

ANSCOMBE AURAIT-ELLE ÉTÉ RELATIVISTE ?

Marie GUILLOT

(CNRS/Institut Jean Nicod)

L'ambition de G.E.M. Anscombe dans « La première personne » est de réfuter la conception dominante de la signification du terme « je ». Comme le rappelle Brian Garrett dans l'étude incisive¹ qu'il a consacrée à l'article, la position aujourd'hui majoritaire à l'égard de « je » comporte deux aspects : une thèse référentielle et une thèse indexicale.

La *thèse référentielle* veut que la fonction de « je » soit de référer à un objet particulier. « Je » serait ainsi, comme les noms propres, un désignateur singulier.

La *thèse indexicale* spécifie le mode de signifier qui permet à « je » de remplir cette fonction. Elle traite « je » comme un terme indexical, à l'instar de « ceci », « tu », « maintenant », « ici », « à droite », « hier matin », etc. La particularité des expressions indexicales tient à ce que leurs référents (les objets qu'elles visent au sein du réel) varient de façon systématique avec le contexte d'énonciation, au lieu d'être fixés une fois pour toutes par une convention baptismale, comme c'est le cas pour les noms propres (ainsi « Molière » et « Madonna » désignent toujours Molière et Madonna, respectivement). Plutôt qu'une convention de baptême, les termes indexicaux exploitent, pour viser leur référent, une règle de forme « token-réflexive »² indiquant quel genre de chose leurs occurrences peuvent représenter, en fonction de la relation concrète existant entre cette chose et lesdites occurrences. Par exemple, chacune des occurrences de « maintenant », « ici » et « toi » désigne, respectivement, le moment, le lieu et le destinataire de l'énonciation de *cette* occurrence. La thèse indexicale au sujet de « je » affirme qu'il en va de même pour ce dernier terme. Le référent de « je » serait alors déterminé par la règle suivante : chacune des occurrences de « je » renvoie à celui qui a énoncé *cette* occurrence.

Anscombe exprime son opposition à la thèse référentielle en des termes sans équivoque : «*Je* n'est ni un nom, ni aucune autre espèce

¹ Garrett (1997), auquel la structure de la présente introduction et la section 2 ci-dessous doivent beaucoup.

² Le terme évoque une relation réflexive impliquant l'occurrence (*token*, en anglais) de l'expression indexicale elle-même.

d'expression dont le rôle logique serait de référer.» (p. 60). Il ne s'agit pas simplement de dire que « je » échoue dans une prétention à faire référence, comme ces noms vides – « Yéti », « Père Noël » – dont les référents supposés n'existent pas. Il est certes, pour « je » comme pour ces noms fictionnels, des raisons métaphysiques à la vacuité référentielle : le seul référent que « je » (adéquatement compris) pourrait avoir tâche de viser serait, argue Anscombe, un *Ego* cartésien, conclusion qu'elle rejette comme tributaire d'une ontologie du sujet par trop improbable. Mais la raison la plus profonde est d'ordre sémantique³ : référer ne serait en aucune manière la fonction de « je », pas plus qu'elle ne l'est pour « il » ou « ça » dans « Il pleut » ou « ça pique ». Le rejet de la thèse indexicale découle de la négation de la thèse référentielle – si « je » ne réfère pas, à plus forte raison ne peut-il référer sur un mode indexical – mais il fait l'objet d'une discussion à part. Pour établir chacune de ses deux thèses négatives, Anscombe déploie un maître-argument, lui-même adossé à une expérience de pensée : l'exemple de la communauté des porteurs du nom « A » pour la thèse anti-indexicale, et le scénario de la cuve d'insensibilité pour la thèse anti-référentielle – deux ingénieuses fictions destinées à frapper pour longtemps l'imaginaire philosophique.

Ce qui suit vise quatre buts. Rappeler, tout d'abord, les raisons sérieuses que linguistes et philosophes du langage ont dès longtemps trouvées de résister à la thèse anti-référentielle d'Anscombe, restée minoritaire. Chercher, par conséquent, les failles de son argumentaire, dont je ne pointerai que quelques-unes. Proposer chemin faisant quelques clarifications touchant des notions séminales qu'Anscombe a, pour certaines, eu l'immense mérite de rendre centrales dans le débat philosophique, mais dont l'emploi parfois ambigu qu'elle en fait est en partie responsable des défauts de l'argumentation. (La principale source de confusion, comme on le verra, est l'absence de distinction entre l'auto-référence linguistique et la « perspective en première personne » dans la pensée.) Montrer enfin qu'une fois ces clarifications faites, et à la lumière du relativisme sémantique contemporain, il est possible de reformuler la position anti-référentielle d'Anscombe en une version plus prometteuse mais aussi plus modérée, et limitée au cas de la première personne *mentale*.

³. Force est néanmoins de remarquer qu'Anscombe ne fournit pas d'argument indépendant pour cette interprétation sémantique et plus radicale de la vacuité référentielle de « je ». Le seul argument offert, de facture métaphysique, concerne « l'absurdité » qu'il y aurait à tenir le sujet pour un *Ego* cartésien.

1. Une position contre-intuitive

Sa thèse anti-référentielle place Anscombe dans une lignée d'auteurs de tradition anti-cartésienne : Lichtenberg, Schlick et Wittgenstein, collectivement visés par la critique de Strawson (1959)⁴ contre ce qu'il appelle la doctrine « de l'absence de sujet » ; mais aussi Hume, Mach et Carnap. Les adversaires de la thèse anti-référentielle sont toutefois beaucoup plus nombreux que ses défenseurs. C'est que les raisons positives de tenir « je » pour un terme référentiel abondent, rendant la position d'Anscombe contre-intuitive et, aux yeux de beaucoup, inacceptable.

Notons tout d'abord le fait, concédé par l'auteure, que « je » possède la même « place syntaxique » – autrement dit la position de sujet dans des énoncés de forme prédicative – qu'un terme référentiel. Anscombe objecte qu'il en va de même pour « il » dans « Il pleut », sans qu'il vienne à l'idée de quiconque de traiter « il », dans ce cas, comme un désignateur singulier. La syntaxe, plaide-t-elle, ne serait pas un guide fiable pour s'orienter sur le terrain de la référence. Mais l'objection, souligne Garrett⁵, est superficielle. Le caractère non-référentiel de « il » dans « Il pleut » est corroboré par d'autres tests, qui ne s'appliquent en revanche pas à « je ». Ainsi, de « Il pleut » on ne saurait inférer « Quelque chose pleut » ; tout au contraire, de « J'ai mal » il s'ensuit bien que « Quelqu'un a mal ».

Par ailleurs, le dense réseau de liens vériconditionnels unissant les énoncés comportant « je » à différents énoncés indiscutablement référentiels est difficile à expliquer si « je » ne désigne rien. Comment comprendre, par exemple, que de « Je suis gauchère » et « Je suis Elizabeth » il découle qu'« Elizabeth est gauchère », si « je » et « Elizabeth » ne sont pas réputés référer au même objet ? Et comment, sans cette même supposition, comprendre que dans tout énoncé incluant « je » prononcé par la personne S, « je » puisse être remplacé *salva veritate* par « S » ? De même, une robuste intuition nous fait penser que si Elizabeth dit « Je suis triste », Ludwig (à Elizabeth) « Tu es triste » et Sydney (au sujet d'Elizabeth) « Elle est triste », alors ces trois personnes ont, en un certain sens évident, *dit la même chose*. Or Ludwig et Sydney ont indéniablement parlé d'un individu défini ; Elizabeth, par conséquent, aussi. La position négative d'Anscombe

⁴. Strawson (1959), p. 95 sq.

⁵. Garrett (1997), p. 510.

l'obligerait à refuser l'intuition que les trois protagonistes disent la même chose. Le corollaire, malaisément réconciliable avec une théorie plausible de la communication, serait que le locuteur d'une phrase contenant « je » et son allocutaire ne donnent pas la même interprétation à l'énoncé, même s'ils parviennent à se comprendre.

A ces différents tests fondés sur la vériconditionnalité – inférences, substitutions *salva veritate* et intuitions relatives au fait de « dire le même » – s'ajoute enfin la caution des reprises anaphoriques autorisées par « je ». Considérons : « Elizabeth et moi sommes inséparables. Nous ne nous quittons jamais. » Dans cet exemple, le pluriel de l'anaphorique « nous » indique que plusieurs référents de discours ont été introduits dans la phrase précédente, contrairement à ce qui serait le cas si seul « Elizabeth », et non « moi », référerait.

Une dernière raison *prima facie* de s'opposer à la thèse anti-référentielle au sujet de « je » est que la cohérence commanderait alors de généraliser la position à tous les indexicaux – « ceci », « tu », « ici », « maintenant », etc., dont « je » partage les principales caractéristiques syntaxiques, sémantiques et épistémiques, en particulier la référence garantie⁶. Il serait cependant difficile de nier que « maintenant » désigne

⁶ Anscombe dénie explicitement cette dernière propriété à « ceci », et semble implicitement en réserver l'extension au cas de « je ». C'est là une position très contestable, même si l'ambiguïté de la notion de 'référence garantie' utilisée dans l'article (cf. la section 3 *infra*) en rend la discussion complexe. La 'référence garantie', entendue tout d'abord comme une garantie contre une référence sémantique vide, caractérise en principe tous les indexicaux simples (par opposition aux indexicaux complexes, comme « cette boîte », « cette figure ») ; Anscombe elle-même reconnaît d'ailleurs que tout emploi de « ceci » doit nécessairement « accrocher sur quelque chose ». Mais les mots utilisés par l'auteure suggèrent qu'elle identifie plutôt la 'référence garantie' à « l'immunité aux erreurs d'identification » mise en évidence par Wittgenstein. (Comme on le verra, cette assimilation est certainement une erreur.) Or, Shoemaker (1968) et Evans (1982) ont montré, au moyen d'arguments qu'il serait ici trop long de rappeler, que cette dernière propriété s'étend bien aux démonstratifs comme « ceci » ; McGinn (1983) l'étend aussi aux indexicaux temporels et démonstratifs. Une troisième option consiste à comprendre, par 'référence garantie', l'impossibilité de toute dissociation entre référence sémantique (ce que le mot employé veut dire) et référence du locuteur (ce que ce dernier voulait dire en employant le mot). Anscombe fournit un contre-exemple à cette forme de garantie dans le cas des démonstratifs : « Voici tout ce qui reste du pauvre Jones », prononcé en indiquant une boîte dont le locuteur croit, à tort, qu'elle renferme les cendres du malheureux. La référence sémantique est la boîte ; la référence du locuteur, suggère Anscombe, les cendres. Deux objections sont ici possibles. D'une part, comme le montre Coliva (2003), de tels exemples de dissociation sont possibles avec « je », dans l'exacte mesure où l'on accepte leur possibilité pour « ceci » (cf. *infra*) ; l'homogénéité de « je » et des autres indexicaux n'est donc nullement remise en cause par la 'référence garantie' ainsi entendue. D'autre part, on peut arguer que l'intention référentielle primaire du locuteur, en tant qu'il maîtrise le sens des démonstratifs, est de désigner la chose, quelle qu'elle soit, que son geste ostensif discrimine ; la boîte donc, plutôt que les cendres. Là encore, l'asymétrie entre « je » et les autres indexicaux est douteuse.

un temps, « ici » un lieu, ou « tu » une personne.

Les conséquences de la thèse anti-référentielle sont incompatibles avec les observations ci-dessus, ce qui conduit à s'interroger sur les failles probables de l'argumentaire d'Anscombe. J'examine tour à tour, dans ce qui suit, les deux fils dialectiques principaux de l'article.

2. « Alphabet city » : l'argument contre la thèse indexicale

Le premier argument d'Anscombe est dirigé contre la thèse indexicale. Il est incorporé à un argument plus général, destiné à montrer que toute théorie voulant expliquer le sens de « je » est soit circulaire, soit insuffisante. D'après Anscombe, toute approche de type référentiel doit rendre compte des effets d'opacité créés par l'emploi de « je », ou de « lui-même » au discours indirect : on ne saurait par exemple remplacer « Smith sait qu'il⁷ n'est autre que Smith » par « Smith sait que Smith n'est autre que Smith », sans passer à côté d'un élément essentiel de la signification (et ce bien que la substitution préserve la valeur de vérité de la phrase, comme on l'a appelé plus haut). Or, toute tentative de spécifier le sens de la première personne pour capturer cet élément est vouée, d'après Anscombe, à devoir le faire en des termes qui réintroduisent « je » – par exemple en précisant que ce que Smith sait le dispose à accepter « Je suis Smith » –, présupposant ainsi cela même qui est en question.

Ce n'est pas le cas de la théorie indexicale, d'après laquelle le sens linguistique de « je » est épuisé par la règle suivante : une occurrence donnée de « je » réfère à quiconque en est l'énonciateur. Mais, d'après Anscombe, cette théorie n'échappe à la circularité qu'au prix de son insuffisance.⁸ En effet, cette approche ignore un trait essentiel de l'usage de « je », à savoir qu'il exprime la conscience de soi du sujet. L'argument s'appuie sur une première expérience de pensée :

⁷ Conformément aux observations pionnières de Castañeda, le pronom « il » est ici un *réflexif indirect*, forçant une lecture « *de se* » de la phrase : on doit comprendre que Smith pense ici à lui-même « en tant que lui-même ».

⁸ L'article d'Anscombe précède l'essor des théories de l'indexicalité à partir de la fin des années 1970, notamment chez Evans, Perry, Kaplan et Stalnaker. C'est en partie pour cette raison que l'attaque menée par Anscombe contre la théorie indexicale peut paraître réductrice aux yeux du lecteur contemporain. Ainsi, Anscombe n'est pas en mesure de faire la distinction entre un supposé « sens » de « je », dans l'acception frégréenne, et le « caractère » de « je », tel que le définira Kaplan par la suite. Surtout, l'accusation d'insuffisance peut, rétrospectivement, sembler expéditive au regard de la sophistication des théories de l'indexicalité qui ont vu le jour après la publication de l'article. Une discussion beaucoup plus approfondie serait aujourd'hui requise.

«Imaginons une société dans laquelle chacun est étiqueté au moyen de deux noms. L'un apparaît sur le dos et en haut de la poitrine et ces noms, que le porteur ne peut pas voir, sont divers: disons de 'B' à 'Z'. L'autre nom, 'A', est imprimé sur la face interne des poignets, et identique pour chacun. Lorsqu'il s'agit de rapporter les actions d'autrui, chacun emploie le nom qui figure sur la poitrine ou le dos, s'il peut voir ce nom ou s'il est habitué à le voir. Chacun apprend aussi à réagir à l'énoncé qui figure sur sa poitrine ou son dos, dans les circonstances et selon la manière dont nous répondons nous-mêmes à l'énoncé de notre nom. Les rapports [de chacun] sur ses propres actions, délivrés directement par observation, sont formulés en employant le nom qui figure sur le poignet.» (Anscombe, 1975, p. 49)

Le scénario imaginé par Anscombe est conçu pour émuler l'usage de « je » tel que les tenants de la thèse indexicale le comprennent. Dans la communauté évoquée, en effet, « A » est le nom que chacun emploie pour se désigner lui-même ; « A » devrait donc pouvoir remplacer « je » dans la règle réflexive avancée par la thèse indexicale. Or, dans la situation décrite, tout jugement de la forme « A est F » se fonde exclusivement sur des sources d'information publiquement accessibles : l'observation du corps et du comportement, l'inférence et le témoignage d'autrui (si B entend dire que « B est F », par exemple, il pourra en déduire que « A est F » dès lors qu'il connaît son nom individuel, et peut juger que « A est B »). Il serait donc compatible avec le scénario d'Anscombe que ses personnages, bien qu'utilisateurs compétents de « A », soient tous dénués de conscience de soi. Mais tout usage compétent de « je » exprime la conscience de soi. Donc la thèse indexicale, qui traite « je » sur le modèle de « A », est fautive.

Deux des prémisses contenues dans l'argument ci-dessus s'avèrent contestables : (1) l'affirmation selon laquelle le scénario illustre la thèse indexicale, et (2) l'assertion que tout emploi correct de « je » exprime la conscience de soi. Par ailleurs, même à supposer que (1) et (2) soient vraies, il ne s'ensuivrait pas que la thèse indexicale soit intenable. Garrett⁹ souligne l'ambiguïté de la conclusion, qui admet deux interprétations : une conclusion *faible*, disant que la thèse indexicale *échoue* à expliquer le fait

⁹ Garrett (1997), p. 508.

que l'usage de « je » manifeste la conscience de soi ; et une conclusion *forte*, voulant que la thèse indexicale soit *incompatible* avec ce fait. Or, l'argument ne supporte nullement la conclusion forte, mais seulement, dans le meilleur des cas, la conclusion faible.

Pour poser la conclusion forte et réfuter la thèse indexicale, Anscombe aurait besoin d'une prémisse supplémentaire, selon laquelle une théorie du fonctionnement sémantique de « je » doit expliquer la corrélation entre les jugements à la première personne et la conscience de soi. Mais il n'y a pas de raison *a priori* de penser que ce problème épistémologique relève de la juridiction du sémanticien. Si l'auteure le suppose, c'est probablement parce qu'elle utilise, dans tout l'article, l'expression « première personne » pour désigner indifféremment l'auto-référence au moyen de « je » dans le langage, et la perspective « en première personne » dans la pensée. C'est le second phénomène qui est inséparable de la conscience de soi, et c'est, de la part d'Anscombe, une simple pétition de principe que de supposer implicitement la même chose du premier : aucun argument n'est fourni à l'appui de la thèse selon laquelle la « première personne » ne correspondrait qu'à un seul et même phénomène dans les domaines linguistique et psychologique. Or, à première vue, autre chose est de pouvoir penser à soi comme à soi-même ; autre chose, de pouvoir dire « je ». Il est de bonnes raisons de penser que, chez les humains, la première capacité est indépendante de la première, qu'elle précède et conditionne¹⁰. Si l'on évite la confusion entre première personne verbale et première personne mentale qu'Anscombe entretient tout au long de l'article, la prémisse (2) cesse également d'être évidente. L'auteure traite comme des anomalies les occurrences de « je » qui ne reflètent nulle conscience de soi, comme dans l'étiquette « bois-moi » sur la bouteille de Lewis Carroll, ou la mention « Je ne suis pas un cendrier ! » sur un pot de fleurs ; mais c'est là seulement une conséquence du présupposé selon lequel auto-référence linguistique et égocentricité mentale ne sont que deux facettes du même phénomène. Une fois ce présupposé identifié et mis de côté, il est tout à fait possible au défenseur de la thèse indexicale de tenir les usages « non-conscients » de « je » pour parfaitement normaux, et de considérer les emplois ordinaires de « je » par des personnes douées de

¹⁰. Voir Bermúdez (1998). Le caractère « en première personne » de la vie mentale d'un sujet exige certainement une forme de conscience de soi au moins implicite, mais pas nécessairement l'auto-référence mentale (laquelle repose sur un concept explicite de soi) ; à plus forte raison n'implique-t-elle pas la capacité à l'auto-référence linguistique.

conscience de soi comme un phénomène plus complexe, mobilisant bien davantage que le déploiement correct de « je »¹¹.

Mais les défauts de l'argument vont plus loin. Comme le montre Garrett¹², il ne soutient pas même la conclusion faible. En effet, la prémisse (1) est fautive. L'expérience de pensée d'Anscombe n'est pas une illustration fidèle de la thèse indexicale. Dans la communauté décrite, la référence d'une occurrence donnée de « A » n'est pas *déterminée* par une règle réflexive comme celle qui gouverne « je » (même si, pour des raisons contingentes, elle est la plupart du temps *conforme* à une telle règle) ; elle se fonde plutôt sur l'exploitation d'informations perceptuelles. On peut vérifier l'hétérogénéité entre « je » et « A » en remarquant que tout usage de « je » désigne celui qui en est la source, qu'il le veuille ou non ; mais qu'en revanche, dans la communauté fictive d'Anscombe, il peut arriver qu'une personne B réfère de manière correcte à une autre personne C sous le nom « A ». C'est le cas si B aperçoit la marque « A » sur le poignet de C, et prend ce poignet pour le sien. Dans ce cas, B peut faire l'affirmation vraie que « A est F » (par exemple, « A a une entaille sur le poignet ») sans pourtant dire quelque chose de vrai sur lui-même. C'est précisément ce que « je », si la thèse indexicale est vraie, rend impossible. En effet, d'après la thèse indexicale, et pour reprendre les mots de l'auteure, « *Si X fait des assertions avec « je » comme sujet, alors ces assertions seront vraies si et seulement si les prédicats ainsi utilisés dans l'assertion sont vrais de X* » (Anscombe, 1975, p. 55). Ainsi, la thèse indexicale exclut absolument que « A » puisse se substituer à notre « je ».

3. La cuve d'insensibilité : l'argument (anti-)cartésien contre la théorie référentielle

Le second argument central de l'article a pour cible la thèse référentielle au sujet de « je ». La ligne d'attaque prend ici la forme d'une

¹¹ Même à prendre au mot les demandes d'Anscombe, on peut du reste contester l'idée qu'une explication à la fois suffisante et non-circulaire du lien entre usage linguistique de « je » et pensée en première personne soit impossible. C'est précisément la tâche qu'Evans (1982) et ses émules se sont fixée. Dans ce type d'approche, le déploiement du concept de « je », et par extension l'usage ordinaire de « je » qui manifeste la conscience de soi, sont fondés non pas sur des sources d'information publiques, mais sur l'exploitation de canaux informationnels exclusivement accessibles au sujet - introspection, proprioception, kinesthésie, perception interne de la douleur et de la faim, etc. C'est là une version beaucoup plus raffinée du scénario des « A-users », et qui ne prête plus (immédiatement) le flanc aux critiques d'Anscombe.

¹² Garrett (1997), p. 509 et 511.

réduction à l'absurde. Schématiquement, d'après Anscombe, si « je » réfère, alors il jouit d'une 'référence garantie'. Mais la seule façon pour « je » de bénéficier d'une référence garantie dans toutes les situations possibles est de faire référence à une entité immatérielle, autrement dit une sorte d'*Ego* cartésien. Or, que « je » désigne un *Ego* cartésien est, pour l'auteure, une conclusion « absurde ». Par contraposition, il s'ensuit que « je » ne réfère pas.

Cet argument, comme le précédent, est vulnérable en plusieurs points, que les limites de cette introduction ne permettent pas de soulever tous. Je m'en tiendrai donc à un petit nombre de critiques sur certaines des prémisses. Au prix de simplifications importantes¹³, mais sans incidence sur la discussion à venir, l'argument peut être reconstruit de la façon suivante :

(P1) « Je » est un terme référentiel.

(P2) Tout usage normal de « je » jouit d'une 'référence garantie'.

(P3) Pour jouir de la référence garantie, un terme référentiel doit désigner son objet via une « conception », ou « manière d'être donné », qui assure à la fois (a) l'existence de l'objet, et (b) sa « présence physique ou réelle » à la conscience.

(P4) Même dans des conditions de privation sensorielle radicale, je serais toujours en mesure d'utiliser « je » normalement, et donc en conformité avec le principe de référence garantie.

(P5) Mais la seule chose à laquelle, dans des conditions de privation sensorielle, je puisse penser d'une manière qui garantisse à la fois l'existence de l'objet et sa « présence réelle » à ma conscience, c'est ma pensée elle-même, soit un *Ego* cartésien.

(P6) Donc, ce à quoi « je » réfère est un *Ego* cartésien.

(P7) Mais la théorie de l'*Ego* cartésien est « absurde » et inacceptable.

(C) Donc, par *modus tollens*, (P1) est fautive. « Je » n'est pas un terme référentiel.

¹³. Hamilton (1991) détaille les différents arguments contre la thèse anti-référentielle que l'on peut extraire de l'article. Je ne présente ici que le plus central, et ce d'une manière très ramassée qui en ignore la complexité et la subtilité.

Une seconde expérience de pensée résume les étapes principales (P1-P6) de l'argument :

«Supposons maintenant que je me retrouve dans un état de 'privation sensorielle'. Ma vision est éliminée, et je suis localement anesthésiée partout, peut-être immergée dans un bassin d'eau tiède; je suis incapable de parler, ou de toucher une quelconque partie de mon corps avec une autre. Je me dis à présent: 'Je ne laisserai pas une pareille chose se reproduire!' Si l'objet signifié par 'Je' est ce corps, cet être humain, alors, dans ces circonstances, il n'est pas présent à mes sens; et de quelle autre manière pourrait-il m'être présent? Ai-je pour autant perdu ce que je veux dire par 'Je'? Cette chose ne m'est-elle pas présente? En suis-je réduite, pour ainsi dire, à 'référer in absentia'? Je n'ai pas perdu ma 'conscience de soi'; et ce que je signifie par 'Je' ne peut pas être un objet qui ne m'est plus présent. Cela semble à la fois juste, en soi-même, et sera requis par la 'référence garantie' que nous sommes en train de considérer.

(...) Rien d'autre qu'un Ego cartésien ne fera l'affaire.»
(Anscombe, 1975, p. 58)

Anscombe jugeant cette conclusion, (P6), inadmissible¹⁴, elle choisit de rejeter (P1) et de préserver (P2)-(P5). Elle admet en particulier (P2), le principe de référence garantie pour « je », mais l'explique de façon déflationniste, en arguant qu'il est satisfait de manière triviale : « je » n'a pas du tout vocation à référer, aussi ne peut-il a fortiori référer « à vide », ni prêter aux erreurs d'identification: «il est exclu que nous saisissons le mauvais objet, et cela nous incite à penser qu'il est garanti que nous

¹⁴. Anti-cartésienne, Anscombe semble plutôt pencher pour une théorie animaliste du soi, dans la lignée de Strawson : «*La connaissance de soi est connaissance de l'objet que l'on est, de l'animal humain que l'on est.*» (p. 62). Les deux principaux arguments qu'elle présente, de façon un peu cursive, à l'encontre de la métaphysique cartésienne du soi sont les suivants. (i) Pour le cartésien, le soi est constitué dans et par le *cogito* ; ses conditions d'individuation dépendent donc essentiellement de l'occurrence présente de la pensée. Mais on voit mal ce qui fonde alors la continuité de l'identité personnelle à travers différents épisodes de pensée, et comment le sujet peut se ré-identifier dans la diachronie. (ii) La théorie cartésienne conduit à des paradoxes, et son interprétation donne lieu à des disputes « interminables, sans issue » et « qui s'auto-alimentent » ; Anscombe y voit le symptôme de son impuissance à fournir une notion univoque et cohérente du soi.

saisissons le bon objet. Mais la raison en est qu'il n'est pas question d'une quelconque saisie d'objet.» (1975, p. 59).

Au vu des considérations invoquées dans la section (1), cependant, le rejet de (P1) paraît excessivement coûteux. Les prémisses (P2) et (P3) semblent beaucoup plus discutables.

Si la prémisse (P2) fait difficulté, c'est d'abord en raison de la polysémie, et de la confusion relative, de la notion de 'référence garantie' (ci-après RG). Anscombe elle-même distingue plusieurs acceptions possibles, dont toutes seraient, d'après elle, satisfaites par « je ». L'étude minutieuse de la notion dans Coliva (2003) porte au contraire à penser qu'aucune, sauf la plus modeste (notée (1) ci-dessous), ne l'est de manière évidente¹⁵.

Anscombe distingue trois notions cumulatives de RG. Premièrement (1), toute énonciation de « je » garantit, par hypothèse, que son énonciateur existe (sans énonciateur, pas d'énonciation). Dans cette première acception minimale, et étant donnée la règle token-réflexive, RG est une *garantie contre la vacuité de la référence sémantique* : ce à quoi une occurrence de « je » prétend référer existe toujours. Deuxièmement (2), ce à quoi l'énonciateur a l'intention de référer en disant « je » existe toujours. Autrement dit, « je » possède une *garantie contre la vacuité de la référence du locuteur*.¹⁶ Troisièmement (3), quiconque dit « je » est assuré que ce à quoi il entend référer par ce mot, est bien ce à quoi le mot réfère. Dans les termes d'Anscombe, une expression X satisfaisant RG en ce dernier sens garantit «*non seulement qu'il y a une telle chose que X, mais aussi que ce que je prends pour X est X.*» (p. 57). RG est, dans cette version forte, une *garantie contre toute divergence entre la référence sémantique et la référence du locuteur*. Par ailleurs, Anscombe identifie ce troisième sens de RG avec (4) *l'immunité aux erreurs d'identification* (ci-après IEM). L'IEM est une propriété, d'abord mise en évidence par Wittgenstein (1958), puis définie par Shoemaker (1968), que possède un jugement de la forme « a est F » lorsqu'il est prémuni contre toute méprise sur l'objet, et ne saurait être vrai d'autre chose que a. Par exemple, quand je juge « j'ai faim » sur la foi de mes sensations internes, alors si ce jugement est vrai, c'est forcément de moi qu'il est vrai. Il n'y a

¹⁵. Coliva propose une ultime définition de RG, qui caractériserait bien « je », mais qui est en revanche absente du texte d'Anscombe.

¹⁶. La *référence sémantique* correspond à ce que le mot employé veut dire ; la *référence du locuteur*, à ce que ce dernier veut dire en employant le mot. La distinction, déjà opérante de façon implicite dans le texte d'Anscombe, a été théorisée deux ans plus tard par Kripke (1977).

nulle place pour une erreur sur la personne : il serait absurde de me demander « certes, quelqu'un a faim, mais est-ce bien moi ? »

Or, les jugements exprimés au moyen de « je », en réalité, ne sont pas tous immunisés contre les erreurs d'identification. Par exemple, lorsque je juge « Je suis le plus jeune de la promotion » en consultant le registre des inscrits, il est tout à fait possible que, lisant mal, je me trompe de personne, et que la propriété prédiquée soit vraie de quelqu'un d'autre. La prémisse (P2) est donc fautive si l'on entend RG au sens (4).

D'autre part, l'identification des sens (3) et (4) de RG est erronée. L'IEM est une propriété des jugements, et a trait à la prédication ; RG est une propriété des termes référentiels, et a trait à la référence. La confusion entre (3) et (4) repose sur un glissement sémantique entre deux notions d'identification. L'IEM concerne la nature des sources épistémiques permettant de déterminer les objets dont on peut prédiquer certaines propriétés ; RG concerne le mécanisme sémantique par le biais duquel certaines expressions sélectionnent leurs référents.

Mais rétablir la distinction ne sauve pas l'argument, car, comme le montre Coliva (2003), « je » ne satisfait pas davantage RG au sens (3) qu'au sens (4). Coliva donne l'exemple suivant. Supposons que j'aperçoive dans un miroir une silhouette, que je crois être la mienne, mais qui est en fait celle d'une autre personne (appelons-la Elizabeth) ; et derrière cette silhouette, un tableau. Si je déclare, sur la foi de cette perception, « Il y a un tableau très laid derrière moi », mon usage de la première personne (« moi ») ne fait pas coïncider la référence sémantique et la référence du locuteur, et contredit ainsi RG au sens (3). En effet, la personne à qui l'occurrence de « je » réfère (la *référence sémantique*) est forcément moi-même ; en revanche, la personne à laquelle j'ai l'intention de référer (la *référence du locuteur*), autrement dit celle que je crois être, est Elizabeth, réfléchi dans le miroir. La prémisse (P2), si l'on envisage RG au sens (3), est donc fautive.

J'ajoute que « je » ne vérifie pas non plus (en toute généralité) RG au sens (2). Dans notre exemple, il se pourrait très bien que l'illusion d'optique aille plus loin encore, et que *personne* ne soit réellement visible dans le miroir ; alors, la personne à qui j'entends référer en disant « je » n'existe pas.

Seule l'acceptation (1) de RG est incontestablement satisfaite par

tout emploi de « je » et rend vraie la prémisse (P2) ; mais, de l'aveu même d'Anscombe, la thèse indexicale n'a aucun mal à l'expliquer très simplement, sans avoir à accepter pour ce faire la prémisse (P3), ni donner la moindre prise à l'argument anti-référentiel.

La prémisse (P3), du reste, pose elle-même un certain nombre de problèmes. Anscombe affirme que la satisfaction de la 'référence garantie' exige l'intervention, dans le processus de référence, d'une « manière pour l'objet d'être donné » – un « mode de présentation » frégéen – qui garantisse à la fois (a) l'existence de cet objet et (b) sa « présence physique ou réelle » à la conscience. Autrement dit, le mode de référence doit être ostensif, ou *démonstratif* : je dois pouvoir penser à moi-même comme à « ce moi ». La condition (a) assure que RG au sens (1) est vérifiée ; la condition (b), que ses variantes (2) et (3) sont satisfaites. Il faut en effet, raisonne Anscombe, que l'objet auquel j'entends référer par « je » soit lui-même donné de façon transparente à ma conscience pour qu'aucune erreur sur son identité ne puisse se glisser dans le mécanisme référentiel – soit qu'un autre objet lui soit substitué à mon insu dans la référence sémantique effective de « je » (ce qu'exclut RG (3)), soit qu'il n'existe pas du tout (ce qu'exclut RG (2)). Or, dans l'état de privation sensorielle absolue qui serait le mien si j'étais plongé dans la cuve d'insensibilité, aucun objet matériel ne serait donné à ma conscience sur un mode ostensif. La seule chose qui serait présentée à ma conscience, et à laquelle je puisse référer de façon démonstrative, serait le fait que je pense. Puisque je pourrais toutefois continuer à dire « je » de façon correcte dans la cuve, le mot ne pourrait alors désigner qu'une chose pensante. Mais, argue Anscombe, cette conclusion de facture cartésienne est inadmissible. L'hypothèse initiale selon laquelle « je » réfère doit donc être fausse.

Diverses présuppositions contestables guident ce raisonnement. On peut tout d'abord mettre en doute l'idée que, pour faire infailliblement référence à quelque chose qui est essentiellement un corps, il faille avoir l'objet en question présent à la conscience *en tant que* corps. Le corps en question peut très bien être visé sous un autre mode de présentation – par exemple, en tant que chose pensante. Pour que l'on fasse à tout coup référence au corps en le saisissant sous l'angle de ses propriétés psychologiques, il suffit d'une garantie métaphysique, telle que l'identité entre la chose pensante et la chose matérielle.

Par ailleurs, l'affirmation qu'une référence démonstrative (même garantie) exige la présence effective de l'objet a été mise en question par Evans (1982). La référence démonstrative exploite des informations sensibles passant par certains canaux spécifiques. Par exemple, pour faire référence à une douleur par ostension, il faut mobiliser ce canal informationnel particulier qu'est la nociception (perception interne de la douleur). De même, pour faire démonstrativement référence à soi-même (en tant que soi-même), il faut posséder des sources d'informations dédiées à la connaissance de cet objet, telles que l'introspection, la proprioception, la kinesthésie, etc. L'exploitation de tels canaux assure la référence garantie. En effet, pour chaque personne, ces canaux informationnels ne peuvent être exploités que pour s'informer sur un seul et même objet – soi-même. (On ne peut connaître les pensées d'autrui par l'introspection, ni la position de ses membres par la proprioception.) Nulle erreur d'identification n'est donc possible. Mais mieux encore, l'existence de tels canaux informationnels dans mon architecture cognitive suffit à garantir que je me réfère bien à moi-même lorsque j'utilise le concept démonstratif (« je ») fondé sur leur exploitation habituelle, quand bien même ces canaux ne délivreraient aucune information au moment où je fais référence. Dans les termes d'Evans, « Un sujet n'a pas besoin que des informations lui soient effectivement accessibles par les canaux pertinents pour savoir qu'il existe exactement un objet auquel il est dispositionnellement relié à travers ces canaux. »¹⁷

Plus radicalement, on peut contester l'idée que la référence effectuée par « je » doive être démonstrative, et donc reposer, de façon immédiate ou au moins dispositionnelle, sur la présence de l'objet à la conscience. Anscombe considère tour à tour les différentes catégories de désignateurs singuliers : noms, pronoms, descriptions définies, démonstratifs, et argue que « je » ne peut appartenir qu'à cette dernière, seule à assurer la référence garantie. Elle ignore toutefois une dernière possibilité, à savoir que « je » soit un indexical « pur » ou « automatique » (Perry 1997). Les indexicaux purs sont ceux dont le référent est déterminé uniquement par une règle token-réflexive, sans qu'un geste démonstratif ou tout autre forme de « pointage » (éventuellement interne) soit en plus nécessaire. Un tel mode de référence bénéficie bien de la référence

¹⁷ Evans (1982), p. 216, note 21.

garantie, dans la mesure où la règle token-réflexive sélectionne le référent de façon mécanique, et où l'intention référentielle primaire du locuteur consiste à viser la chose, quelle qu'elle soit, qui satisfait la règle. Anscombe suppose que son argument contre la thèse indexicale exclut que « je » puisse opérer de cette façon. Mais, comme nous l'avons vu plus haut, l'argument n'est pas concluant. Il est donc loisible de considérer « je » comme un indexical pur et de rejeter le modèle démonstratif, ce qui permet de nier la prémisse (P3) et de bloquer l'argument anti-référentiel.

4. Première personne mentale et relativisme sémantique : perspectives pour une position néo-anscombiennne aujourd'hui

Même averti des failles argumentatives de l'article d'Anscombe, et des développements subséquents des théories de l'indexicalité, le lecteur contemporain ne pourra manquer d'être frappé par la finesse et la complexité de la réflexion, la fécondité philosophique des notions introduites, ni se défaire de la conviction que l'auteure pointe un problème central et profond dans le concept de première personne. L'intuition générale d'Anscombe ne pourrait-elle pas être préservée, indépendamment des modalités spécifiques de son argumentation dans ce texte ?

Comme on l'a vu plus haut, l'un des principaux points d'achoppement de l'article est l'absence de distinction entre deux sens, linguistique et psychologique, de la notion de « première personne ». Anscombe semble surtout s'intéresser à ce que c'est que *penser* à soi-même en tant que soi-même, mais n'aborde cette question qu'à travers celle des conditions de l'auto-référence dans le langage – laquelle peut, mais ne doit en aucune manière, exprimer la pensée en première personne. Or, c'est le traitement conjoint des deux problèmes qui rend la position d'Anscombe intenable. Les objections *prima facie* évoquées dans la section 1 ci-dessous concernent essentiellement la thèse anti-référentielle au sujet du mot « je », dont elles invoquent le comportement sémantique et syntaxique dans la langue. Mais une version limitée de la thèse qui ne concernerait plus que la pensée en première personne échapperait à la plupart de ces critiques. Certes, une partie des arguments ciblant l'anti-référentialisme, en particulier ceux fondés sur les transitions vériconditionnelles, s'appliqueraient aussi à l'auto-référence mentale, et aux pensées

explicitement égologiques. Toutefois, il existe un autre sens, plus fondamental, de la « première personne mentale ». C'est à ce niveau que l'anti-référentialisme a une chance de convaincre.

C'est une présupposition courante, mais rarement argumentée, de la philosophie de l'esprit contemporaine que les pensées en première personne – celles dont la possession fait de nous des sujets – se caractérisent par une certaine façon de se prendre soi-même pour objet. Par contraste avec cette approche classique, un certain nombre de philosophes du langage et de l'esprit – Andy Egan, François Recanati, Tamina Stephenson, Dilip Ninan, parmi d'autres – proposent de faire fonds sur les intuitions pionnières de Loar (1976), Lewis (1979), Chisholm (1981) et Sosa (1983), ainsi que sur les développements contemporains du relativisme sémantique, pour définir les cas primitifs de pensée en première personne, ou '*de se*', par l'absence de toute forme d'auto-référence, du moins explicite. L'intuition centrale est que certains types d'états mentaux – par exemple la perception, l'imagination située, le souvenir épisodique – peuvent être *égocentriques* sans être pour autant *égologiques* ; c'est-à-dire refléter un point de vue subjectif, sans pour autant représenter le sujet dans leur contenu. Au niveau le plus élémentaire, une pensée « en première personne » ne serait donc pas une pensée portant sur moi-même, mais une pensée portant sur un autre objet, saisi *depuis ma perspective subjective* singulière. Ainsi Recanati (2007), par exemple, affirme que le sujet n'est pas représenté dans le contenu (étroit) de ses pensées égocentriques les plus simples, comme « ça fait mal » ou « la chaise est à gauche ». Seule une *propriété* (par exemple « est douloureux ») est représentée dans ce contenu. De telles pensées 'concernent' toutefois le sujet de façon privilégiée, dans la mesure où elles ne peuvent pas être évaluées comme vraies ou fausses absolument, mais seulement *relativement* à leur possesseur. En d'autres termes, le sujet joue le même rôle que d'autres « circonstances d'évaluation » plus communément reconnues par les théoriciens du contenu (mental), telles que le monde possible, et, pour les relativistes, le temps ou le lieu. De ceci découle en particulier, selon Recanati (2010), la solution au problème de l'IEM, ou plutôt, en termes wittgensteiniens, sa « dissolution ». Comme elles ne contiennent pas d'auto-représentation, les pensées *de se* ne requièrent aucun effort particulier d'auto-identification de la part du sujet, et partant, ne laissent

logiquement place à aucune possibilité d'identification fautive. Cette conception modeste et 'a-référentielle' des pensées en première personne primitives a pour attrait supplémentaire de s'accorder avec les considérations d'économie cognitive motivant les théories de la cognition située et incarnée, ce qui lui confère une plausibilité psychologique supérieure à celle de l'approche auto-référentielle classique. On peut voir dans cette prometteuse approche contemporaine un retour à l'intuition anti-référentielle d'Anscombe, restreinte au terrain mental, et précisée grâce à l'apport d'un outillage sémantique plus sophistiqué dans le cadre relativiste.

Références choisies

- Bermudez J.-L. (1998) - *The Paradox of Self-Consciousness* (Cambridge, Mass. : The MIT Press)
- Chisholm R. (1981) - *The First Person* (Minneapolis : University of Minnesota Press)
- Coliva A. (2003) - The First Person: Error Through Misidentification, the Split between Speaker's and Semantic Reference, and the Real Guarantee (*Journal of Philosophy*, vol.100(8), p. 416-431)
- Egan A. (2006) - Secondary Qualities and Self-Location (*Philosophy and Phenomenological Research*, vol.72(1), p. 97-119)
- Evans G. (1982) - *The Varieties of Reference* (Edition posthume dirigée par J. McDowell. Oxford : Clarendon Press)
- Garrett B. (1997) - Anscombe on 'I' (*The Philosophical Quarterly*, vol.47(189), p. 507-511)
- Hamilton A. (1991) - Anscombian and Cartesian Scepticism (*The Philosophical Quarterly*, vol.41(162), p. 39-54)
- Kripke S. (1977) - Speaker's Reference and Semantic Reference (In French, P. A., Uehling Jr T. E. & Wettstein H. K. (eds.), *Studies in the Philosophy of Language*. University of Minnesota Press)
- Lewis D. (1979) - Attitudes *De Dicto* and *De Se* (*The Philosophical Review*, vol.88, p513-543)
- O'Brien L. (2007) - *Self-Knowing Agents*, Chapitre 2, « Arguments for the No-Reference View of 'I' » (Oxford : Oxford University Press)
- Perry J. (1997) - Indexicals and Demonstratives (in Hale, B. & Wright, C. A

- Companion to the Philosophy of Language*, Oxford, Blackwell)
- McGinn C. (1983) - *The Subjective View: Secondary Qualities and Indexical Thoughts* (Oxford : OUP)
- Ninan D. (2010) - *De Se Attitudes: Ascription and Communication* (*Philosophy Compass*, vol.5(7), p. 551-567)
- Récanati F. (2007) - *Perspectival Thought. A Plea for (Moderate) Relativism* (Oxford : OUP)
- Récanati F. (2010) - *Le soi implicite* (*Revue de métaphysique et de morale*, n°4/2010)
- Shoemaker S. (1968) - *Self-Reference and Self-Awareness* (*Journal of Philosophy*, vol.65(19), p. 556-567).
- Sosa E. (1983) - *Consciousness of the Self and of the Present.*(in Tomberlin J. E. (ed), *Agent, Language and the Structure of the World*. Hackett)
- Strawson P. (1959) - *Individuals* (Londres et New York : Routledge)
- Wittgenstein, L. (1958) - *The Blue and Brown Books* (Oxford : Blackwell)