

L'IDENTITÉ PERSONNELLE N'EST PAS CE QUI IMPORTE

Michel Seymour

Professeur honoraire

Département de philosophie

Université de Montréal

Nous allons dans cet article prendre comme point de départ les expériences de pensée proposées par Derek Parfit au sujet de l'identité personnelle, telles que discutées dans *Reasons and Persons*¹. Nous ferons alors quelques remarques visant à rendre cohérents les résultats apparemment contradictoires de ces expériences. Cette réflexion permettra ensuite d'expliquer la sorte de réductionnisme que Parfit cherche à défendre. Nous serons ainsi en meilleure posture pour apprécier la cohérence des changements opérés au sein des différents textes tardifs dans lesquels notre auteur discute de l'identité personnelle.

1. Les expériences de pensée initiales

Parfit commence par proposer une expérience de pensée qui s'appuie sur la possibilité logique d'une téléportation. Dans le premier cas, un scanner copie exactement tous les faits de mon corps et de mon cerveau, l'état de toutes mes cellules, puis détruit ensuite mon corps et mon cerveau (199). L'information qui est ainsi obtenue est envoyée à la vitesse de la lumière en quelques minutes jusqu'à un réplicateur situé sur la planète Mars qui va se charger de reconstituer mon corps et mon cerveau dans ses moindres détails à partir du scan effectué, mais en se servant d'autres échantillons d'un matériau biologique du même type. D'un point de vue physiologique, on a affaire à une sorte de clonage tandis que, du point de vue psychologique, la continuité est parfaite. Selon Parfit, j'ai donc été reconstitué par le réplicateur. On doit conclure que c'est moi qui retrouve conscience quelques instants plus tard. Telle est la première expérience de pensée. Elle assure du moins en apparence le maintien de mon identité. Et puisque ma psychologie est la même dans un corps nouveau du même type, le maintien de mon identité semble être assuré par la continuité psychologique.

Avant de considérer la seconde expérience de pensée, il convient de faire un certain nombre de remarques. Parfit semble présupposer qu'il existe une relation de covariance entre les états du cerveau et les états psychologiques de l'agent. Il semble présupposer

¹Parfit D., *Reasons and Persons*, Oxford, Oxford University Press, 1982. Voir en particulier les chapitres 10 (« What We Believe Ourselves to Be ») et 11 (« How We Are not What We Believe »). Les références à cet ouvrage ne feront intervenir que les numéros de page.

que si le cerveau d'un agent est reconstitué avec les mêmes états physiques, l'agent est alors également reconstitué dans ses états psychologiques. On ne voit pas en tout cas pourquoi la reconstitution à l'identique de mon corps et de mon cerveau, effectuée par le réplicateur, à partir d'un nouveau matériau biologique devrait déterminer un ensemble d'états mentaux qui assurent la continuité psychologique. Même si rien n'interdit la possibilité que mes états psychologiques soient réalisés de plusieurs façons différentes, les états de mon cerveau semblent déterminer pour leur part des états psychologiques spécifiques. Il semble donc que Parfit présuppose une relation de survenance entre les états psychologiques et les états physiques de l'agent. En effet, une fois que le corps et le cerveau sont répliqués, il faut que les états mentaux soient les mêmes pour qu'il y ait continuité psychologique. Si cela est assuré par une reconfiguration à l'identique des états physiques de l'agent il faut alors présupposer qu'il existe une relation de survenance (dépendance) du psychique sur le physique.

Quoi qu'il en soit, si survenance il y a, il semble qu'il s'agit seulement d'une relation de survenance faible, c'est-à-dire intramondaine. Les états du cerveau peuvent peut-être en principe déterminer des états psychologiques différents dans différents mondes possibles, mais ici on présuppose seulement de manière contingente que l'agent a tel ensemble d'états psychologiques dès lors que son cerveau se trouve dans tel ensemble d'états physiques (neurophysiologiques). Si l'expérience de pensée de Parfit permettait d'anticiper par l'intuition une réalité susceptible d'être empiriquement fondée, il faudrait donc à tout le moins être en mesure de valider éventuellement l'existence d'une corrélation statistique entre les propriétés physiques et les propriétés psychologiques d'un agent. On pourrait alors confirmer qu'en répliquant intégralement ce que je suis sur le plan physique, on reproduit sur Mars une personne ayant les mêmes états psychologiques que moi.

Un amendement important doit toutefois être apporté pour s'accorder en principe avec ceux qui, comme Hilary Putnam ou Tyler Burge², admettraient l'anti-individualisme en philosophie de l'esprit. Selon ce point de vue, les contenus d'états mentaux sont déterminés en partie en fonction des règles linguistiques conventionnelles adoptées par la communauté. On peut faire varier le contenu des états mentaux en fonction de changements dans la signification linguistique conventionnellement admise des expressions servant à articuler le contenu de ces états mentaux, et on peut le faire sans que les états du cerveau changent. Ce serait une façon de rejeter la survenance métaphysique des états mentaux

²Putnam, H., «The Meaning of 'Meaning'», in *Language, mind, and knowledge*. Minnesota studies in the philosophy of science, Volume 7, 1975, p. 131-193; Burge, T. "Individualism and the Mental", in P.A. French (Ed.), *Midwest Studies in Philosophy, Vol.4, Studies in Metaphysics*, 1979, 73-121.

sur les états physiques. Pour préserver la survenance et rester neutre au sujet de l'anti-individualisme, il faut proposer l'amendement suivant. Les états mentaux ne préservent leur identité que si les états physiques du cerveau sont préservés ET que le contexte sociolinguistique reste le même. L'amendement proposé permet alors de neutraliser les objections visant à démontrer que la reproduction intégrale des états neurophysiologiques d'un agent n'est pas en soi une garantie de continuité psychologique. Pour répondre à cette objection, il suffit d'admettre une survenance faible et de préciser que l'environnement sociolinguistique est le même sur la planète Mars. On peut dire métaphoriquement que la téléportation est une forme de voyage rapide. En fait, personne ne peut voyager à la vitesse de la lumière. Peu importe la vitesse à laquelle on se déplace, notre vitesse doit toujours être inférieure à 300 000 kilomètres à l'heure. Il est difficile de prétendre que je me suis « déplacé » d'un endroit à un autre à la vitesse de la lumière, et ce, même si le message radio a été envoyé à la vitesse de la lumière. Quoi qu'il en soit, la question suivante se pose : suis-je la même personne une fois que le scanner m'a reconstitué intégralement sur Mars ? Dans cette situation, je me retrouve dans les mêmes types d'états physiques et psychologiques, même si mon corps et mon cerveau particuliers, bien que du même type, ne sont plus les mêmes à cause du matériau nouveau utilisé. Est-ce encore moi ? Parfit estime que notre intuition nous invite à répondre par l'affirmative. Il s'agit ici d'une téléportation simple.

Parfit admet à l'époque de *Reasons and Persons* la possibilité logique et physique de la téléportation d'un corps. Il s'agirait seulement d'une impossibilité technologique actuelle. L'humanité ne sera peut-être jamais capable de développer une technologie lui permettant de réaliser ce genre de choses, mais ce n'est pas physiquement impossible. Seuls des facteurs parfaitement contingents l'empêcheront de le faire. Parfit admet que je puisse en principe réapparaître reconstitué sur Mars par un réplicateur après que l'on m'ait scanné et que l'on m'ait détruit. Pour arriver à ce résultat, il n'est pas nécessaire de supposer que je voyage à la vitesse de la lumière. Pour éviter cette conséquence, il faut renoncer à l'idée que mon identité se conforme au principe de la continuité spatio-temporelle. Dans la téléportation, mon identité subit un sort semblable à celui d'une montre que l'on aurait tout d'abord défaït en de multiples pièces, pour la reconstituer ensuite et être en mesure de dire qu'il s'agit de la même montre.

Pour Parfit, ce serait toutefois se méprendre sur la pertinence des expériences de pensée que de s'employer à les critiquer en s'appuyant sur une thèse d'impossibilité physique de réaliser une téléportation complète du corps humain. Il est vrai que dans l'état actuel de la science on peut tout au plus admettre la téléportation quantique de particules subatomiques à 1200 kilomètres de distance. Même s'il s'avérait que la téléportation du corps

humain entier est impossible, la situation fictive qu'elle met en scène permettrait de tester nos intuitions profondes concernant ce qui est essentiel dans l'identité personnelle. Les expériences de pensée servent à faire ressortir nos intuitions concernant l'individuation des personnes dans notre monde. Tel est l'usage qu'on peut en faire. S'agissant de l'expérience de pensée postulant une téléportation sur la planète Mars, elle permet de mettre en évidence le fait que l'essentiel se trouve du côté de la continuité psychologique.

Considérons toutefois une seconde expérience de pensée. Supposons que, plusieurs années plus tard, le procédé de téléportation s'est raffiné puisque le scanner ne doit plus nécessairement détruire le cerveau et le corps de l'individu sur Terre. On peut le reproduire grâce à un réplicateur sans détruire sur le champ l'original. Je pourrais donc quelques minutes plus tard échanger avec ma réplique sur un écran qui transmet son image à la vitesse de la lumière à partir de Mars. Je reçois quelques minutes plus tard l'image d'une personne qui semble m'être en tout point identique, dans la mesure où elle est en continuité psychologique et elle possède tous mes traits physiques. C'est le cas de la téléportation complexe (*branch line case*) (200-201)

. Dans le cas de la téléportation complexe, ma vie et celle de ma réplique se superposent l'espace de quelques instants. Toutefois, selon Parfit, notre intuition serait à l'effet de dire que ma réplique n'est pas moi. Voilà qui donne l'impression que la théorie initiale, celle portant sur la continuité psychologique, est somme toute fausse. Autrement, il faudrait se demander pourquoi les arguments jugés valables dans le cas de la téléportation simple seraient soudainement falsifiés lorsque la téléportation est complexe ? Si l'essentiel de l'identité repose sur la continuité psychologique, ne faut-il pas conclure que je suis identique à cette réplique, à tout le moins au moment où ma réplique vient tout juste d'apparaître pleinement reconstituée sur la planète Mars ? Ne sommes-nous pas deux exemplaires d'une seule et même personne ? Étant donné que, dans une situation de téléportation complexe, notre intuition est que ma réplique n'est pas moi, il semble qu'il faille renoncer au critère d'identité proposé dans la téléportation simple.

Pour bien comprendre ce qui se passe, il faut saisir le fait que Parfit ne cherche pas au fond à formuler un bon critère d'identité personnelle. Il cherche plutôt à nous convaincre du fait que l'identité personnelle n'est pas ce qui importe. Il faut plutôt comprendre que, selon Parfit, la continuité psychologique importe plus que l'identité personnelle. Telle est la leçon essentielle à retenir, y compris au sujet de la téléportation simple. Si dans la téléportation simple Parfit donne à penser que je suis ma réplique, c'est d'abord et avant tout pour montrer que la continuité psychologique est ce qui importe.

Si la téléportation simple sert à mettre en évidence l'importance de la continuité psychologique, la téléportation complexe sert à souligner le fait que la même situation peut

se produire sans que l'identité personnelle ne soit maintenue. Au lieu de voir dans la seconde expérience de pensée une position qui contredit bêtement une thèse portant sur l'identité personnelle, la leçon essentielle est de mettre l'accent sur la continuité psychologique pour indiquer que c'est cela qui importe. Le fait que dans la téléportation complexe notre intuition nous incite à dire que ma réplique n'est pas moi ne doit donc pas être vu seulement comme une simple réfutation de la thèse initiale, mais bien comme une confirmation que l'essentiel de la thèse initiale n'est pas le maintien de l'identité personnelle, mais bien le maintien de la continuité psychologique. Une préoccupation à l'égard de ce qui importe vraiment se trouve au cœur de la réflexion parfitienne sur l'identité personnelle.

Supposons que, dans le cas d'une téléportation complexe, le scanner provoque des effets secondaires conduisant à la mort quelques jours plus tard. Même si, dans ce cas, je ne suis pas ma réplique, la vie de ma réplique est ce qui se rapproche le plus de la survie ordinaire de ma personne.

On peut certes interroger avec Kathleen Wilkes³ et d'autres la pertinence des avancées philosophiques qui exploitent notre imagination et qui prennent la forme d'expériences de pensée plus ou moins fantaisistes, mais la controverse soulevée par Parfit se trouve ailleurs, précisément dans l'articulation entre ce que nous sommes et ce qui importe dans ce que nous sommes. Après tout, ceux qui doutent de la téléportation peuvent se rabattre sur le clonage qui est une potentialité technologique bel et bien réelle. Nous verrons plus loin comment il est possible de concevoir une nouvelle expérience de pensée en s'appuyant sur le clonage ou sur des jumeaux identiques.

2. L'identité personnelle : en quoi consiste-t-elle ?

Pour interpréter correctement la position de Parfit, il faut tout d'abord l'appréhender à partir de trois réponses à des questions essentielles et incontournables. Il faut en premier lieu se demander en quoi consiste une personne, c'est-à-dire, quelles en sont les composantes essentielles. Ensuite, il faut se prononcer sur la question de l'unicité des personnes. Qu'est-ce qui explique le caractère unique de chaque personne ? Quel est notre critère d'identité numérique ? Enfin, il y a la question de savoir ce qui assure le maintien de l'identité à travers le temps. C'est ici que Parfit développe un critère de continuité psychologique (202).

³Wilkes K., *Real People: Personal Identity without Thought Experiments*, Oxford, Clarendon Press, 1988. Je reviens plus loin sur le caractère problématique des expériences de pensée en examinant les critiques de Sydney Shoemaker.

La première question soulève le problème de l'identité qualitative de la personne. Pour Parfit, la réponse est : un corps, un cerveau et des évènements psychiques et physiques interreliés. Plus exactement, il faut dire que c'est *l'existence* d'une personne qui se compose de ces différents éléments.

Avec la deuxième question, on demande à quelles conditions *x* est identique à *y*, et c'est au fond le critère d'identité numérique qui est ici en cause. On pourrait être tenté de dire en première approximation que pour Parfit l'ensemble de ses qualités spécifiques est ce qui rend la personne unique. En l'occurrence, c'est au niveau de ses propriétés psychologiques que la personne est unique. Quiconque partagerait la même configuration psychologique que moi serait moi. Toutefois, ainsi qu'on l'a déjà indirectement démontré avec la seconde expérience de pensée, Parfit rejette l'identité des indiscernables. Même s'il est réductionniste, il ne va pas jusqu'à affirmer que si un individu *m*'est indiscernable, alors il est moi, et ce, même si je suis encore en vie. Tout être humain vivant et conscient fait l'expérience de son unicité. C'est ce qui explique la réticence à admettre que ma réplique sur Mars est Michel Seymour. Ma réplique sur Mars ou mon clone dans la vraie vie ne me sont pas identiques même s'ils me sont indiscernables. Comme nous le verrons, Parfit sera plus tard amené à dire que même si les personnes n'existent pas comme séparées de leurs propriétés, la distinction entre une personne et l'ensemble de ses propriétés est un élément essentiel de notre schème conceptuel.

La troisième question pose le problème de la persistance. Quand on sait ce qui rend compte de l'identité, on sait aussi ce qui fait que la personne reste la même à travers le temps. Ce sera pour Parfit la continuité psychologique. Plus tard, ce critère sera amendé pour inclure aussi le maintien d'une partie suffisante du cerveau.

Il y a plusieurs critères d'identité à travers le temps. Pour certains, le critère de la continuité spatio-temporelle est le critère déterminant. Pour parler d'une seule et même personne, il faut qu'un espace-temps continu ait été occupé. Ce critère peut être satisfait par un organisme qui est tout d'abord un œuf, puis une chenille, puis une chrysalide et enfin un papillon. Le critère ne fonctionne toutefois pas tout le temps. S'il fallait absolument admettre le critère de la continuité spatio-temporelle pour tous les objets, il faudrait alors conclure qu'une montre qui serait défaite en pièces et qui serait ensuite reconstituée ne serait pas la même montre. Notre intuition va pourtant dans le sens contraire. Il semble donc que le critère de la continuité spatio-temporelle ne s'applique pas aux montres. La question se pose cependant de savoir si ce critère est une condition nécessaire de l'identité des personnes à travers le temps. Si l'on répondait par l'affirmative, il faudrait rejeter l'idée que ma réplique sur Mars est bel et bien moi, y compris dans le cas d'une téléportation simple, car pendant les quelques minutes que dure le processus de téléportation, je

suis en quelque sorte aux abonnés absents. Nous reviendrons plus tard sur le thème de la continuité spatio-temporelle pour les personnes.

Un autre critère serait que la personne n'est rien d'autre qu'un cerveau et un corps particuliers. Pour que la personne reste la même, il faudrait que son cerveau et son corps soient préservés. Il s'agirait ici une identité entre la personne et les choses qui la constituent. Le critère supposerait que la personne doit avoir nécessairement le même cerveau et le même corps pour rester la même. Ainsi que nous allons le voir, Parfit récuse le réductionnisme de l'identité.

Considérons maintenant le critère physique qui est une version sophistiquée du critère précédent (203-204). Pour que la personne soit la même, il faut que subsistent suffisamment des composantes du cerveau et du corps de la même personne. Une personne X au temps T1 est la même personne que Y au Temps T2 si et seulement si le cerveau et le corps de X ont suffisamment subsisté et peuvent être maintenant localisés chez Y et personne d'autre n'existe qui ait aussi suffisamment du cerveau de X. L'identité personnelle dans le temps consisterait dans ces faits physiques⁴. Parfit a aussi récusé ce critère à l'époque de *Reasons and Persons*. Il en retiendra plus tard certains aspects dans la version amendée de son critère d'identité. Les tenants de ce troisième critère rejettent l'idée selon laquelle le résultat de la téléportation est la même personne.

Un autre critère, ayant trait à la dimension psychologique, pourrait être mobilisé. Il s'agirait de postuler une substance spirituelle, ce qui aurait pour conséquence de traiter la personne comme séparée de l'ensemble de ses propriétés. Parfit rejette cette option cartésienne. On peut même dire que son ouvrage a d'abord et avant tout comme objectif de nous libérer de cette façon de penser.

On peut ensuite mentionner aussi un critère psychologique, celui qui est associé à John Locke⁵. Le critère de Locke est fondé sur la continuité psychologique, mais il repose essentiellement sur la mémoire. C'est le critère en vertu duquel il est supposé que l'unité de la personne est garantie par la présence d'une mémoire directe de l'ensemble de nos expériences. Le critère de Locke est très certainement faux, car je suis la même personne que j'étais à 3 ans, mais j'ai très peu de souvenirs directs de mes expériences d'alors.

Le critère de Parfit lui-même peut maintenant être introduit (205-209). La personne apparaît en tant que phase (*phase sortal*) dans la carrière d'un être humain. En tant que personnes, nous ne sommes pas des êtres humains⁶. La personne apparaît après que l'être

⁴Thomson J. J., « People and their bodies », in J. Dancy (dir), *Reading Parfit*, Oxford, Basil Blackwell, 1997, p. 202-229.

⁵Locke J., *An Essay Concerning Human Understanding*, New York, Dover 1959, Livre 2, Chapitre 27, section 16.

⁶Parfit, D. « We are not Human Beings », *Philosophy*, vol. 87, no.1, 2012, p. 5-28.

humain soit né et il arrive souvent qu'elle cesse d'exister même si l'être humain n'est pas encore cliniquement mort. Qu'est-ce qui ferait qu'à partir d'un certain temps, la personne apparaît dans la vie d'un être humain ? Comme on l'a vu, la continuité psychologique est l'élément central avancé par Parfit. Rappelons que l'intérêt véritable de Parfit pour la continuité psychologique est à mettre en relation avec ce qui à ses yeux importe vraiment. Toutefois, au départ, il le présente comme ce qui se trouve au centre de l'identité personnelle.

Le critère de Parfit suppose l'existence d'une vie mentale consciente dont la continuité est assurée par un éventail de facteurs psychologiques qui vont bien au-delà de la mémoire. Supposons quand même tout d'abord l'existence de connexions mémorielles directes entre des ensembles d'états mentaux. Il peut y avoir une continuité de la mémoire pourvu qu'il existe une chaîne d'ensembles d'états mentaux entre lesquels subsistent des connexions mémorielles directes. L'ensemble des événements et états mentaux présents à un moment donné forme une classe d'éléments pouvant être majoritairement en connexion mémorielle directe avec une seconde classe d'éléments. Si à son tour cette seconde classe entretient un même type de relation avec une troisième classe et que ce genre de connexion remonte dans le temps jusqu'à ma tendre enfance, on se fait alors une assez bonne idée de la continuité psychologique mémorielle. Il doit cependant y avoir d'autres faits que des connexions mémorielles qui lient les événements mentaux entre eux. Ici Parfit s'éloigne de la tradition inaugurée par Locke tout en restant à la recherche d'un critère fondé sur la continuité psychologique. Selon le critère psychologique, les amnésiques peuvent en principe être la même personne pourvu que la continuité psychologique soit préservée sous d'autres formes. Cela est possible pour Parfit parce que la continuité psychologique ne dépend pas seulement des connexions mémorielles. Elle dépend également du maintien des désirs, intentions, croyances, valeurs, finalités et projets. Si l'amnésique perd non seulement contact avec le passé, mais n'est même plus mû par les mêmes croyances, les mêmes désirs, les mêmes intentions et les mêmes projets, il ne peut plus être la même personne.

On peut parler de connexions psychologiques pour décrire l'ensemble des connexions pertinentes. Et on peut parler de continuité psychologique lorsqu'il existe des chaînes superposées de connexité (*connectedness*) entre des ensembles formés de sensations, d'états, d'actes et d'actions. Pour qu'il y ait une connexité entre les événements psychiques au temps t et les événements psychiques au temps t' , il faut qu'un agent ait des connexions directes (semblables à celles qui existent chez un agent normal) entre les événements psychiques qu'il entretient en t et ceux qu'il entretient en t' . Lorsqu'il y a suffisamment de connexions directes, la connexité est robuste (*strong connectedness*).

Les évènements mentaux M peuvent être en connexion robuste avec des événements mentaux M' et les M' en connexion robuste avec des événements mentaux M'' sans que les M soient en connexion robuste avec les M''. La connexion n'est donc pas une relation transitive. Elle n'est donc pas à elle seule un critère d'identité personnelle parce que cette dernière est une relation transitive. Si A est la même personne que B et B la même personne que C, alors A est C. Je suis identique à celui que j'ai été il y a vingt ans, mais je n'ai pas actuellement de connexion robuste avec ce que j'ai été à cette époque. Toutefois, lorsqu'il y a une superposition d'ensembles d'états psychologiques en séquences et que tous les couples d'ensembles sont liés entre eux par une connexion robuste, il y a une continuité psychologique. La continuité psychologique ainsi définie est ce qui, en première analyse, assure l'identité d'une personne à travers le temps. La connexion directe peut prendre différentes formes. Des relations peuvent exister entre les croyances, puis entre des croyances et des intentions, des intentions et des désirs, entre des actions et des intentions, des actions et des souvenirs, des actions et des plans, etc. Ces différents types de connexion ponctuent notre vie mentale.

On vient de voir que le critère psychologique d'identité personnelle repose sur la continuité psychologique qui est comprise comme supposant des chaînes superposées de connexion robuste entre des ensembles d'états mentaux. Un critère prenant la forme de conditions nécessaires et suffisantes semble alors pouvoir être formulé : X est la même personne que Y ssi (i) X est dans une relation de continuité psychologique avec Y (la téléportation simple répond à ce critère), (ii) cette continuité est assurée par la bonne sorte de cause (cela reste à préciser) et (iii) il n'existe pas deux personnes différentes qui sont dans une continuité psychologique avec Y (207) (Ce dernier scénario serait celui de la téléportation complexe qui, comme on l'a vu, prendrait autrement en défaut le critère proposé).

La condition (ii) fait intervenir le concept de « bonne sorte de cause ». La bonne sorte de cause peut être la cause normale, habituelle. Dans ce cas, les événements mentaux que j'ai sont dans une continuité qui en même temps est produite (causée, déterminée) par mon cerveau. Mais on peut admettre aussi des causes fiables ou n'importe quelles causes garantissant aussi la même continuité.

Prenons le cas de la mémoire pour illustrer le critère qui ferait intervenir un concept de bonne sorte de cause normale. Je crois me rappeler avoir eu une expérience. J'ai de fait eu cette expérience et ma mémoire apparente est de façon normale causalement dépendante de cette expérience passée. Mais parfois on se rappelle un événement que l'on a vécu parce que l'on nous l'a raconté. En de telles circonstances, on se rappelle d'un événement qui est effectivement survenu mais que l'on ne se remémore pas de façon normale.

Voilà un premier exemple de cause anormale ayant produit un état mental de souvenir. Pour Parfit, de tels exemples peuvent malgré tout être considérés comme de bonnes sortes de cause. Il s'agit de sources fiables. Un autre exemple peut être signalé pour illustrer des causes anormales qui sont malgré tout de bonnes sortes de causes. Si je vois grâce à un œil artificiel, je vois quand même, et ce, bien que la cause de l'expérience ne soit pas une cause normale. Se rappeler de mon expérience passée suite à un cas de téléportation serait alors un autre cas de cause anormale. La cause de la continuité psychologique entre moi et ma réplique n'est pas la cause normale mais c'est une cause acceptable.

Il y a les causes normales et les causes anormales mais fiables. Puis il y a n'importe quelles autres causes. Parfit estime pouvoir se contenter de n'importe quelles causes pour exposer son critère d'identité (208).

Pour que la personne reste la même, il faut qu'il y ait continuité psychologique. Ce n'est peut-être pas une condition suffisante, mais il semble bien que cela soit pour Parfit une condition nécessaire. Mais puisque les connexions qui assurent cette continuité ne sont pas des relations transitives, la question se pose : une seule et même personne pourrait-elle vivre des changements importants dans ses croyances, valeurs et plans tout en restant la même ? Le maintien d'un même ensemble de croyances, valeurs et plans ne serait donc pas essentiel au maintien de l'identité personnelle ? À l'époque de *Reasons and Persons*, alors que n'importe quelles causes sont admises, il admet une version supposant une large portée au critère psychologique de continuité (*Wide Psychological Criterion*). Cela s'oppose au critère étroit commandant toujours une cause normale impliquant donc le cerveau (*Narrow Psychological Criterion*). À cette époque, la continuité psychologique repose entièrement sur le maintien d'une masse critique des mêmes croyances, désirs, intentions, buts, projets et souvenirs. Un être humain amnésique pourrait alors être la même personne pourvu que l'essentiel de ces autres traits psychologiques demeure. Il faut alors dire que, pour Parfit, lorsque des changements draconiens surviennent et que la majorité de nos traits psychologiques se transforment, on a affaire à un seul et même être humain, mais qui se transforme en diverses personnes (327-329).

Ce qui assure l'identité à travers le temps, est-ce le fait que les changements soient progressifs ? Il semble bien que oui. Certes, rien ne semble interdire au départ un changement relativement rapide et radical, comme une conversion religieuse, par exemple. Il est clair cependant que si cela se produit, le même être humain devient une autre personne.

Ceci dit, on ne peut, selon Parfit, parvenir à formuler de manière tranchée un critère d'identité de la personne, et donc à préciser une fois pour toutes ce à quoi elle est identique. Même si Parfit cherche à formuler un critère d'identité personnelle qui met d'abord et avant tout en scène la continuité psychologique, il admet pour n'importe quel critère

donné qu'il n'y a pas de réponse tranchée par oui ou par non à la question : est-ce la même personne ? Même dans le cas du critère psychologique, il y a de l'indétermination parce que la connexité suppose qu'il y ait *suffisamment* de connexions directes. Le terme « suffisamment » a une application floue dans le cas qui nous occupe. Cette indétermination généralisée fait en sorte de ne pas toujours être capable de répondre par oui ou par non à la question de savoir si l'individu est le même ou non. Pour nous convaincre de cela, Parfit va construire des spectres psychologique (231-233), physique (234-236) et combiné (à la fois psychologique et physique) (236-237) dans lesquels les propriétés psychiques et/ou physiques d'un individu sont modifiées au point où il devient méconnaissable ou alors finit par ressembler à une autre personne. Ainsi, par exemple, un être humain peut être amené à se prendre pour Napoléon. À aucun moment devons-nous conclure qu'il est Napoléon, mais la question se pose de savoir s'il est encore la même personne. On peut être tenté de répondre que c'est la même personne, affligée d'un désordre psychologique, mais jusqu'où peut-on aller ? De la même manière, un individu peut voir plusieurs de ses organes transplantés et plusieurs de ses membres amputés et on peut se demander si c'est encore la même personne. On peut être tenté de répondre par l'affirmative et ce bien qu'il soit physiquement très transformé, mais là encore, la question se pose de savoir jusqu'où on peut aller. Enfin, même si l'on adopte un critère fondé sur la continuité psychologique et physique, Parfit imagine une situation où à un bout du spectre combiné, il y aurait Derek Parfit tandis qu'à l'autre bout, suite à des modifications successives autant physiques que psychiques, il y aurait Greta Garbo. Entre les deux, il y aurait une zone grise indéterminée.

L'indétermination ne fait pas que prendre en défaut nos critères d'identité à travers le temps. Elle affecte également notre identité personnelle proprement dite. Le spectre psychologique prend en défaut le critère psychologique, mais il ne faut pas le faire en présupposant qu'il doit bien y avoir une personne déterminée que le critère ne parvient pas à cerner. De la même manière, on peut grâce au spectre physique prendre en défaut le critère physique, mais il ne faut pas le faire en présupposant qu'il doit bien y avoir une personne déterminée que le critère ne parvient pas à cerner. L'objectif de Parfit est de montrer que l'identité personnelle est indéterminée. Malgré l'intuition que nous avons que l'essentiel de ce que nous sommes est préservé dans la téléportation simple, nos intuitions vont dans le sens contraire si je suis maintenu en vie et confronté à ma réplique quelques minutes plus tard. Ce résultat contradictoire vise à ébranler nos convictions profondément ancrées au sujet de notre identité personnelle. L'indétermination n'est pas que sémantique et ne se résume pas à identifier un problème dans notre emploi du mot 'personne'. Il s'agit d'une indétermination ontologique.

Pour certains, l'important est le fait que les états psychologiques aient un propriétaire et c'est la présence d'un même propriétaire qui explique l'unité de l'expérience. Dans ce cas, l'identité personnelle est ce qui importe et la personne est une entité séparée qui est de surcroît, une entité déterminée. Si au contraire, on s'engage sur le terrain du réductionnisme, les expériences de pensée qu'il propose servent à montrer que l'identité personne est indéterminée. Parfit se sert de cette conclusion pour montrer que ce qui importe, c'est l'existence d'une relation R, c'est-à-dire d'ensembles superposés de connexions psychologiques directes suffisamment nombreuses assurées par la bonne sorte de causes et qui conduisent à la continuité psychologique (215).

Une dernière remarque s'impose avant d'aller plus loin. Puisque le critère favori de Parfit est psychologique, on pourrait être tenté de prétendre que Parfit souscrit à une certaine variante du dualisme. Sans admettre une substance spirituelle séparée, il semble admettre une psychologie séparée (ce qui s'appelle en fait un dualisme des propriétés) et semble donc être dualiste. Toutefois, Parfit rejette explicitement cette interprétation. Il écrit :

« Les physicalistes pourraient accepter le critère psychologique. Et ils pourraient accepter la version qui permet n'importe quelle cause fiable, ou n'importe quelle cause. Ils pourraient donc croire que, dans la téléportation simple, ma réplique serait moi. Ils seraient ici alors en train de rejeter le critère physique. »⁷ (209)

La téléportation des états psychologiques dans un autre cerveau est compatible avec l'hypothèse que les états psychologiques soient réductibles à des états physiques du cerveau. On doit seulement supposer que les mêmes états du cerveau sont reproduits dans un autre cerveau.

3. Le réductionnisme constitutif de Parfit

Le rejet du réductionnisme en matière d'identité personnelle suppose que la personne est une entité séparée de l'ensemble de ses propriétés. La réponse réductionniste classique consiste à prendre le contre-pied de cette position et à affirmer que la personne n'est rien d'autre que l'ensemble de ses propriétés. Il faut non seulement affirmer l'indiscernabilité des identiques, mais également l'identité des indiscernables. Il faut non seulement ad-

⁷« Physicalists could accept the Psychological Criterion. And they could accept the version that allows any reliable cause, or any cause. They could thus believe that, in Simple Teleportation, my Replica would be me. They would here be rejecting the Physical Criterion. »

mettre que si A et B sont identiques, alors ils ont les mêmes propriétés. Il faut aussi admettre que si leurs propriétés sont identiques, alors A et B sont également identiques. Si l'identité des personnes repose sur le partage de qualités particulières plus fondamentales, alors l'identité d'une personne dépend de ses qualités particulières, ce qui va dans le sens d'une réduction de la personne à un ensemble de propriétés. Or, Parfit récuse ce second critère lorsqu'il est appliqué à l'identité personnelle, car placé en situation de téléportation complexe, il admet que ma réplique, bien qu'indiscernable, ne m'est pas identique. Son réductionnisme n'est donc pas classique.

Le critère de Parfit est réductionniste pour deux raisons :

- a. L'identité des personnes à travers le temps consiste en l'existence de faits qui sont plus particuliers (*particular facts*).
- b. Ces faits peuvent être décrits, sans avoir à exprimer explicitement l'existence des personnes, ou sans présupposer l'existence des personnes. Ces faits peuvent être décrits de façon impersonnelle (210).

Ceux qui, comme Bernard Williams et Judith Jarvis Thomson, mettent en avant le critère physique défendent eux aussi une variante du réductionnisme. Ces deux positions (affirmant respectivement le critère physique et le critère psychologique) s'accordent sur l'essentiel (le réductionnisme), mais diffèrent dans des cas comme ceux de la téléportation.

L'existence d'une personne consiste dans un cerveau et un corps particuliers ainsi que dans une série d'événements mentaux et physiques interreliés. À partir de ce qui existe, on peut certes inférer que la personne a un cerveau et un corps particulier (213). Mais puisque la description impersonnelle de ce qui existe est exhaustive et que la proposition inférée n'ajoute aucun fait nouveau, le réductionnisme est vrai.

À l'époque où il écrit *Reasons and Persons*, Parfit soutient que l'unité de la conscience et de la vie entière peut être expliquée en faisant état seulement des relations qui existent entre les expériences et un cerveau sans jamais faire référence à des personnes. Il pense que l'on peut faire une description factuelle parfaitement impersonnelle de la réalité et que celle-ci est exhaustive et n'omet aucun fait. Selon Parfit, on peut décrire ces expériences et les relations qu'elles entretiennent entre elles sans faire intervenir le fait qu'elles soient possédées par une personne. Les faits pertinents peuvent donc être décrits de manière impersonnelle. Même si, en un certain sens, on peut affirmer que les personnes existent, on peut décrire complètement cette existence sans avoir à affirmer qu'elles existent. L'affirmation qu'elles existent n'ajoute rien comme fait supplémentaire.

Il peut arriver que l'on fasse référence à un individu quand on décrit les faits psychologiques impliqués dans la continuité psychologique, mais cela ne fait aucunement référence à un sujet séparé qui posséderait des états mentaux (210). Certes, plusieurs anti-réductionnistes croient que les personnes existent comme des entités séparées. Ils peuvent croire cela parce qu'ils souscrivent au dualisme et acceptent un ego cartésien. Ils pourraient aussi souscrire à cela parce qu'ils admettent l'existence d'une substance physique séparée (210). D'autres auteurs croient que les personnes sont en partie un cerveau et un corps ainsi que des événements mentaux et physiques interreliés. Mais ils croient qu'il faut ajouter un fait supplémentaire, l'existence de la *personne*. Nous verrons plus loin que telle est la position de Sydney Shoemaker. Ce dernier endosse un réductionnisme semblable à Parfit, mais il estime que les personnes sont identiques à des « systèmes » constitués de cerveaux, de corps, ainsi que d'événements mentaux et physiques interreliés.

Comme Shoemaker, Parfit soutient tout d'abord que les personnes ne sont pas des entités séparées de leurs propriétés. Elles ne sont pas des êtres ayant une existence indépendante et qui ont par ailleurs de façon accidentelle la propriété de posséder un corps, un cerveau et des croyances, désirs et intentions interreliés. En tant qu'entités existantes, elles ne sont *constituées* de rien d'autre qu'un cerveau, un corps, ainsi que des événements mentaux et physiques interreliés. Pour Parfit, l'existence des personnes ne fait intervenir (*involves*, 216) rien d'autre que cela. Leur existence *consiste* (211) dans le fait qu'il existe des cerveaux et des corps, ainsi que des événements mentaux et physiques interreliés. En tant qu'entités séparées, les personnes n'existent pas. Jusque là, son point de vue s'accorde parfaitement avec celui de Shoemaker.

Le point de vue du sens commun s'appuie sur une intuition pré-philosophique différente. Nous croyons en l'existence des personnes comme entités séparées. Nous pensons que nous *avons* un cerveau, un corps, des croyances, désirs et intentions et que, pour cette raison, nous sommes distincts de toutes ces choses. Nous croyons que notre existence n'est pas réductible à ces choses que nous avons. Nous avons tendance à nous fier à la grammaire qui nous permet de dire que Derek *a* un corps, un cerveau, des croyances, des désirs et des intentions. Il semble que le référent de 'Derek' doive être différent de toutes ces choses qu'on lui attribue. Il semble que l'unité de la conscience et d'une vie entière requiert qu'on les explique en invoquant justement la personne comme fait séparée. Peu de philosophes souscrivent explicitement à l'idée d'une substance mentale séparée d'un

corps physique, c'est-à-dire au dualisme des substances mentale et physique de René Descartes⁸, mais nous avons tout de même tendance tout naturellement à penser les choses en ce sens. Parfit veut tenir compte du sens commun et de ce que nous avons tendance à dire, mais il veut aussi désamorcer les fausses hypothèses métaphysiques qui peuvent en découler et c'est à ce dernier niveau que se situe son réductionnisme. Il ne s'agit pas d'un réductionnisme de l'identité, car un tel type de réductionnisme nous forcerait à dire que la personne est identique à ce corps, ce cerveau et ces évènements mentaux et physiques qui la constituent. Or, Parfit récuse explicitement l'idée que la personne soit identique à ce qui la constitue.

Il faut ensuite distinguer deux choses. Il y a l'idée que la personne puisse être une entité séparée. C'est une idée que Parfit rejette, mais il y a aussi le fait que nous soyons enclins à dire que Derek a un corps, des croyances, des désirs, des intentions, etc. Parfit estime que nous sommes en droit de dire les choses de cette façon et ce, même si ce n'est peut-être au fond qu'« une façon de parler » (223, 226).

Pourquoi ne pas dire que, dans les faits, la personne est identique à ce qui la constitue ? Si la seule chose qui nous fait résister au réductionnisme de l'identité est la prise en compte d'une façon de parler, alors on ne voit pas ce qui gêne Parfit à adopter la thèse de l'identité entre ce qu'est une personne en réalité et ce qui la constitue. Le problème est toutefois que, selon lui, les personnes ont certes une part d'existence réelle, et cette existence ne consiste en rien d'autre que ce qui a été indiqué plus haut, mais elles semblent aussi avoir en partie une part d'existence fictionnelle. Il y a tout d'abord le fait que, pour Parfit, nos énoncés psychologiques à la première personne ne sont, comme on vient de le remarquer, que des façons de parler (223, 226). Il y a aussi la comparaison qu'il établit entre la personne et ce que Hume a à dire au sujet de la république ou de la nation (211-212). Chez Hume, il en va de la nation et de la république comme de la causalité. On projette sur le réel des faits qui n'existent pas vraiment. Il y a ensuite le fait que l'indétermination n'affecte pas que le mot 'personne'. Ce n'est pas qu'une indétermination sémantique. C'est notre identité personnelle elle-même qui est indéterminée (213). Se demander si X est la même personne peut dans certains cas être une question vide (213). Et surtout, comme nous allons le voir, tous les critères d'identité personnelle semblent être confrontés à des contre-exemples. Aussi, même si l'existence réelle des personnes ne consiste en rien d'autre que la présence d'un cerveau et d'évènements physiques et mentaux interreliés, une personne n'est pas identique à ce qui la constitue parce qu'elle ne se réduit

⁸Descartes R., *Méditations Métaphysiques*, Jean-Marie Beyssade, Michelle Beyssade (dir), Paris, Garnier Flammarion, 2011.

pas à sa part d'existence réelle. Elle est aussi en partie fictionnelle. Sans être une pure fiction, elle ne se ramène pas non plus à ce qui la constitue dans sa matérialité. Pourquoi parler de réductionnisme alors ? Tout simplement, parce qu'il s'agit de réduire sa part d'existence réelle à n'être rien d'autre que les composantes qui la constituent. Et surtout, puisque l'autre dimension de la personne n'est que fictionnelle, une description exhaustive de la réalité peut être impersonnelle. C'est l'autre élément essentiel à la base de son réductionnisme, tel qu'annoncé au début de cette section.

Parfit croit que :

- c. l'existence d'une personne ne consiste en rien de plus que l'existence d'un cerveau et d'un corps ainsi que des événements mentaux et physiques interreliés.

Il rejette l'idée que :

- d. la personne n'est rien d'autre qu'un cerveau particulier et qu'un corps particulier ainsi que des événements mentaux et physiques interreliés.

Il soutient également qu'en un sens :

e. la personne est une entité distincte d'un corps et d'un cerveau. Même si son existence n'est constituée par rien de plus que l'existence d'un cerveau et d'un corps et de l'occurrence d'événements mentaux et physiques, on peut dire qu'en un sens elle est distincte de ceux-ci, sans pour autant être obligé de prétendre qu'elle existe comme une entité séparée ou qu'elle est un fait supplémentaire attaché à un corps et un cerveau. La raison semble être qu'elle comporte une part virtuelle, abstraite et fictionnelle (211).

Parfit reconnaît que (e) semble être contradictoire avec (c). L'interprétation que nous en faisons rend cependant parfaitement cohérents les énoncés (c-e). Si c'est la part fictionnelle des personnes qui nous autorise à affirmer (e), ce dernier énoncé est compatible avec ce qui est affirmé en (c), car il est alors question de l'existence réelle des personnes. La théorie réductionniste à laquelle Parfit adhère est celle qui combine (c) et (e) (212).

Parfit compare les personnes avec les peuples ou les nations. L'existence d'une nation implique seulement celle de ses citoyens entretenant des relations diverses entre eux sur un territoire donné. Cela ressemble à (c). On rejette toutefois l'identification de la nation à cette population et ce territoire. La population se reproduit dans le temps et est progressivement remplacée par une autre génération. Elle pourrait aussi être déportée sur un autre territoire. Si on affirmait qu'elle est identique à des personnes entretenant des relations diverses sur un territoire, cela ressemblerait à (d). Mais on rejette l'idée que la nation ne soit rien de plus que celle des citoyens et d'un territoire. Elle est une entité qui est distincte de composantes telles que des citoyens et un territoire. Dire que la nation est

telle et telle, c'est dire plus que simplement affirmer que des citoyens existent sur un territoire. Cela rappelle (e) (211-212).

Comment peut-on expliquer la cohérence de (c) et (e) ? Comment la personne peut-elle être distinguée de son cerveau et de son corps, ainsi que des événements psychiques et/ou physiques interreliés, tout en admettant que son existence ne consiste en rien de plus ? La seule réponse semble être que l'élément résiduel est fictionnel. Si la personne est en partie une fiction en plus d'être un corps et un cerveau, alors elle se distingue de ces composantes tout en n'étant rien d'autre sur le plan ontologique existentiel que l'ensemble de ces composantes. Parfit n'emploie jamais l'adjectif « fictionnel » pour rendre compte de la composante résiduelle qui distingue la personne de son existence ontologiquement réelle, mais il est difficile de caractériser autrement cet élément résiduel.

Il est amené à affirmer (e) parce que dans le cas de la téléportation complexe, on a affaire à une réplique qui est un corps, un cerveau et qui est en continuité psychologique avec moi, mais qui n'est pas moi. Notre intuition nous pousse à dire que si j'étais placé dans cette situation, j'affirmerais que ma réplique n'est pas moi. Mon autoreprésentation résisterait à l'affirmation que ma réplique est moi. La personne est donc aussi en partie le résultat d'une autoreprésentation. Si la représentation que je me fais de moi-même me pousse à nier que ma réplique m'est identique, ce n'est pas à cause de mon rejet de la thèse métaphysique de la bilocalisation. Après tout, la Honda Civic et le Camembert existent en plusieurs exemplaires différents et existent simultanément en différents endroits. Si on refuse cette idée pour les personnes, c'est à cause du lien étroit qui existe entre le concept de personne et celui d'ipséité, c'est-à-dire la conscience de mon unicité. Le rejet de la bilocalisation pour les personnes s'explique par le fait que les personnes n'existent pas indépendamment de la conscience qu'elles se font d'elles-mêmes. Or, puisque cette autoreprésentation n'est rien d'autre au fond qu'une façon de parler, le sujet apparaît dans le contenu de cette auto-représentation comme une personne-fiction. Parfit semble reconnaître cela tout en précisant au fond que cette seule propension à réaliser une auto-attribution devrait nous inciter à reconnaître que l'identité personnelle n'est pas ce qu'il y a d'important. L'important est ce qui existe réellement, à savoir la continuité psychologique. La survie de ma réplique après ma mort est ce qui se rapproche le plus de ma propre survie.

Bien que l'existence d'une personne consiste dans un corps, un cerveau, et des croyances, désirs et intentions, l'énoncé que la personne a telle ou telle propriété n'est pas identique à l'énoncé qu'un corps, un cerveau, et des croyances, désirs et intentions interreliés ont telle ou telle propriété. Ces deux énoncés ne disent pas la même chose. Notre

concept de personne nous pousse à refuser de reconnaître comme étant la même personne une réplique qui serait pourtant en parfaite continuité psychologique avec nous. La raison est que la notion de personne va de pair avec un mécanisme d'auto-attribution permettant de mieux cibler ce que l'on est.

En même temps, puisque l'élément résiduel ayant permis d'identifier cette différence entre moi et ma réplique n'est pas une personne séparée ni le fait supplémentaire d'une personne attachée à ses propriétés, on peut dire que ce qui importe vraiment est la continuité psychologique. Si je mourais dans dix minutes après la téléportation complexe, ma réplique assurerait pour ainsi dire l'essentiel de ma survie. Cela prouve que l'auto-attribution du statut de personne n'est pas ce qui importe. La question serait certes très différente si je subsistais pendant une trentaine d'années après la téléportation, mais la raison en est que ma réplique serait alors très différente de moi et ne serait plus en continuité psychologique avec moi.

La version du réductionnisme à laquelle Parfit souscrit est cette version qui tient compte de notre concept actuel de personne (212) ; d'où le réductionnisme constitutiviste. La personne est constituée d'un corps, d'un cerveau, et des croyances, désirs et intentions interreliés sans leur être identique. Cette position est semblable à celle de Shoemaker, tout en différant sur un point essentiel. Alors que Shoemaker considère la personne comme un fait additionnel, Parfit considère que la référence à des personnes n'est qu'une façon de parler.

Admettre l'indétermination ontologique, c'est reconnaître qu'en un certain sens, les personnes n'existent pas. Si on n'a pas tout à fait tort d'affirmer qu'elles existent, c'est parce que ce qui les constitue existe (la thèse (c) mentionnée plus haut). Mais la personne est à distinguer de ce qui la constitue (la thèse (e) mentionnée plus haut). Si les constituants des personnes existent bel et bien, les personnes elles-mêmes n'existent pas. Si une description totale de la réalité peut être impersonnelle, c'est parce que les personnes n'existent pas. Qu'il y ait une dimension fictionnelle dans notre concept de personne est ce qui permettrait de le démontrer. Au cœur de la réflexion de Parfit, il y a cette idée que l'identité personnelle n'est pas ce qui importe. Si cela importe peu, c'est parce qu'au fond les personnes n'existent pas.

4. La théorie reformulée

Nous avons jusqu'ici montré la cohérence des expériences de pensée proposées par Parfit et explicité quel sens il fallait donner à sa posture réductionniste. Il faut maintenant s'attarder aux changements majeurs survenus au sein de sa pensée. Ils affectent sa position réductionniste et même le statut et l'importance de ses expériences de pensée. Avant de

les considérer et pour mieux les comprendre, il convient de s'attarder à certaines objections formulées à l'endroit de sa théorie. Déjà à l'époque de *Reasons and Persons*, Parfit considérait le vieil argument du Bishop Butler selon lequel la personne ne peut être réduite à des expériences, parce que ces expériences presupposent elles-mêmes l'existence des personnes (219-223). Même si Butler a écrit à une autre époque, la position qu'il défend revient à soutenir que l'on ne peut affirmer la continuité psychologique sans presupposer l'identité personnelle. S'il a raison, Parfit a tort de prétendre être en mesure de décrire la continuité psychologique de façon impersonnelle.

Butler affirme que la conscience présuppose cette identité. L'idée est que l'on ne peut se rappeler que de ses *propres* expériences. Comment cela affecte-t-il l'argument de Parfit ? On se rappellera que ce dernier reconnaît l'existence de connexions robustes entre les événements mentaux d'un agent. Parmi celles-ci, il y a le fait de se rappeler de ses expériences antérieures. Or, comment une telle mémoire s'articule-t-elle ? *Je me rappelle* être sorti hier, d'avoir pensé à telle ou telle chose, d'avoir voulu utiliser mon parapluie, et d'avoir finalement oublié de l'apporter ? Dans ces souvenirs successifs, une référence à la première personne apparaît. Il s'agit d'une présupposition qui fait intervenir ce que la base de réduction cherche à éviter de faire intervenir. Autrement dit, si l'objection de Butler pouvait être transposée dans le contexte de la discussion de la théorie de Parfit, cela voudrait dire que ce dernier se trompe en prétendant être capable de faire une description impersonnelle fondée sur la continuité psychologique.

Pour répondre à cette objection, Parfit a tout d'abord soutenu que l'on peut en principe se rappeler les états mentaux et les expériences d'un autre. Dans la mémoire que nous avons de nos expériences, il faut donc distinguer ce qui impliquerait une référence à soi et ce qui constitue l'expérience en tant que telle. Dans les faits, nous n'avons pas des expériences de mémoire des expériences d'autrui. Mais il est techniquement envisageable que des neuro-chirurgiens soient un jour en mesure d'implanter des expériences de mémoire de X dans le cerveau de Y. On décrit habituellement avec raison les expériences dont on se rappelle comme nos expériences. Mais puisqu'il est au moins en principe possible d'imaginer le transfert d'un souvenir dans la tête d'une autre personne, celle-ci serait dans ce cas en mesure de se rappeler ce qu'une autre a vécu. Cela veut dire qu'il n'est pas métaphysiquement nécessaire de décrire nos souvenirs comme des souvenirs portant sur nos expériences. Sans s'objecter à des expressions comme « se rappeler ses propres expériences », Parfit propose de construire une notion, celle de la quasi-mémoire. Qu'est-ce que la quasi-mémoire ? Il y a trois composantes à mentionner :

- Je crois qu'une expérience a eu lieu.
- Quelqu'un a eu cette expérience.

-Ma mémoire est causalement dépendante (et de la bonne manière) de cette expérience (220).

Dès lors que je crois qu'une expérience a bel et bien eu lieu, que quelqu'un a eu cette expérience et que ma mémoire est causalement dépendante de cette expérience, alors j'ai une expérience de quasi-mémoire. Ici, on a expulsé du contenu de la mémoire toute présence d'éléments faisant référence à ma personne. Selon Parfit, mes propres souvenirs forment une instance particulière d'expériences de quasi-mémoire. Il s'agit du cas particulier dans lequel ce quelqu'un n'est nul autre que moi-même. Le lien avec l'existence de ma personne est toutefois contingent. La même expérience aurait pu être celle d'un autre.

Selon Parfit, la mémoire présuppose l'identité personnelle seulement parce que nous faisons usage de notre concept de mémoire. Notre concept de mémoire fait intervenir en plus de la quasi-mémoire une croyance à l'effet que ces souvenirs sont nos propres expériences. Toutefois, il s'agit d'une croyance séparée :

« Puisque nous n'avons pas de quasi-souvenirs au sujet des expériences passées des autres, nos souvenirs apparents ne font pas que se présenter à la première personne. Ils sont accompagnés d'une croyance supposant que s'ils ne sont pas illusoires, ils sont au sujet de nos propres expériences. Toutefois, en ce qui a trait aux souvenirs d'expériences, il s'agit d'une croyance détachable.⁹ » (222)

Le concept de mémoire que Parfit propose renvoie dans la réalité à rien d'autre que de la quasi-mémoire. Nous projetons en plus une auto-représentation de nous-mêmes sur des faits de quasi-mémoire.

Le fait que notre mémoire présuppose l'identité est-il un indice que quelque chose ne tourne pas rond dans la théorie de Parfit ? Sa réponse est que le discours de l'auto-attribution fait intervenir tout au plus notre concept de personne, telle que celui-ci s'articule dans une auto-représentation qui n'est qu'une façon de parler, et cela ne veut donc pas dire que l'on présuppose une personne réelle. Les expériences de mémoire présupposent en effet la présence d'une personne, mais il s'agit d'une personne-fiction, projetée sous la forme d'une croyance autoreprésentationnelle portant sur des expériences de quasi-mémoire. Si elles ne font pas intervenir une référence à une personne réelle, ce n'est pas seulement parce qu'il est en principe possible d'avoir des expériences de quasi-mémoire.

⁹« Because we do not have quasi-memories of other people's past experiences, our apparent memories do not merely come to us in the first-person mode. They come with a belief that, unless they are delusions, they are about our own experiences. But, in the case of experience-memories, this is a separable belief. »

C'est d'abord et avant tout parce que la croyance réflexive est détachable et prend la forme d'une projection sur une réalité qui, elle, ne fait intervenir que de la quasi-mémoire.

L'identité personnelle réelle reposerait en partie sur des chaînes superposées d'ensembles de connexions directes robustes de quasi-mémoire. Deux ensembles A et B de souvenirs entraînent une connexité robuste si les connexions directes de quasi-mémoire entre A et B sont en nombre suffisant. Il faut pour cela qu'au moins la moitié des expériences de quasi-mémoire de B soient connectées à des expériences vécues en A. Des chaînes d'ensembles de connexions directes assurent la continuité de la quasi-mémoire. Ce concept de quasi-mémoire ne présupposerait donc pas l'identité personnelle. Ou plus exactement, une expérience mémorielle inclurait la quasi-mémoire ainsi qu'une auto-représentation de soi comme sujet-fiction. La quasi-mémoire serait alors la seule chose qui existe réellement dans toute expérience de mémoire.

La question est cependant loin d'être résolue. Dans une recension de *Reasons and Persons*, Sydney Shoemaker, qui défend lui aussi un critère d'identité néo-lockéen, prend parti en faveur d'une position semblable à celle que nous venons de décrire au sujet du Bishop Butler¹⁰. Selon ce point de vue, la personne consiste dans une continuité psychologique fondée sur la superposition d'ensembles de connexions directes, mais on doit faire intervenir des expériences conçues comme des « *experiencings* », c'est-à-dire des objets qui requièrent la présence d'un propriétaire. Les expériences sont des « *experiencings* » qui ont un caractère adjectival supposant donc un terme qu'elles modifient, et donc un sujet d'expérience (Shoemaker, 447).

Il affirme sa préférence pour une conception qui explique l'unité de nos expériences en invoquant un seul et même propriétaire, à savoir les personnes conçues comme étant attachées à ces expériences. Il s'agit d'une conception qui rejette l'idée des personnes comme entités séparées, mais qui suppose néanmoins que l'existence des personnes est un fait supplémentaire, ajouté aux faits concernant ses éléments constitutifs. Cette personne semble être le système comprenant comme éléments le cerveau, le corps, et des états physiques. Les éléments constituant la personne sont organisés en un système fonctionnel. Tel est le fait supplémentaire qu'est la personne. Cette dernière est le système constitué d'un corps, d'un cerveau et d'évènements mentaux et ou physiques interreliés. Nous sommes toujours en présence d'une position réductionniste qui est constitutiviste, mais la personne est considérée comme identique à ce système fonctionnel :

¹⁰Shoemaker S., «Critical Notice of *Reasons and Persons*», *Mind*, 94: 1985, 443-453; reproduit sous le titre « Parfit on Identity », dans Jonathan Dancy (dir), *Reading Parfit*, Oxford, Basil Blackwell, 1997, p.135-148.

« Si une sorte d'explication causale ou fonctionnelle du mental est correcte, ce qui constitue un évènement ou un état mental donné comme étant d'un type mental particulier (par exemple une expérience ou une croyance ayant un certain contenu) est le lien qu'il entretient avec un système plus large dans lequel il peut jouer un certain «rôle causal» fonctionnel. Et l'existence d'un tel système sera exactement ce que Parfit considère comme constituant l'existence d'une personne¹¹. » (Shoemaker, 446)

La personne, comprise comme système, est le fait supplémentaire attaché aux constituants. De cette façon, Shoemaker peut rendre compte des conditions objectives de vérité des énoncés psychologiques à la première personne. Ces énoncés ne sont pas seulement 'projetés' sur le réel, car ils décrivent des états de fait objectifs. Ce ne sont pas seulement des façons de parler. Toutefois, ces conditions de vérité ne nous contraignent pas à admettre des personnes séparées ou une notion de personne qui soit conceptuellement ou ontologiquement irréductible. Les énoncés de la forme « Je pense (crois, désire, veux, etc.) ceci ou cela » affirment l'existence d'un système comprenant tel ou tel état psychologique de pensée (croyance, désir, volonté, etc.).

À cela, Parfit peut bien entendu répondre que la réduction ontologique des personnes, désormais considérées comme identiques à des systèmes, ne change pas la donne, car elle engendre les mêmes résultats dans une situation de téléportation complexe. Dans une telle situation, on aurait affaire à deux systèmes indiscernables réalisés dans des cerveaux, des corps et des évènements mentaux et physiques indiscernables, et pourtant nous ne dirions pas que ce sont deux exemplaires de la même personne. La position de Shoemaker va à l'encontre des intuitions profondes qui nous sont révélées par la seconde expérience de pensée.

Cependant, comme nous allons le voir, Shoemaker va formuler une seconde critique qui vise justement à neutraliser l'impact de ces expériences de pensée. Il récuse un traitement de l'identité personnelle fondé sur ce qui est logiquement possible et qui semble être parfois impliqué dans ce genre d'expérience de pensée. Il entend pour sa part se garder de discuter de ce qui est logiquement possible ou impossible. Son réductionnisme est un réductionnisme faible, s'appuyant tout au plus sur la nécessité métaphysique, et non

¹¹« If any sort of causal or functional account of the mental is correct, what constitutes a given mental state or event as being of a particular mental kind (for example an experience or belief having a certain content) is its being so related to a larger system to which it belongs as to be apt to play a certain 'causal role' in the workings of that system, and the existence of such a system will be just what Parfit regards as constituting the existence of a person. »

un réductionnisme s'appuyant sur la nécessité logique. Les expériences de pensée de Parfit nous placent souvent en face de situations purement imaginaires et, lorsqu'elles sont interprétées de cette façon, elles n'ont guère d'utilité. Pire, elles privent d'une vision plus juste et plus scientifique de l'identité personnelle.

Avant d'examiner plus avant ce second problème, on peut examiner la réponse que Parfit donne à la première critique. Dans un article daté de 1999, Parfit offre une réponse nuancée à l'argumentaire déployé par Shoemaker¹². Tout en continuant de défendre le réductionnisme constitutiviste, il renonce tout d'abord à l'une des deux thèses qui caractérisaient à ses yeux ensemble ce type de réductionnisme, à savoir la possibilité de produire une description impersonnelle qui serait en même temps exhaustive de la réalité. Il soutient maintenant que la référence à des personnes fait partie de notre schème conceptuel. Toute description du monde devra inclure des énoncés faisant référence à des personnes et affirmant leur existence. Cela donne en partie raison à Shoemaker. Dans cette nouvelle version, les deux auteurs admettent le réductionnisme constitutiviste. Les deux admettent que nous sommes distincts de nos cerveaux et de nos corps, sans toutefois être des entités séparées (Parfit 1999, 218) Contrairement à Shoemaker cependant, il ne s'agit pas d'une existence réelle. Parfit n'accepte pas de réduire les personnes à des systèmes fonctionnels. On se ferait une meilleure idée de la position adoptée si on la comparait à celle qui est proposée par Daniel C Dennett. Ce dernier compare les personnes à des centres de gravité en physique¹³. Il s'agit d'abstractions au sens d'entités fictionnelles qu'il est cependant très utile de postuler, ne serait-ce que pour être en mesure d'attribuer une responsabilité morale aux agents qui performent certaines actions et posent certains gestes. Les personnes sont en quelque sorte les centres de gravité de la trame narrative que les organismes humains tissent. Or, les êtres humains tissent tout au long de leurs vies plusieurs trames narratives distinctes. Dans l'exemple de l'Aristocrate russe discuté par Parfit (327), on est incité à admettre qu'un être humain est susceptible de modifier tout au long de sa vie ses croyances, valeurs et finalités. Il faut admettre un parcours dans lequel défilent en lui plusieurs personnes. Or, la présence de plusieurs personnes au sein d'un seul être humain contribue à accroître la conviction que celles-ci sont en fait de pures fictions.

Parfit ne renonce donc pas à son idée que les personnes distinctes de ce qui les constituent ne sont que des « façons de parler » (223), mais il admet que le concept de personne

¹² Parfit D., « Experiences, Subjects, and Conceptual Schemes », *Philosophical Topics*, 26, 1999, p. 217-70.

¹³ Dennett D., « Why everyone is a novelist », *Times Literary Supplement*, 4459 (Sept.16-22, 1988): 1016, 1028-29; reproduit sous le titre « The Self as the Center of Narrative Gravity », dans F. Kessel, P. Cole, and D. Johnson (dir), *Self and Consciousness : Multiple Perspectives*. Hillsdale, NJ, Erlbaum, 1992.

fait irrémédiablement partie de notre schème conceptuel. Désormais, les énoncés « Les personnes existent », « Les personnes ont des cerveaux », « Les personnes ont des corps », etc. doivent figurer dans notre description du monde. Nous ne pouvons donc plus faire une description impersonnelle de la réalité. Il soutient toutefois que cela dépend de notre schème conceptuel et qu'il existe un monde possible dans lequel la personne ne fait pas partie de notre schème conceptuel. En somme, même si Parfit accuse le coup et retient quelque chose des critiques formulées par Shoemaker, sa réponse ne saurait satisfaire ce dernier. Parfit admet certes qu'une description exhaustive de la réalité doit désormais faire référence à des personnes, ce qui l'oblige à renoncer à son type initial de réductionnisme, mais leur statut d'entités partiellement fictionnelles n'est pas remis en cause. À l'opposé de ce point de vue, Shoemaker semble réduire la personne à un système fonctionnel. Pour ce dernier aussi, la personne est distincte de ce qui la constitue. Elle n'est pas identique à un ensemble particulier d'états du cerveau, mais elle est identique aux réalisations multiples d'un même ensemble fonctionnel.

Le différend initial entre les deux philosophes ne s'arrête pas là. Venons-en donc au deuxième problème évoqué plus haut. Shoemaker récuse l'utilité d'expériences de pensée purement imaginaires alors que Parfit en fait l'élément central de sa démonstration. Les expériences de pensée ont pour but de faire échouer tous les critères d'identité proposés, ce qui permet ensuite à Parfit de montrer que l'identité personnelle n'a pas d'importance. Pour Shoemaker, les expériences de pensée initiales de Parfit, interprétées comme de simples explorations de l'espace logique, n'ont rien à voir avec ce qui est prescrit par le réalisme scientifique.

L'auteur de *Reasons and Persons* est jusqu'à un certain point sensible à cette critique. Pour bien comprendre sa nouvelle position, il convient de prendre connaissance du nouveau critère d'identité qu'il propose. Il retient toujours la continuité psychologique, mais il ajoute que, pour être la même personne, il faut aussi préserver au moins un hémisphère du cerveau. L'identité personnelle est assurée par la continuité psychologique et la continuité d'une portion significative du cerveau¹⁴. C'est le critère psychologique ayant sa base dans le cerveau (*brain based psychological criterion*) discuté dans la section 3 de Parfit (1995). Cela correspond aussi au critère psychologique étroit (*Narrow Psychological criterion*) de *Reasons and Persons* qui requiert le maintien de suffisamment de cerveau en plus de la continuité psychologique (208).

Ce nouveau critère entraîne des changements d'importance. Tout d'abord, il ne permet plus l'expérience de pensée de la téléportation qui supposait une discontinuité spatio-

¹⁴« The Unimportance of Identity », dans Henry Harris (dir), *Identity*, Oxford, University Press, 1995.

temporelle de l'identité. Désormais, puisque l'agent doit disposer d'une partie importante de son cerveau pour continuer d'exister, il semble que l'identité personnelle doive aux yeux de Parfit aller de pair avec la continuité spatio-temporelle. En vertu de ce nouveau critère, une personne continue d'exister si et seulement si (a) une partie suffisante du cerveau continue d'exister et continue d'être le cerveau d'un être vivant, (b) s'il y a une continuité psychologique, (c) si cette continuité est assurée par une cause normale (le cerveau) et (d) il n'existe pas une personne différente qui est en continuité psychologique avec elle (208).

Un autre changement important doit être remarqué. Puisque les seules causes acceptables sont les causes normales, la continuité repose sur les liens causaux survenant au sein du cerveau et non plus nécessairement sur le maintien des mêmes croyances, valeurs, désirs, intentions et finalités. Le maintien d'un même ensemble de croyances, valeurs et plans n'est donc pas essentiel au maintien de l'identité personnelle. En vertu du nouveau critère adopté qui coïncide parfaitement avec le critère de continuité psychologique étroit discuté dans *Reasons and Persons*, rien ne semble interdire au départ un changement radical, comme une conversion religieuse, par exemple. (207) Les connexions psychologiques peuvent dans ce cas exister, même si notre ensemble ancien de croyances, valeurs et finalités nous semble désormais insatisfaisant. Puisque la conversion consiste à vouloir s'en distancier, on peut dire que l'ancien ensemble joue un certain rôle causal dans l'apparition du nouvel ensemble. Dans le cas d'une conversion radicale, en effet, si la mémoire perdure et que l'ensemble ancien de croyances, valeurs et finalités sert de repoussoir à l'individu par rapport à un ensemble nouveau, il y a bel et bien des connexions directes entre les ensembles. À l'opposé, chez l'amnésique radical pour qui le problème aurait eu un impact négatif sur les autres types de connexions directes, chez le schizophrène ainsi que chez les personnes souffrant du syndrome de personnalités multiples, les connexions directes disparaissent presque totalement.

On vient de voir que Parfit a fait quelques pas dans la direction des positions adoptées par Shoemaker. Il admet que la référence à des personnes fait partie de notre schème conceptuel. Il propose un critère d'identité révisé qui renoue avec la continuité spatio-temporelle et qui fait reposer la continuité sur les causes normales. À partir de cette version révisée, il est cependant aussi possible d'imaginer une reformulation de ses expériences de pensée qui s'éloigne sensiblement de la pure science-fiction. Au lieu d'envisager les mondes possibles et les alternatives disponibles dans l'espace logique, Parfit considère des éventualités scientifiquement envisageables et donc, des alternatives épistémiques qui sont davantage ancrées dans la pratique de la science réelle. Toutefois, ces

expériences de pensée vont permettre à Parfit de revenir à la charge et de défendre à nouveau l'indétermination de l'identité personnelle.

Selon Shoemaker, il est incohérent de prétendre qu'il y a dans le monde deux sortes d'entités : des choses dont l'identité est indéterminée et des choses dont l'identité est déterminée. (1451) L'indétermination de l'identité est peut-être seulement un phénomène sémantique et non métaphysique. (1450) Il n'y aurait donc pas d'indétermination ontologique de l'identité. Pour créer des zones grises comme à mi-chemin dans le spectre combiné, il faut admettre des changements progressifs dans les états mentaux des agents, ce qui suppose une conception atomiste ou moléculariste des états mentaux. Or, Shoemaker endosse une conception holiste qui n'autorise pas des changements progressifs venant modifier par couches successives les ensembles d'états mentaux. Un système compris de façon holiste relève du tout ou rien (*all or nothing*). S'agissant de la matière grise, il n'y a pas, selon lui, de zone grise.

Qu'à cela ne tienne, Parfit est en mesure de provoquer le même genre de remise en question en faveur de l'indétermination de l'identité. Et qui plus est, il peut y parvenir avec des expériences de pensée ne relevant plus de la pure science-fiction. Considérons le clone d'un être humain ou un jumeau identique. Le clonage est une possibilité physique bien plus réelle que la téléportation, mais il n'est même pas nécessaire de mobiliser cette idée. On peut se contenter d'imaginer le cas de jumeaux identiques ayant tous les deux été victimes d'un accident. Supposons donc que l'un des deux jumeaux (nommons le A) soit devenu paraplégique et que, l'autre jumeau (nommons le B) subit une mort cérébrale, même si ses organes vitaux sont maintenus en vie. Il faut ensuite postuler la possibilité d'exécuter une transplantation d'au moins un hémisphère du cerveau de A dans la boîte crânienne de B après que l'on ait extrait le cerveau de B. Ce genre d'opération n'est pas encore réalisable mais il n'est pas impossible de l'envisager comme éventuellement possible. Après tout, le neurochirurgien italien Sergio Canavero a déjà considéré la possibilité de réaliser une telle opération en Chine en 2017 sur Valeri Spiridonov, un individu russe qui était paraplégique¹⁵.

Supposons donc qu'une opération de ce genre puisse être réalisée. J'évacue ici pour les fins de la discussion les enjeux éthiques qu'une telle opération soulève. On suppose que la continuité psychologique de A sera assurée, parce qu'une bonne partie du cerveau de A sera maintenue et que la transplantation a lieu dans le corps d'un clone ou d'un jumeau identique. Pour Parfit, la personne qui se réveille au terme de l'opération est bel et bien A.

¹⁵http://www.lepoint.fr/editos-du-point/anne-jeanblanc/la-premiere-greffe-totale-de-tete-pour-bientot-23-11-2017-2174616_57.php. Consulté le 28 juin 2018.

Aux yeux de ceux qui défendent un critère de continuité physique, A serait mort, puisque l'essentiel de son corps est mis de côté, mais pour le défenseur du critère psychologique amendé, il est clair que A est bel et bien vivant dans ce nouveau corps. On peut d'autant plus l'affirmer que, contrairement à la téléportation, la continuité spatio-temporelle est de cette manière elle aussi assurée. Ce nouveau critère s'appuie sur la continuité psychologique, ainsi que sur le maintien d'au moins la moitié du même cerveau, au sein d'un corps génétiquement identique et respectant la continuité spatio-temporelle. Si à l'époque, le cerveau reconstitué à partir d'un nouveau matériau pouvait constituer une bonne sorte de cause, les bonnes causes ne sont cette fois-ci rien d'autre que celles qui sont engendrées par le même cerveau.

Considérons maintenant la possibilité que l'on produise deux clones de A, à savoir B et C, ou que A ait B et C comme jumeaux identiques. Imaginons ensuite que lors de l'incident, B et C subissent tous les deux une mort cérébrale. Imaginons enfin que l'un des deux hémisphères de A soit transplanté dans la boîte crânienne de B et que l'autre hémisphère soit transplanté dans la boîte crânienne de C. Comment devons-nous évaluer la situation ? Est-ce que, suite à la double opération, A se réveille dans le corps de B ou dans le corps de C ? Il semble problématique d'affirmer que A se réveille seulement dans l'un des deux corps, mais il semble tout autant impossible de supposer que A survit dans les deux corps et que les organismes humains vivants soient deux exemplaires de A. Manifestement, chacun dirait de l'autre qu'il n'est pas lui. Si on faisait subir à B une épreuve physiquement douloureuse, C ne ressentirait rien et vice-versa.

Pour certains, cette seconde expérience de pensée, qui met en jeu une double transplantation, annule ni plus ou moins la plausibilité du critère de continuité psychologique dans sa version nouvelle et il faut alors se mettre à la recherche d'un meilleur critère d'identité personnelle. Au contraire, pour Parfit, la double transplantation montre que l'identité personnelle n'a pas d'importance.

On comprendra ainsi comment Parfit réagit à la critique de Shoemaker au sujet de la valeur des expériences de pensée imaginaires qu'il avait initialement conçues. Les nouvelles expériences de pensée permettent de montrer que la réflexion métaphysique initiale de Parfit était valable. Dans la nouvelle version proposée, les hypothèses postulées sont beaucoup moins extravagantes, mais elles vont dans le même sens que les expériences initiales. Parfit ne s'appuie plus sur des mondes possibles relevant de la pure imagination. Il ne mobilise même pas des mondes possibles physiquement accessibles, c'est-à-dire sur des mondes possibles très différents du nôtre mais qui obéissent aux mêmes lois physiques. Il prend appui seulement sur des alternatives épistémiques cliniquement

envisageables. Toutefois, puisque les nouvelles expériences corroborent les conclusions des anciennes expériences, elles permettent du même coup d'en justifier la pertinence.

5. Conclusion : la personne comme artefact culturel

Parfit s'intéresse au critère d'identité personnelle parce qu'il veut en arriver à montrer que la continuité psychologique est ce qui importe. Ce qui compte vraiment dans l'identité personnelle, c'est la continuité psychologique et non l'ipséité, à savoir l'expérience de son unicité. Il ne faut pas perdre de vue que, pour Parfit, s'il y a ce qui assure l'identité au travers du temps, il y a aussi ce qui importe dans l'identité. S'il s'est intéressé à l'identité personnelle, c'est d'abord et avant tout pour identifier ce qui à ses yeux est important. Or, ce qui est important au travers du temps peut ne pas concorder avec ce qui assure la continuité de l'identité personnelle. S'agissant de déterminer ce qui importe le plus, il se pourrait bien que cela soit le maintien des mêmes croyances, valeurs et finalités. On sait très bien que dans la téléportation complexe ou dans la transplantation dans des crânes différents de nos deux hémisphères, l'identité personnelle ne coïncide plus avec la continuité psychologique. Parfit n'en conclut pas qu'il faille renoncer à l'importance accordée à la continuité psychologique, bien au contraire. Il faut plutôt accorder peu d'importance à l'identité personnelle. Si à un certain niveau, l'identité personnelle importe du point de vue de l'éthique, de l'altruisme et de la mort, à un niveau plus profond encore, la chose la plus importante est la poursuite de certaines croyances, valeurs et finalités qui peuvent survivre à notre propre mort.

Parfit ne veut pas seulement épingle ce qui importe dans l'identité des personnes. Il veut également montrer que l'identité des personnes est indéterminée. Il y a des cas où on ne peut répondre par oui ou par non à la question de savoir si tel individu est la même personne. Il y a même des cas où l'on ne peut répondre par oui ou par non à la question de savoir si telle ou telle personne est sur le point de mourir. Ces questions sont parfois vides de sens. Admettre l'indétermination, c'est reconnaître que l'on est parfois dans l'impossibilité de répondre par oui ou par non à des questions comme « est-ce la même personne ? ». Même lorsqu'il y a une continuité psychologique, que le cerveau de la personne est en partie maintenu en plus d'être dans un corps génétiquement identique et qu'il y a une continuité spatio-temporelle, l'identité peut être indéterminée.

Que devons-nous penser de cette théorie ? Ainsi que Paul Ricœur l'a souligné, Parfit a peut-être une compréhension imparfaite de l'ipséité qui est, comme on vient de le souligner, l'expérience de notre unicité¹⁶. La faiblesse, principale ombre au tableau, concerne le caractère fictionnel de l'identité ainsi comprise et qui n'est selon lui qu'une façon de parler. Parfit aurait dû s'interroger davantage sur cette « façon de parler » qui est la nôtre lorsqu'il est question de l'usage que nous faisons des énoncés psychologiques à la première personne, surtout lorsqu'il s'agit de décrire nos états intentionnels. Ces énoncés ne servent pas seulement à rapporter la directionalité de certains états (le fait qu'ils renvoient à des objets ou états de chose qui peuvent ne pas exister). Ils n'ont pas non plus pour seuls traits caractéristiques le fait d'échouer aux différents tests d'extensionnalité (substitution des identiques, généralisation existentielle, vérifonctionnalité). Ils traduisent aussi une capacité réflexive qui est constitutive de toute expérience intentionnelle complète. Et surtout, ils servent à exprimer l'autorité de la première personne. En effet, si je crois intentionnellement que p, alors je sais que je crois que p. Du moins est-ce ainsi qu'il faut comprendre une intentionnalité qui existe peut-être à divers degrés dans la réalité mais qui comporte la directionalité, l'intensionnalité, la réflexivité et l'autorité de la première personne lorsque nous sommes parvenus aux degrés ultimes d'une expérience intentionnelle pleine et entière.

D'où origine cette autorité de la première personne ? Doit-on invoquer un accès épistémique privilégié ? Une autre explication serait la suivante. Quand je crois, pense, juge, désire intentionnellement que p, j'ai une propension à *statuer* que je crois, pense, juge et désire que p. En tant qu'états intentionnels, la croyance, la pensée, le jugement ou le désir que p sont des états qui sont pour ainsi dire sous mon contrôle. Pour le dire plus précisément, les états intentionnels vont de pair avec des dispositions à la performance d'actes illocutoires déclaratifs à l'occasion desquels (i) j'exprime ma décision de croire, penser, juger ou désirer que p, (ii) je présuppose une autorité particulière que j'ai à cet égard et (iii) j'implicite que ces états sont rendus vrais par le fait de mon énonciation et ce, à cause de cette autorité que j'ai. Plusieurs états de croyance, pensée, jugement et désir sont inconscients, mais certains sont sous notre contrôle. La classe des états mentaux relevant du volontarisme doxastique n'est pas vide.

On a beaucoup trop négligé le fait que les énoncés psychologiques à la première personne pouvaient servir à la performance de différents types d'actes de langage. On peut

¹⁶ Ricœur P., *Soi-même comme un autre*, Paris, Seuil, 1990. Voir la Cinquième étude, L'identité personnelle et l'identité narrative, et la Sixième étude, Le soi et l'identité narrative.

s'en servir pour performer des actes illocutoires assertifs, mais on peut s'en servir également pour performer des actes illocutoires expressifs et surtout, même parfois, des actes illocutoires déclaratifs. Or, les actes illocutoires déclaratifs rendent vrais des états de fait. Ils instituent certains faits institutionnels.

On peut trouver inspiration auprès de Daniel C Dennett et de ses remarques naturalistes sur l'origine des personnes pour faire la généalogie de cette propension à déclarer tel ou tel état de fait. On peut remonter loin dans le temps pour en retracer les origines¹⁷. Dennett s'interroge au fond sur le processus conduisant à l'affirmation de son identité. En l'occurrence, cela semble s'inscrire dans le prolongement de l'évolution des espèces qui construisent leur propre niche écologique et tracent de cette manière une frontière entre l'extérieur et l'intérieur. La toile d'araignée, la termitière, la fourmilière, le barrage construit par des castors en sont des exemples probants. L'institutionnalisation d'une identité narrative prenant la forme d'une autobiographie n'est pas autre chose qu'une niche écologique de ce genre. Il en va de la toile narrative comme de la toile d'araignée.

Certes, Dennett se méprend à son tour au sujet du statut ontologique de la personne en la réduisant lui aussi à une entité fictionnelle. Il ne semble pas apercevoir que l'explication généalogique qu'il fournit permet de rendre compte d'un processus d'institutionnalisation. L'identité narrative ne saurait se réduire à un traitement de la personne comme « centre de gravité » au sein d'une trame narrative. La personne n'est pas que pure fiction et la trame narrative n'est pas qu'une suite d'assertions qui ne seraient qu'une façon de parler lorsqu'il est question du 'je', car la narration institue globalement un artefact culturel réel. La narration met en place un dispositif qui statue autant sur ce que l'on veut être que sur ce que l'on croit être. Une identité narrative mobilise une forme spécifique de niche écologique qui est à la fois sociale, culturelle, symbolique et langagière. Le sujet reconstitué, que j'ai discuté au chapitre 12 de mon ouvrage *Pensée, langage et communauté*, en 1993, est un sujet institué¹⁸. Or, les institutions ne sont pas des fictions. L'argent, le mariage, la citoyenneté, la nation¹⁹ et l'État sont des institutions bel et bien réelles. La statue que le sculpteur crée à partir d'un morceau de glaise est peut-être le résultat de son

¹⁷ Dennett D., Daniel « The Origins of Selves », *Cogito*, 3, 1989, 163-73. Reproduit dans D. Kolak et R. Martin, (dir.), *Self & Identity, Contemporary Philosophical Issues*, Macmillan, 1991

¹⁸ Seymour M., *Pensée, langage et communauté*, Paris-Montréal, Vrin-Bellarmin, 1993.

¹⁹ Le même enjeu existe au sujet de la nation que d'aucuns considèrent trop souvent comme une entité fictionnelle en se fondant sur une interprétation superficielle de la théorie de Benedict Anderson développée dans son livre *L'imaginaire national*, Éditions de la Découverte, 2007. Contrairement à l'opinion courante, Anderson ne traite pas la nation comme une fiction. Il s'agit plutôt d'un « artefact culturel ». Or, les artefacts culturels existent bel et bien. Ce sont des biens institutionnels réels.

imagination. N'empêche, le résultat est un artefact culturel bel et bien réel. Les états intentionnels qui nous habitent sont des états que l'on institue par le langage et sont en ce sens eux-mêmes des artefacts culturels bel et bien réels. La personne instituée n'est donc pas qu'une façon de parler. Elle est un artefact culturel construit par nous. L'ipséité, comme expérience de notre unicité, peut se vivre dans les œuvres narratives que nous créons et ce, même si au départ, ce sont des œuvres qui avaient seulement été imaginées.