dire C sans dire A. Le fait que ces situations imaginaires soient farfelues n'ôte rien à l'argument, puisque si A et B (ou A et C) avaient la même signification, il faudrait que Mathilde accepte ou rejette conjointement A et B (ou A et C) dans toutes les situations possibles. Je ne vais pas entrer dans le détail des arguments ici ; il suffit pour mon propos de noter qu'il est largement admis que les locutions en première personne comme « je » ne peuvent pas être éliminées de l'expression de certaines attitudes, ce qui conduit à considérer ces attitudes comme constituant une classe à part⁵.

On peut formuler à nouveau le problème des analyses dominantes des locutions en première personne : ces conceptions présupposent que les attitudes *de se* constituent un sous-ensemble des attitudes *de re* (cf. p.ex. Maier, 2009). Autrement dit, elles présupposent un caractère

⁴ La discussion sur la légitimité des attitudes *de re* est étroitement liée au conflit entre internalisme et externalisme au sujet du mental, lui-même étant étroitement lié aux débats sur la nature du mental. Je n'aborde pas ces questions ici.

⁵ Castañeda a introduit un pronom spécial pour formaliser les attributions d'attitudes *de se* : pour distinguer l'attitude *de se* que Mathilde exprimerait à l'aide de A, du sens affaibli qu'elle exprimerait à l'aide de B ou C, Castañeda aurait écrit : « Mathilde croit qu'*elle** est une psychanalyste renommée ».

relationnel des attitudes *de se*, dont la *res* a ceci de particulier d'être le sujet de l'attitude. Ce point de vue comporte une idée qui n'est pas fautive : une attitude *de se* porte bien sur le sujet de l'attitude, au sens où le sujet figure, *es qualité*, dans le contenu ; et l'attitude est bien focalisée sur le sujet, ce dont ne rend pas compte une simple attitude *de dicto*⁶. Mais le contenu d'une attitude *de se* doit-il être scindé en une *res* (Moi) dans le monde, et une qualité attribuée à cette *res* (être une psychanalyste renommée) ? Considérer cela, c'est aller très largement au-delà des compétences d'une théorie sémantique. C'est en effet supposer résolu un débat ontologique et métaphysique, portant sur l'existence et la nature du Moi. En considérant les attitudes *de se* comme des cas particuliers d'attitudes *de re*, on répond positivement à la question de l'existence du Moi, et on en fait une substance (une *res*) susceptible de porter des attributs au même titre que les autres substances qui peuplent le monde comme les arbres, les chiens, voitures, les tables et les chaises. Très concrètement, les sémanticiens proposent des formalisations où chaque locuteur de « je » désigne une entité du domaine de l'interprétation du langage, sur laquelle on peut quantifier (donc cette entité existe) ; cela « résout » à très peu de frais des questions philosophiques pourtant débattues depuis des siècles⁷.

Je vais dans ce qui suit proposer ma propre analyse des locutions en première personne. Pour cela je vais passer par les attributions d'attitudes *de se* (« Mathilde croit qu'elle* est une psychanalyste renommée ») et je reviendrai seulement dans un second temps à l'expression directe de ces attitudes (Mathilde énonçant : « Je suis une psychanalyste renommée »).

3. Formaliser les attitudes : la logique modale épistémique, doxastique, et autre

Comme je l'ai dit en introduction on doit à Hintikka l'application de la logique modale à l'analyse formelle de la connaissance et de la croyance. L'ouvrage de 1962, *Knowledge and Belief*, a inauguré un champ de recherche extrêmement fructueux, que cela soit en philosophie pendant quelques décennies, en économie, ou en informatique.

L'idée de départ est de considérer que les verbes d'attitude (*VA*) ne doivent pas être formalisés comme des prédicats relationnels (entre un sujet et un contenu), mais comme des opérateurs préfixant des formules. Ainsi, « Mathilde croit que Lacan est un psychanalyste renommé » n'aura pas la forme CROIT(Mathilde, Lacan est un psychanalyste renommé), mais plutôt CROIT_{Mathilde} (Lacan est un psychanalyste renommé), soit sous une forme (semi-formalisée) plus compacte :

B_M PsychanalysteRenommé(Lacan)

où $B_M\phi$ se lit « Mathilde croit que (*believes that*) ϕ ». De même, l'énoncé « Mathilde sait que Lacan est un psychanalyste renommé » aura la forme :

⁶ Ma conception s'éloigne donc sensiblement de celle défendue par Chauvier, 2001.

⁷ Les linguistes ont des arguments pragmatiques en faveur de cette sémantique : il faut bien un locuteur pour énoncer une assertion en première personne. Je critique ces arguments dans Rebuschi, 2011.

K_M PsychanalysteRenommé(Lacan)

$K_M\phi$ se lisant : « Mathilde sait que (*knows that*) ϕ ». Les locutions intentionnelles sont donc considérées comme étant du même type que les locutions modales, par exemple « Lacan aurait pu ne pas être un psychanalyste renommé », ou « Lacan devait nécessairement être un psychanalyste renommé », que l'on formalise respectivement par :

$\diamond \neg$ PsychanalysteRenommé(Lacan)

\square PsychanalysteRenommé(Lacan)

où $\diamond\phi$ se lit « il est possible que ϕ », et $\square\phi$ « il est nécessaire que ϕ ».

Au-delà des motivations proprement logiques, notamment l'intensionnalité des locutions intentionnelles, l'intérêt de cette formalisation est qu'elle permet de transposer l'appareillage technique de la logique modale à l'analyse des locutions intentionnelles. Cet appareillage, développé simultanément par différents auteurs (dont Hintikka), c'est la sémantique des mondes possibles⁸.

L'idée est d'évaluer les phrases modales par rapport à un ensemble de mondes possibles. En simplifiant, $\diamond\phi$ sera vraie si et seulement s'il y a au moins un monde possible où ϕ est vraie ; et $\square\phi$ est vraie si et seulement si ϕ est vraie dans tous les mondes possibles — autrement dit s'il n'y a pas de monde possible où ϕ est fausse. Quand on entre un peu plus dans le détail, on ne considère pas les mondes possibles dans leur ensemble mais uniquement les mondes possibles *accessibles* suivant une modalité donnée. On dira que $\diamond\phi$ est vraie dans un monde donné w , si et seulement s'il y a au moins un monde *accessible depuis* w où ϕ est vraie ; et de même, $\square\phi$ est vraie dans un monde donné w , si et seulement dans tous les mondes *accessibles depuis* w , ϕ est vraie. Pour les modalités aléthiques, c'est-à-dire le possible et le nécessaire, il est fréquent de considérer que la relation d'accessibilité est totale, c'est-à-dire que depuis n'importe quel monde on peut accéder à n'importe quel monde (y compris lui-même). Dans ce cas, chaque monde possible devient un « monde possible accessible », et les définitions sont simplifiées.

Si l'on en vient aux modalités épistémiques et doxastiques, alors les relations d'accessibilité prennent tout leur sens. On dira que Mathilde sait que ϕ dans un monde w , si et seulement si ϕ est vraie dans tous les mondes épistémiquement accessibles depuis w ; l'opérateur K_M (« Mathilde sait que ») se comporte donc comme un opérateur de nécessité (\square). Que sont les mondes épistémiquement accessibles depuis w ? Ce sont les mondes indiscernables de w pour autant que Mathilde le sache ; ou dit autrement, les mondes que le savoir (ou la connaissance) de Mathilde ne lui permet pas de discriminer du monde w , ou encore les mondes compatibles avec ce que Mathilde sait (ou connaît).

⁸ L'appellation de *monde possible* ne doit pas obligatoirement être prise pour argent comptant à la manière de Lewis, 1986 : si l'on souhaite rejeter l'idée suivant laquelle il y a des mondes possibles au même titre que le monde dans lequel nous vivons, on peut considérer que *monde possible* n'est qu'une manière de parler des états possibles du monde.

Considérons le cas où Mathilde sait que p (Lacan est un psychanalyste renommé), mais elle ne sait pas si q (Il pleut le 21 novembre 2011 à Londres). On a quatre possibilités, que l'on peut représenter par quatre mondes possibles (voir Fig.1). Supposons qu'il ait plu à Londres le 21 novembre 2011, et donc que dans le monde actuel w les deux phrases soient vraies (on peut noter : $w = \{p, q\}$). Aucune des deux phrases ne paraît nécessairement vraie, ce qui correspond au fait qu'on peut trouver au moins un monde où p est fausse, et un monde où q est fausse. Tous ces mondes sont mutuellement accessibles suivant la modalité aléthique : la relation est totale (elle n'est pas représentée sur la figure). En revanche, d'après ce que sait Mathilde, il est vrai que p : dans tous les mondes épistémiquement accessibles, p est vraie ; donc les mondes où p est fausse sont peut-être aléthiquement possibles, mais ils ne sont pas accessibles suivant la relation correspondant à K_M (relation représentée par des flèches pleines sur la figure). Comme par ailleurs Mathilde ne sait pas s'il pleut à Londres, q et $\neg q$ sont toutes deux épistémiquement possibles : les deux mondes $\{p, q\}$ et $\{p, \neg q\}$ sont indiscernables pour Mathilde, étant donné ce qu'elle sait (ou « pour autant qu'elle sache »).

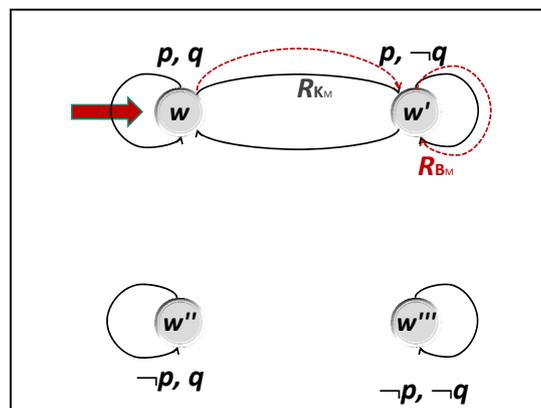


Fig.1. Possibilités aléthiques, épistémiques et doxastiques

La croyance (au sens de ce que l'on pense vrai) fonctionne de manière analogue à la connaissance : si Mathilde *croit* (ou *pense*) qu'il ne pleut pas à Londres, p.ex. parce qu'il lui a semblé l'entendre dire à la radio ce matin, cela sera modélisé par le fait que dans tous les mondes doxastiquement accessibles, $\neg q$ (« il ne pleut pas à Londres ») est vraie. De même que les mondes épistémiquement accessibles pour Mathilde sont les mondes compatibles avec ce qu'elle sait, les mondes doxastiquement accessibles (reliés par la flèche en pointillés sur la figure) sont les mondes compatibles avec ce qu'elle *croit* (ou *pense*). On a donc deux relations d'accessibilité entre mondes possibles, qui correspondent aux deux opérateurs K_M et B_M , et on pourrait dire à deux types de possibilités subjectives pour Mathilde. Suivant cette recette, on peut en principe dupliquer ces opérateurs pour introduire plusieurs sujets (« Antoine sait que » K_A , « Antoine croit que » B_A), et les concaténer pour attribuer des attitudes complexes (comme par exemple : « Antoine croit que Mathilde croit savoir qu'il pleut à Londres », $B_A K_M B_M q$). Mais on peut également créer de nouveaux opérateurs (et corrélativement de nouvelles relations entre mondes possibles) pour d'autres verbes d'attitudes, comme « Mathilde voit que », « Mathilde désire que », « Mathilde craint que », etc.

Il est utile de se pencher sur les propriétés des relations d'accessibilité associées aux opérateurs, et donc associées aux attitudes que ces opérateurs visent à formaliser. Ces propriétés correspondent en effet, au moins pour certaines d'entre elles, à des caractéristiques des opérateurs.

On peut l'illustrer à l'aide de deux exemples. Le savoir est traditionnellement caractérisé comme étant une croyance vraie et justifiée. Si l'on admet cette conception, le savoir implique la croyance. Par exemple, si Mathilde sait que Lacan est un psychanalyste, alors elle le croit. Autrement dit, si Mathilde sait que φ , c'est-à-dire si φ est vraie dans tous les mondes *épistémiquement* accessibles, alors elle croit que φ , c'est-à-dire que φ n'est fausse dans aucun monde *doxastiquement* accessible. Pour garantir cela quelle que soit φ , il faut supposer que tous les mondes doxastiquement accessibles sont épistémiquement accessibles. On a donc l'inclusion : $R_B \subset R_K$. Ce qui lie les deux relations d'accessibilité est en correspondance directe avec ce qui relie les deux concepts formalisés.

Un autre exemple, qui va nous intéresser par la suite, est celui de la *factivité* de la connaissance. Si l'on sait que φ , alors φ est vraie. Ce n'est pas seulement une conséquence de la caractérisation traditionnelle : c'est une règle à laquelle il semble impossible de déroger. Si Mathilde sait que Paul est psychanalyste, alors Paul est psychanalyste. Si elle le croit, si elle croit le savoir, si nous croyons qu'elle le sait, évidemment cela n'entraîne pas la vérité du contenu ; mais si elle le sait, alors c'est vrai. On ne peut pas concevoir que quelqu'un sache quelque chose de faux : la vérité est inscrite dans notre concept de connaissance. La connaissance est ainsi une attitude factive, de même que la vision, alors que la croyance ou le désir n'en sont pas.

La factivité d'un opérateur modal O est caractérisée par un axiome (connu sous le nom d'axiome T) : $O\varphi \rightarrow \varphi$. On considère généralement que l'opérateur de nécessité aléthique est factif ($\Box\varphi \rightarrow \varphi$), et que l'opérateur de connaissance l'est également ($K\varphi \rightarrow \varphi$), mais pas l'opérateur de croyance. A la factivité d'un opérateur correspond la *réflexivité* de la relation d'accessibilité : chaque monde est accessible à partir de lui-même. Pour le commenter de manière informelle : avec la réflexivité on est assuré que le monde actuel figure parmi les mondes accessibles depuis lui-même. Ce qui est vrai dans tous les mondes accessibles, est alors vrai dans le monde actuel. Si on considère la connaissance, ce qui est connu, c'est-à-dire vrai dans tous les mondes épistémiquement accessibles, est donc vrai dans le monde actuel. Réciproquement, si la relation d'accessibilité n'est pas réflexive, il y a alors au moins un monde w qui n'est pas accessible à lui-même : rien n'empêche qu'une phrase fausse dans w soit vraie dans tous les mondes accessibles depuis w ; on perd donc la factivité⁹.

⁹ Cette correspondance entre validité modale (ici la validité de l'axiome T) et propriété de la relation d'accessibilité (ici la réflexivité) n'est pas informelle, même si je la présente ici sur ce mode : elle se prouve rigoureusement.

4. L'extension IF, et les attitudes *de objecto*

Revenons au contraste entre attitudes *de dicto* et attitudes *de re*. En logique modale, le contraste est réputé se réduire à une question de portée. Reprenons l'exemple des croyances de Ralph à propos du voisin espion, en le simplifiant. En logique du premier ordre, en utilisant un langage semi-formalisé, on peut représenter la phrase : « il y a (au moins) un espion » : $(\exists x) (x \text{ est espion})$. On contraste alors de la manière suivante les deux types de croyances :

- *de dicto* : Ralph croit qu'il y a (au moins) un espion [*dictum* complet] :

$B_R (\exists x) (x \text{ est espion})$

- *de re* : Ralph croit d'une personne (au moins) qu'elle est un espion [*dictum*

incomplet + *res*] : $(\exists x) B_R (x \text{ est espion})$

La différence réside donc dans la portée syntaxique de l'opérateur doxastique (B_R) relativement au quantificateur existentiel ($\exists x$) : la portée est large et englobe tout le *dictum* dans le premier cas, étroite et se contente du prédicat dans le second. On peut évidemment contraster exactement de la même façon la connaissance *de dicto* de la connaissance *de re*¹⁰.

L'interprétation des deux formules à l'aide de la sémantique des mondes possibles est immédiate, même si elle soulève des difficultés. Dans le cas de la croyance *de dicto* (voir Fig.2), la formule est vraie dans le monde actuel si et seulement si, en allant dans n'importe quel monde doxastiquement accessible (B_R), on peut trouver un individu ($\exists x$) qui est un espion. Il n'y a ici aucune obligation (ni interdiction) de choisir le même individu dans chacun des mondes, ce qui est illustré sur la figure par la position de personnages distincts selon les mondes : il n'y a pas un individu déterminé qui serait l'objet de la croyance de Ralph.

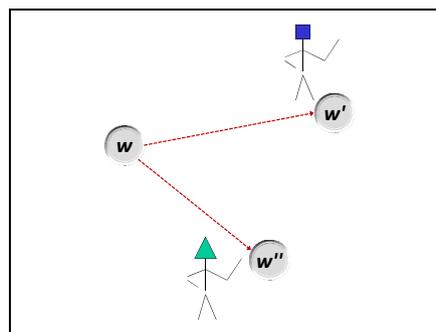


Fig.2. Croyance *de dicto*

A l'inverse, pour la croyance *de re* (Fig.3) il faut tout d'abord choisir un individu ($\exists x$) puis, en allant dans n'importe quel monde doxastiquement accessible (B_R), vérifier que *cet* individu est un espion. L'attitude de Ralph est alors bien focalisée sur un individu particulier. Du point de vue de la sémantique des mondes possibles, cela suppose qu'un même individu peut exister

¹⁰ On admet ici le présupposé méta-ontologique de Quine suivant lequel *être, c'est être la valeur d'une variable (quantifiée)*. Il est ainsi assumé que la *res* est un objet du monde (et pas une entité d'un éventuel domaine fictionnel ou autre).

dans différents mondes possibles¹¹. Pour Hintikka, la condition se résume à la supposition que Ralph puisse identifier l'individu concerné.

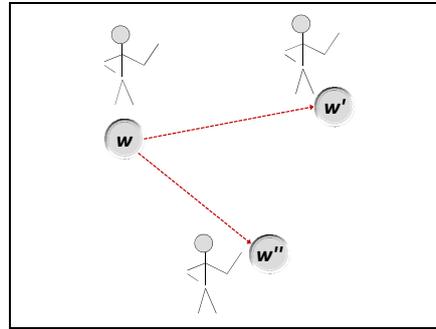


Fig.3. *Croyance de re*

Un problème surgit immédiatement : à la différence de nos connaissances, nos croyances peuvent être fausses parce que nous attribuons à des individus des qualités qu'ils n'ont pas (p.ex., nous croyons d'un voisin qu'il est un espion, alors qu'il ne l'est pas). Mais nos croyances peuvent être erronées d'une autre manière : dans le cas où nous attribuons des qualités à des individus qui n'existent pas. Ralph peut ainsi croire que James Bond est un espion, alors que James Bond n'existe pas. Or le contraste entre *de dicto* et *de re* ne permet pas de rendre compte des attitudes focalisées sur des inexistantes. Si Ralph croit que James Bond est un espion (en supposant qu'il croie que James Bond existe), cela ne peut pas être traité comme une croyance *de re* car il faudrait alors choisir un individu *dans le monde actuel* dont Ralph croie qu'il s'agit de James Bond et qu'il est un espion. On devrait donc se rabattre sur une croyance *de dicto*, alors que les croyances de ce type sont générales et ne peuvent viser un individu particulier.

Il a cependant été observé qu'une extension de la logique doxastique permettait d'étendre la typologie des attitudes et d'introduire le chaînon manquant : les attitudes que nous avons qualifiées de *de objecto* (Rebuschi & Tulenheimo, 2011). Cette extension est en fait la transposition à la logique modale d'une extension de la logique du premier ordre développée par Hintikka et Sandu, et dont Hintikka (2003) a proposé l'application à la logique épistémique pour des raisons que j'indiquerai brièvement.

L'extension en question est connue sous le nom de logique IF, pour *independence-friendly*, et souvent traduite par *faite pour l'indépendance*. Je me contente d'indiquer l'idée générale de la logique IF à partir d'un exemple.

Dans la formule $(\forall y) (\exists x) (x > y)$, le choix d'une valeur pour x dépend du choix de la valeur pour y , et la formule est vraie relativement à l'ensemble des nombres naturels. De façon générale en logique du premier ordre standard, les quantificateurs existentiels dépendent des quantificateurs universels dans la portée syntaxique desquels ils se trouvent. On passe à

¹¹ Il y a là matière à discussion, mais dans une approche « instrumentaliste » et antiréaliste de cette sémantique on se contentera de considérer qu'il s'agit bien du même individu envisagé dans des situations contrefactuelles.

l'extension IF en introduisant un marqueur syntaxique d'indépendance. Ainsi la formule : $(\forall y) (\exists x/\forall y) (x > y)$ est semblable à la précédente, à la différence près que le choix d'une valeur pour x doit être effectué de manière *uniforme* relativement au choix d'une valeur pour y . Autrement dit, le choix d'une valeur pour x doit être effectué sans savoir quelle valeur a été préalablement choisie pour y . Cette indépendance informationnelle rend la seconde formule fausse relativement au même modèle.

En fait, cette seconde formule est équivalente à une formule en logique du premier ordre standard (non IF) : $(\exists x) (\forall y) (x > y)$, puisqu'on peut obtenir l'indépendance informationnelle du quantificateur existentiel en le plaçant hors de portée du quantificateur universel. L'intérêt de la logique IF est cependant de permettre l'expression de formules qui ne sont pas équivalentes à des formules du premier ordre standard. Un exemple souvent donné est celui des formules à quantificateurs de Henkin, dits *branchants*, qui peuvent être représentés en logique IF mais pas en logique du premier ordre habituel (avec une valeur appropriée pour φ) :

$$\begin{aligned} & (\forall x) (\exists y) && &> \varphi(x, y, z, t) \\ & (\forall z) (\exists t) && & \\ & (\forall x) (\exists y) (\forall z) (\exists t/\forall x) \varphi(x, y, z, t) \end{aligned}$$

Cette extension peut être appliquée à la logique modale, et particulièrement à la logique épistémique. Hintikka défend ainsi l'idée que le contraste entre connaissance *de dicto* et connaissance *de re* est au fond une affaire d'indépendance, ce qui offre une explication satisfaisante de l'ambiguïté de certaines expressions en langue naturelle (comme « Ralph sait qu'il y a un espion », ambiguë entre *de dicto* et *de re*) du fait de l'absence de marqueur d'indépendance informationnelle — le slash n'étant explicite que dans le langage formel de la logique.

$$\begin{aligned} - \textit{de dicto} : & & K_R (\exists x) (x \text{ est espion}) \\ - \textit{de re version IF} : & & K_R (\exists x/K_R) (x \text{ est espion}) \\ - \textit{de re} : & & (\exists x) K_R (x \text{ est espion}) \end{aligned}$$

L'opérateur épistémique étant au plan sémantique un quantificateur universel (sur des mondes possibles), un quantificateur existentiel dans sa portée syntaxique tombe sous sa dépendance sémantique : dans le cas *de dicto*, le choix de la valeur pour x dépend du choix du monde pour K_R . L'extension IF permet de rendre un quantificateur existentiel informationnellement indépendant d'un opérateur épistémique. Ainsi dans la seconde formule ci-dessus, le choix de la valeur pour x doit être uniforme relativement au choix du monde possible pour K_R . Cela signifie qu'une fois dans un monde possible, on doit choisir une valeur pour x sans savoir dans quel monde on se trouve. Comme Hintikka l'a exposé, cela revient à inverser l'ordre des opérateurs, donc à effectuer le choix d'une valeur pour x avant d'aller dans d'autres mondes possibles pour vérifier si l'individu choisi est un espion.

Mais... l'équivalence entre la formule IF et la formule *de re* ne fonctionne que parce que l'opérateur est factif ! Ce qui fait que le choix peut être effectué indifféremment avant ou

après le voyage entre mondes possibles, c'est que le monde actuel (au départ) fait partie des destinations possibles, c'est-à-dire des mondes possibles accessibles. Le cas est totalement différent dès que l'on a affaire à une attitude non factive comme la croyance, puisqu'alors on peut faire le choix uniforme d'un individu qui existe dans les mondes doxastiquement accessibles (les mondes compatibles avec ce que Ralph croit), bien qu'il n'existe pas dans le monde actuel (voir Fig.4) : c'est le cas de la croyance *de objecto*.

- *de dicto* : $B_R (\exists x) (x \text{ est espion})$
- *de objecto* : $B_R (\exists x/B_R) (x \text{ est espion})$
- *de re* : $(\exists x) B_R (x \text{ est espion})$

Ainsi Ralph peut-il avoir des croyances focalisées sur James Bond, des désirs focalisés sur des particuliers inexistant (comme celui de peindre un tableau précis, pas un autre ni n'importe lequel, *avant* de l'avoir peint), des craintes focalisées sur des inexistant (comme la crainte de rencontrer la sorcière qui se cache au fond de son jardin, qu'il distingue de celle qui habite au fond de sa cave) ou encore des hallucinations. Dans chacun de ces cas, on a affaire à des attitudes *de objecto*, focalisées sur ce qu'on peut appeler un *objet intentionnel* comme le sont les attitudes *de re*, mais sans exigence quant à l'existence de cet objet.

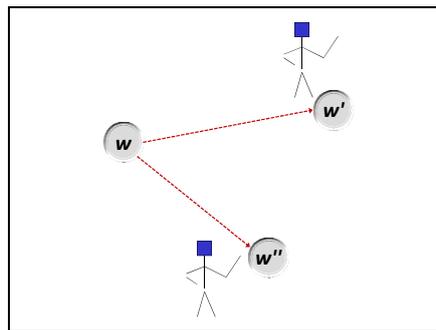


Fig.4. *Croyance de objecto*

On a bien affaire à trois types d'attitudes et non pas à deux. Pour une attitude A donnée, on aura de façon générale :

- *de dicto* : $A (\exists x) \varphi(x)$
- *de objecto* : $A (\exists x/A) \varphi(x)$
- *de re* : $(\exists x) A \varphi(x)$

Il faut cependant noter que si l'attitude A est factive (i.e. si $A\varphi \rightarrow \varphi$ est valide), les cas *de objecto* et *de re* se confondent, puisque le monde actuel étant accessible à lui-même, l'objet de l'attitude est assuré de s'y trouver. Autrement dit, l'objet intentionnel (la valeur de x) est un individu actuellement existant.

Si l'attitude n'est pas factive, les cas *de re* et *de objecto* peuvent être distingués, mais ils ne peuvent pas être distingués par le sujet de l'attitude. Du point de vue interne au sujet, les objets intentionnels d'attitudes comme la croyance ou le désir sont considérés de manière

similaire qu'ils existent ou non. Un enfant peut désirer de son oncle (dont il a entendu parler mais qu'il n'a jamais vu) qu'il vienne chez lui, et il peut également désirer du Père Noël (dont il a également entendu parler sans l'avoir jamais vu) qu'il vienne chez lui, sans que ces deux désirs soient foncièrement différents. Pour rendre compte du point de vue du sujet, on est conduit à privilégier l'attitude *de objecto* comme le type le plus unifiant. L'enfant désirerait *de objecto* du Père Noël qu'il vienne lui rendre visite, et il désirerait également *de objecto* de son oncle qu'il vienne lui rendre visite. Le fait que son oncle, l'individu visé par son attitude, existe dans le monde actuel, assurerait à nouveau la réduction de cette attitude particulière (le désir que son oncle vienne chez lui) du cas *de objecto* au cas *de re*.

Le clivage fondamental entre types d'attitudes peut ainsi être déplacé : on l'envisage comme un clivage entre attitudes *de dicto* (générales) et attitudes *de objecto* (focalisées¹²). Si l'objet de l'attitude existe actuellement, en particulier (mais pas exclusivement) dans le cas où l'attitude est factive, alors les attitudes *de objecto* ne sont rien d'autre que des attitudes *de re*. Les attitudes *de re* sont un cas particulier d'attitudes *de objecto*.

5. La connaissance de soi est-elle possible ?

Reprenons. Le défaut des analyses sémantiques habituelles du pronom de la première personne est de considérer que les attitudes *de se* sont un cas particulier d'attitude *de re*. Je propose ici de considérer que les attitudes *de se* sont un cas particulier d'attitude *de objecto*. Je vais indiquer suivant quelles directions cette analyse peut être développée.

« Je » est un indexical : sa valeur, l'individu qu'il désigne, dépend du contexte d'énonciation de chacune de ses occurrences. Chaque occurrence de « je » désigne le locuteur qui la produit. Si je dis « je », alors cette occurrence de « je » désigne MR; si c'est Jaakko Hintikka qui dit « je », alors cette occurrence de « je » le désigne lui et pas moi. C'est une différence importante entre indexicaux et noms propres, puisqu'on peut considérer avec quelque approximation¹³ que chaque occurrence d'un nom propre désigne le même individu indépendamment du contexte. Mais noms propres et indexicaux ont un point commun : ce sont des *désignateurs rigides*, c'est-à-dire qu'ils désignent le même individu dans tous les mondes possibles. Kripke (1972) a défendu cela au sujet des noms propres. Si je dis « Jaakko Hintikka aurait pu être psychanalyste », alors j'envisage un monde possible, autre que le monde actuel, où Jaakko Hintikka (l'individu actuellement désigné par ce nom) est psychanalyste¹⁴. Une fois la valeur de l'indexical « je » déterminée par le contexte, on a la

¹² J'hésite à parler d'attitudes singulières, puisque les attitudes dont le contenu est une pensée singulière sont généralement considérées comme portant sur un objet actuellement existant (c'est-à-dire *de re*). On devrait parler d'attitudes « quasi-singulières » à propos des attitudes *de objecto*.

¹³ L'approximation est liée aux homonymes, dont la distinction exige également de recourir au contexte. Pour une discussion détaillée de cette question, voir Fiengo & May, 2006.

¹⁴ Il faut différencier cela du cas où je désignerais Hintikka comme « le plus renommé des philosophes finlandais vivants », c'est-à-dire par l'usage d'une description définie. Dire que « Le plus renommé des philosophes finlandais vivants aurait pu être psychanalyste » est ambigu entre deux sens, l'un où la description continue de désigner Hintikka, l'autre où cette description désigne quelqu'un d'autre. Les descriptions définies sont des désignateurs que l'on peut qualifier de *flexibles*.

même rigidité : si Hintikka dit « J'aurais pu être psychanalyste », il faut envisager un monde possible autre que le monde actuel où Jaakko Hintikka lui-même (l'individu actuellement désigné par l'occurrence de « je », c'est-à-dire le locuteur actuel de la phrase) est psychanalyste.

L'un des réquisits d'une analyse sémantique de « je » est donc de rendre compte de son caractère de désignateur rigide. Le traiter comme un terme référentiel, au même titre qu'un nom propre, qui désigne dans tous les mondes possibles, est alors une solution facile qui nous conduit sur la pente des analyses dominantes. Nous n'avons pas besoin d'autant. Il suffit d'exiger que « je » désigne le même individu dans tous les mondes possibles *où il désigne quelque chose*. Si le mot « je » désigne le Moi, alors certes il le désigne de manière rigide, mais cela n'implique pas que le Moi existe dans le monde actuel. Ce qui est important, c'est que dans tous les mondes où le Moi existe, « je » le désigne.

Les déclarations en première personne peuvent ainsi être analysées comme impliquant un opérateur d'attitude implicite. Dans des conditions de sincérité ordinaires, l'affirmation d'une phrase peut être comprise comme l'expression d'une croyance du locuteur. Si j'affirme : « Jaakko Hintikka est un philosophe », j'exprime ainsi ma croyance que Jaakko Hintikka est un philosophe. Dire alors « Je crois que Jaakko Hintikka est un philosophe », à conditions de sincérité égales, n'ajoute pas grand-chose quant au contenu exprimé¹⁵. Dans un premier temps, je propose donc de reformuler les déclarations en première personne en les préfixant de « Je crois que ». Par exemple, pour analyser « Je suis un espion » on formule en « Je crois que je suis un espion ». La formalisation *de objecto* sera alors la suivante :

$$B (\exists x/B) (x = \mathbf{je} \ \& \ x \text{ est un espion})$$

où **je** est une constante individuelle symbolisant « je », et dénotant le Moi dans chaque monde doxastiquement accessible. L'intérêt de cette analyse est qu'elle offre une position ontologiquement neutre : on peut supposer que le Moi existe dans le monde actuel, comme on peut supposer l'inverse. Cette analyse n'entraîne ni le réalisme ni l'antiréalisme du Moi.

La limite qui surgit immédiatement, c'est que l'antiréalisme ne semble pas pouvoir résister au passage de la croyance à la connaissance de soi. Comme je l'ai dit, le cas *de objecto* se réduit à un cas *de re* dès qu'on a affaire à une attitude factive. La pensée à propos de Moi serait compatible avec l'antiréalisme uniquement dans les cas non factifs de la croyance à propos de Moi, du désir à propos de Moi, etc. mais envisager la connaissance de Moi exigerait de retourner à une conception réaliste. *Tout ça pour ça ?*

Il semble pourtant que l'on puisse envisager une connaissance véritable au sujet d'objets intentionnels. On peut ainsi dire que Mathilde *sait* que le monstre qu'elle craint est bleu. Si l'on peut être sceptique sur un exemple comme celui-ci, alors on peut envisager cet autre

¹⁵ Selon Chauvier « une croyance de second ordre est moins une croyance à propos d'une croyance de premier ordre qu'un état de croyance distinct des états de croyances du premier ordre » (Chauvier, 2001, p.76). Quand nous avons des *raisons* de croire que *p*, nous jugeons que nous le croyons (alors que ce n'est pas le cas quand nous sommes simplement causés à croire que *p*). Si je dis : « je crois que *p* », j'exprime simplement la (re)formation rationnelle de ma croyance (de premier ordre) que *p* (que j'aurais pu exprimer en déclarant : « *p* »).

exemple : Mathilde *sait* que la douleur qu'elle ressent dans le bras n'est pas la même que celle qu'elle ressentait hier. Le monstre craint par Mathilde, ou la douleur qu'elle ressent, sont des objets intentionnels qui n'existent pas dans le monde. Comment rendre compte de connaissance à leur propos ? Il faut enchâsser les opérateurs :

$K_M B_M (\exists x/B_M) (x \text{ est un monstre } \& x \text{ est bleu})$

$K_M B_M (\exists x/B_M) (x \text{ est une douleur } \& x \text{ est dans le bras})$

La croyance n'est cependant pas systématiquement appropriée quand on doit envisager les attitudes *de se*. Plutôt que de considérer que chaque déclaration en première personne doit être redoublée d'un « Je crois que... », je propose d'introduire un nouvel opérateur noté **Int**. Cet opérateur correspond au mode de discours intentionnel, sans forcément traduire un type particulier d'attitude mentale du sujet¹⁶. On peut lire **Int** ϕ de la manière suivante : « D'après l'interprétation [du sujet], ϕ », ou plus simplement : « D'après [le sujet], ϕ ». La déclaration « Je suis en colère » doit ainsi être reformulée en « D'après moi, je suis en colère », puis formalisée par :

Int $(\exists x/\text{Int}) (x = \text{je } \& x \text{ est en colère})$

L'évaluation sémantique des formules préfixées par cet opérateur se fait de la même façon que pour les opérateurs introduits jusqu'ici : on ajoute une relation d'accessibilité R_{Int} , qui relie entre eux les mondes intentionnellement possibles (i.e. compatibles avec l'interprétation du sujet). Ainsi, **Int** ϕ est vraie dans le monde actuel (c'est-à-dire : il est vrai que d'après l'interprétation du sujet, ϕ), si et seulement si dans tous les mondes compatibles avec l'interprétation du sujet, ϕ est vraie.

On peut considérer que le mode de discours intentionnel est un mode de discours déclaratif parmi d'autres qui doit viser une description adéquate du monde : dans ce cas, il faut supposer que l'opérateur **Int** est factif, puisqu'une déclaration de ϕ selon ce mode sera réputée vraie à condition que ϕ soit vraie, et le cas *de objecto* se réduira au cas *de re* correspondant. On aura retrouvé l'interprétation réaliste du Moi. A l'opposé, on peut considérer que le mode de discours intentionnel n'est pas descriptif, notamment en s'inspirant de la stratégie de l'interprète de Dennett (1987). En adoptant le mode de discours intentionnel, je peux prêter des croyances et des désirs à toutes sortes de créatures : à moi-même et à mes interlocuteurs, mais aussi aux chiens (« (D'après moi,) il veut attraper la balle »), aux ordinateurs (« (D'après moi,) il veut protéger sa reine »), etc. Si les déclarations à la première personne sont comprises de cette manière, cela signifie que le cas *de objecto* ne se réduit pas au cas *de re* et que l'on peut concevoir le discours et la pensée égologiques comme portant sur le Moi sans pour autant considérer que le Moi existe¹⁷. Dans ce cas, l'opérateur **Int** est donc supposé non

¹⁶ De ce point de vue, il est semblable à d'autres opérateurs modaux comme l'opérateur proposé par Lewis, 1978, pour analyser le discours fictionnel (« D'après l'Odyssée, ... »). Il s'agit bien alors d'un mode de discours particulier plutôt que de l'expression d'une attitude particulière.

¹⁷ Passer à un mode non descriptif de discours n'implique pas qu'il faille renoncer à tout concept sérieux de vérité, si ce mode de discours s'avère indispensable pour notre interaction pratique avec le monde. Cf. Rebuschi & Renaud, 2011.

factif¹⁸. Dans ce qui suit, je fais cette hypothèse pour détailler ce que pourrait être une conception antiréaliste du Moi.

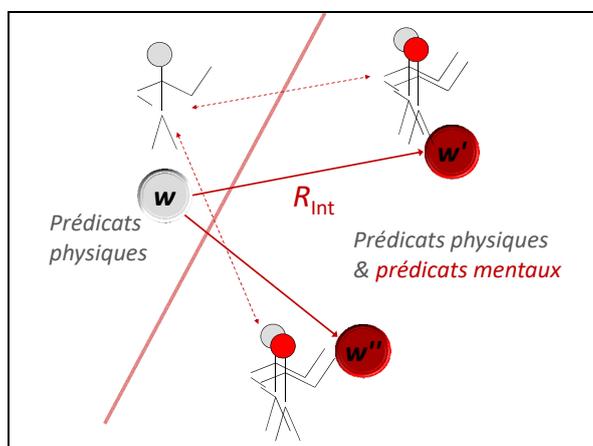


Fig.5. *Sphère physicaliste et sphère intentionnelle*

L'image qui en ressort est celle d'une séparation entre la sphère intentionnelle (l'ensemble des mondes colorés en rouge sur la Figure 5, qui sont les mondes compatibles avec l'interprétation du sujet), et la sphère physicaliste qui comprend, notamment, le monde actuel. Le personnage en gris représente le corps du sujet, le personnage en rouge représente son Moi. Dans le monde actuel le sujet se réduit ainsi à son corps. Le Moi est un objet purement intentionnel : il n'existe qu'à l'intérieur de la sphère intentionnelle et en particulier, il n'existe pas dans le monde actuel.

Mais comme l'a noté Castañeda (1968), certains prédicats attribués en première personne semblent concerner son corps : si un sujet déclare « Je pèse 160 livres », il ne parle pas de son Moi. J'emprunte à Castañeda l'idée d'une séparation nette entre prédicats mentaux (comme « être en colère » ou « croire qu'il va pleuvoir ») et prédicats physiques, lorsqu'ils sont attribués à la première personne. Les premiers s'appliquent au Moi, les seconds s'appliquent au corps ou à une combinaison du corps et du Moi. Cela suppose qu'une relation particulière lie le corps du sujet à son Moi, de telle sorte que le sujet s'identifie avec son Moi (Perry, 1998).

Cette relation d'identification (représentée par des lignes en pointillés sur la figure) lie des individus situés dans différents mondes possibles. C'est un cas particulier, parmi d'autres, de relation *inter-mondaine*. Les relations inter-mondaines peuvent être appréhendées par une simple extension du vocabulaire et de la sémantique de la logique modale quantifiée habituelle (Wehmeier, 2012). Il ne peut pas s'agir de relations physiques car on ne voit pas bien ce que pourrait signifier une interaction physique voire causale entre différents mondes possibles (Lewis, 1986). Il s'agit donc de relations non-physiques, qu'il faut comprendre, étant donné le type de mondes possibles en jeu, comme des relations intentionnelles.

¹⁸ Il semble qu'à défaut de l'axiome T : $\mathbf{Int} \varphi \rightarrow \varphi$ (restreint aux formules φ closes), l'opérateur pourrait se conformer à un axiome plus faible, employé en logique déontique pour formaliser l'obligation : $\mathbf{Int} (\mathbf{Int} \varphi \rightarrow \varphi)$.

Lorsqu'un sujet entretient une attitude *non propositionnelle* à l'égard d'un individu (quand le sujet *admire* cet individu, qu'il le *désire*, qu'il le *hait*), il ne faut pas chercher dans le monde actuel une relation (physique) entre le sujet et cet individu objet de son attitude. C'est une relation intentionnelle, non-physique, qui relie le sujet du monde actuel à cet individu dans les mondes de la sphère intentionnelle (que cet individu existe ou non dans le monde actuel, c'est-à-dire qu'il s'agisse d'un individu réel ou d'un individu purement intentionnel).

Par la pensée en première personne, le sujet s'*identifie* au Moi. Il peut par ailleurs *admirer* ou *détester* son Moi, ce qui met en jeu autant d'autres relations du sujet à son Moi. Mais il doit préalablement procéder à cette identification. Dans le cadre sémantique proposé ici, la relation d'identification est une extension inter-mondaine de l'identité intramondaine, et elle est donc conçue comme une relation intentionnelle¹⁹. Le cadre sémantique ne spécifie rien de plus quant à sa constitution ou sa construction. Il s'agit (techniquement) d'une *ligne de monde* suivant l'appellation de Hintikka²⁰, qui dans son approche correspond à un mode d'identification inter-mondaine (*cross-identification*) reposant sur le cadre de référence du sujet (cf. Hintikka, 1970) : le Moi est un objet *perspectivement individué*, et non un objet publiquement individué; il est appréhendé par accointance et non par description (cf. Hintikka, 1996). La méthode d'identification centrée sur le sujet rend compte de la spécificité de l'identification du sujet au Moi.

En disant cela, on n'a toujours cependant rien affirmé concernant le processus d'identification. S'agit-il d'une ré-identification à un Moi préexistant ? La relation est-elle à l'inverse une construction par le sujet, constitutive de l'identité du Moi ? L'option antiréaliste qui fait du Moi un objet purement intentionnel fait pencher en faveur de la seconde hypothèse. Mais il reste de l'espace pour une approche compréhensive de la construction effective de la relation d'identification et du Moi : on est alors au-delà de ce que le cadre sémantique peut prétendre éclairer.

6. Conclusion

Récapitulons. A l'encontre des conceptions dominantes qui considèrent les attitudes *de se* comme des cas particuliers d'attitude *de re*, j'ai proposé de les considérer comme des cas particuliers d'attitudes *de objecto* au moyen de l'introduction d'un opérateur modal supposé enrégimenter le mode de discours intentionnel. Grâce à cette stratégie, doublement inspirée des idées de Hintikka, j'ai atteint l'objectif d'une analyse sémantique de « je » qui soit ontologiquement neutre, puisqu'elle est compatible avec un réalisme au sujet du Moi, comme avec un antiréalisme. Si l'on veut une sémantique adaptée à sa position, réaliste ou

¹⁹ Cf. Rebuschi, 2012 pour une discussion plus fournie de la correspondance entre relations inter-mondaines et relations intentionnelles, comme la mobilisation de l'identité inter-mondaine pour rendre compte des attitudes *de se*.

²⁰ Ce n'est pas le cas des autres relations intentionnelles comme *admirer* ou *détester*, d'où ma préférence pour l'approche en termes de relations inter-mondaines plus générale que celle en termes de lignes de mondes.

antiréaliste, il faut trancher sur la factivité de l'opérateur, donc émettre quelque hypothèse complémentaire sur la nature du mode de discours intentionnel.

Je pense avoir rempli le programme minimum annoncé en introduction, c'est-à-dire avoir offert une clarification de la structure conceptuelle des pensées égologiques. Le sujet appréhende son Moi avant tout comme un objet intentionnel, au même titre que les autres objets. Il a une relation particulière vis-à-vis de ce Moi puisqu'il s'identifie à lui. Mais cela ne fait pas pour autant du Moi une entité dans le monde, même si le sujet est dans le monde. *Ensuite*, on peut défendre l'une ou l'autre option. On peut être réaliste, considérer finalement que le Moi est dans le monde, et que nos attitudes *de se* sont des attitudes *de re* (on retrouve ainsi l'analyse sémantique dominante); mais on peut également être antiréaliste et défendre avec Castañeda (1968), dans une veine wittgensteinienne, que le Moi « est une entité à l'extérieur du monde qui doit être identifiable en termes d'entités dans le monde », et que les attitudes *de se* sont irréductibles à des attitudes *de re*. Aucune de ces deux positions n'est impliquée par la seule sémantique : il faut avoir fait le choix du réalisme ou de l'antiréalisme.

Ma stratégie remplit-elle le programme maximum ? Eclaire-t-elle d'une quelconque façon la nature du Moi ? Par principe cela paraît impossible : on ne peut pas d'une part prétendre offrir une sémantique ontologiquement et métaphysiquement neutre, et d'autre part vouloir en tirer des conséquences ontologiques ou métaphysiques. On peut cependant observer quelques caractéristiques du Moi pour qui reprendrait cette sémantique tout en adoptant une conception antiréaliste du Moi, donc en supposant, comme je l'ai fait à la fin de la section précédente, que l'opérateur **Int** est non factif.

Le Moi est alors un objet purement intentionnel. Il ne fait pas partie de l'ontologie, de l'ameublement du monde actuel, ce n'est pas une substance mais une entité qui dépend des attitudes intentionnelles des sujets. Mon propre Moi est construit par mes attitudes intentionnelles, au travers de mes auto-attributions de prédicats mentaux. Pour reprendre une expression de Dennett (1992), proche en cela de Ricœur, il est comme un centre de gravité narrative. C'est en racontant des histoires à la première personne que le Moi se constitue.

En étant analysé comme un objet uniformément choisi dans un ensemble de mondes possibles plutôt que comme une entité existant dans le monde actuel, le Moi est un objet *essentiellement* incomplet²¹. Il n'est pas épistémiquement incomplet, au sens où on ne parvient pas à en avoir une connaissance complète — à ce compte-là, la plupart des objets ordinaires, actuellement existants, sont peut-être épistémiquement incomplets. Il est essentiellement incomplet au sens où certains prédicats mentaux ne sont ni vrais ni faux de lui, tant qu'ils ne lui ont pas été attribués ni déniés.

La narration n'est pas un monologue : le Moi se constitue par les dialogues (comme ce qui est désigné par un « tu »), et par les histoires que d'autres racontent à son propos (comme ce qui est désigné par un « il » ou un « elle »). Le Moi dépend de la perspective d'un sujet au point qu'il est n'est pas aisé de savoir quel sens donner à l'idée que deux sujets appréhendent le

²¹ Le Moi apparaît ainsi très proche des objets fictionnels selon les conceptions habituelles. C'est lié à la proximité (mentionnée plus haut) de l'opérateur **Int** avec les opérateurs fictionnels.

même Moi²². Quand je m'adresse à Mathilde, j'interprète ce qu'elle dit comme traduisant des attitudes de son Moi ; j'ai donc une image de son Moi, et un autre interlocuteur construira une autre image de son Moi. Si le Moi de Mathilde n'est pas une entité substantielle au-delà de ce que nous interprétons, est-ce que cela a bien un sens de dire que ses deux interlocuteurs appréhendent le Moi de Mathilde ? Est-ce que nous ne construisons pas deux Moi distincts ?

Il semble que l'on puisse s'en tirer par une décision. C'est une norme de l'interprétation que de s'accorder intersubjectivement sur l'identité d'un Moi. Mais ce n'est pas une norme arbitraire. Elle s'appuie sur la relation particulière qui lie Mathilde à son propre Moi : quand Mathilde s'exprime à la première personne, comme je l'ai souligné dans la section précédente, elle s'identifie à ce Moi. Cela fait autorité et si Mathilde n'est pas en plein délire, ses interlocuteurs acceptent et se plient à cette identification. De ce point de vue aussi il y a bien une spécificité du discours en première personne : l'identification du sujet à son Moi est une relation qu'elle ne partage pas avec les autres sujets. L'autorité du sujet s'étend certainement à ses états mentaux qualitatifs (comme la douleur, la tristesse, la joie). Mais l'autorité du sujet sur son propre Moi ne va guère au-delà. L'attribution de nombreux prédicats mentaux au Moi par le sujet lui-même, notamment concernant ses croyances, désirs, ou intentions, n'est ni plus ni moins vraie que celle effectuée par d'autres de ses interprètes²³.

²² Le problème est celui de l'identité intentionnelle, pointé par Geach, 1967, avec l'exemple fameux de Hob-Nob : qu'est-ce qui fait que deux sorcières (inexistantes) redoutées par Hob et Nob sont distinctes ou identiques ?

²³ Le contenu de ce travail a été partiellement présenté au colloque Semantics & Philosophy in Europe (SPE4, Bochum, 26 septembre-1er octobre 2011), et au colloque « Lacan avec Hintikka » (Institut Finlandais de Paris, 21-22 novembre 2011). Je tiens à remercier l'audience de ces colloques, ainsi qu'Elizabeth Leméteil, Marion Renauld, Tero Tulenheimo et Denis Vernant, pour leurs discussions et commentaires utiles sur des versions antérieures de cet article.

Références bibliographiques

Castañeda, H.-N.

- 1968 « On the Phenomeno-Logic of the I ». *Proceedings of the XIV International Congress for Philosophy* 3, 260-266.

Chauvier, S.

- 2001 *Dire "Je". Essai sur la subjectivité*. Paris, Vrin.

Dennett, D.

- 1987 *The Intentional Stance*. Cambridge Mass., The MIT Press
1992 « The Self as a Center of Narrative Gravity ». In F. Kessel, P. Cole & D. Johnson (eds.), *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ, Erlbaum.

Fiengo, R., May, R.

- 2006 *De Lingua Belief*, Cambridge Mass., The MIT Press.

Geach, P.T.

- 1967 « Intentional Identity ». *The Journal of Philosophy* 74, 627—632.

Hintikka, J.

- 1962 *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell, Cornell University Press.
1970 « On Attributions of "Self-Knowledge" ». *The Journal of Philosophy* 67, 3, 73-87.
1996 « *Cogito ergo quis est?* » *Revue Internationale de Philosophie*, 50, 195, 5-21.
2003 « A Second-Generation Epistemic Logic and its General Significance ». In V.F. Hendricks *et al.* (eds.), *Knowledge Contributors*, Dordrecht, Kluwer 2003: 33-55.

Hintikka, J., Sandu, G.

- 1997 « Game-Theoretical Semantics ». In J. van Benthem & A. ter Meulen (eds.), *Handbook of Logic and Language*, Cambridge (Mass.), MIT Press, 361-410.

Kripke, S.

- 1972 « Naming and Necessity ». In D. Davidson & G. Harman (eds.), *Semantics of Natural Language*, Dordrecht, Reidel, 253-355. Tr. fr. P. Jacob & F. Récanati, *La logique du nom propre*, Paris, Minuit, 1982.

Lewis, D.

- 1978 « Truth in fiction ». *American Philosophical Quarterly* 15 (1), 37-46.
1986 *On the Plurality of Worlds*. Oxford, Basil Blackwell.

Maier, E.

- 2009 « Presupposing Acquaintance: A Unified Semantics for *De Dicto*, *De Re* and *De Se* Belief Reports ». *Linguistics and Philosophy* 32, 429-474.

Perry, J.

- 1998 « Myself and I ». In M. Stamm (ed.), *Philosophie in Synthetischer Absicht. Synthesis in Mind*, Stuttgart, Klett-Cotta, 83-103.

Quine, W.V.O.

- 1956 « Quantifiers and Propositional Attitudes ». *Journal of Philosophy* 53, 177-187.

Rebuschi, M.

2011 « Le cogito sans engagement ». *Igitur*, Vol. 3, n° 2, 1-25.

2012 « Représentations de soi et modalités ». In F. Drapeau Contim & S. Motta (eds.), “Modalités : Sémantique & épistémologie”, *Philosophia Scientiae* 16/2, 173-194.

Rebuschi, M., Renauld M.

2010 « Fiction, Indispensability and Truths ». In F. Lihoreau (ed.), *Truth in Fiction*, Frankfurt, Ontos Verlag, 245-286.

Rebuschi, M., Tulenheimo, T.

2011 « Between *De Dicto* and *De Re: De Objecto* Attitudes ». *The Philosophical Quarterly*, Vol. 61 No. 245, 2011, 828-838

Wehmeier, K.

2012 « Subjunctivity and Cross-World Predication ». *Philosophical Studies*, 159/1, 107-122.

Wittgenstein, L.

1964 *Philosophische Bemerkungen*, Oxford, Blackwell. Tr. Fr. J. Fauve, 1975, *Remarques philosophiques*, Paris, Gallimard.