

PARFIT ET L'INSIGNIFIANCE DE L'IDENTITÉ

Grégoire Lefftz (SND, Sorbonne Université)

En soutenant une thèse audacieuse, provocatrice et solidement argumentée sur la question de l'identité personnelle, Derek Parfit s'est imposé comme une référence qu'aucun philosophe abordant ces questions ne peut omettre de discuter¹. Cette thèse est, dans les termes de Parfit, que l'identité n'est pas ce qui importe² (« *identity is not what matters* »), ou encore, qu'elle est insignifiante³ (« *unimportant* »). Que j'existe encore ou non dans dix heures, ou dans dix ans, ne doit pas me soucier, car ma relation avec moi-même n'est pas fondamentalement différente de celle que j'entretiens avec un autre. À peine est-elle plus étroite – mais la différence est alors de degré, et aucunement de nature. Dans ce cas, c'est cette « étroitesse » qui importe – étroitesse qui, bien qu'il reste encore à la caractériser précisément, ne peut en aucun cas être confondue avec l'identité numérique. Que celui qui sera mon « continuant » soit heureux ou en bonne santé doit me soucier ; quant à savoir s'il me sera numériquement identique ou non, cela n'importe pas.

Parfit parvient à dériver cette thèse contre-intuitive en deux étapes cruciales. Il s'attache tout d'abord à défendre une certaine conception de ce qu'est l'identité personnelle, qu'il appelle le « réductionnisme », et selon laquelle, en un mot, notre identité n'est pas un fait. Elle n'est qu'une manière de décrire certains faits impersonnels, et ne prend donc pas place parmi les éléments objectifs du monde, que nous trouverions déjà constitués, et qu'il n'y aurait pas à contester. Remarquons que Parfit ne paraît pas vouloir défendre ici une thèse ontologique générale en l'étendant, par exemple, à l'identité de l'ensemble des objets ordinaires. C'est l'identité personnelle qui l'intéresse. C'est pour autant que le réductionnisme n'est pour lui qu'un moyen de dériver certaines conclusions morales – ce qui constitue la seconde étape de son raisonnement. Si l'identité n'est pas un fait, mais une simple manière de décrire les faits, alors elle ne peut être ce qui importe. Ce sont ces faits eux-mêmes qui importent, et non leur description.

1 On le vérifiera aisément en consultant les livres de Marya Schechtman ou d'Amy Kind cités dans la bibliographie.

2 D. Parfit, *Reasons and Persons* (dorénavant *RP*), Oxford, Oxford University Press, 1984, chap. 12.

3 D. Parfit, « The Unimportance of Identity » (dorénavant *TUOI*), in R. Matin & J. Barresi (dir.), *Personal Identity*, Malden, Blackwell Publishing, 2003 (publié pour la première fois en 1995).

L'organisation de cet article suivra celle du raisonnement de Parfit : les deux premières sections seront ainsi consacrées à exposer la thèse réductionniste avec les arguments qui l'étayent, puis la conclusion selon laquelle l'identité n'importe pas, assortie de ses principales conséquences morales. Dans une troisième section, on s'attachera toutefois à interroger le raisonnement de Parfit, en adoptant la stratégie suivante. En supposant que la prémisse réductionniste est vraie (et les arguments de Parfit la rendent en effet très plausible), on contestera qu'il s'ensuit que l'identité n'importe pas. Le réductionnisme de Parfit semble fournir au contraire les outils qui permettent de comprendre pourquoi notre identité nous importe, justement parce qu'elle n'est pas déterminée par une succession de faits qui nous arrivent, mais bien par le sens que nous leur donnons.

I. Le réductionnisme en matière d'identité

I.A. La thèse réductionniste

Étant donné le nom dont la baptise Parfit, il serait facile de se méprendre sur le sens de sa thèse réductionniste. On pourrait croire en effet que, l'identité « se réduisant » à un ensemble de faits, elle *serait* elle-même l'ensemble de ces faits – or Parfit, malgré l'ambiguïté de certaines de ses formulations, veut dire précisément l'inverse. L'identité (personnelle) n'est pas un fait. Pour le comprendre, il importe de distinguer plusieurs cas. Les situations les plus ordinaires sont bien sûr celles dans lesquelles un énoncé d'identité numérique est clairement vrai ou clairement faux. Ainsi, il est tout simplement vrai que je suis la même personne qu'il y a cinq minutes ; et il est tout aussi vrai que je ne suis pas Socrate ni Emmanuel Macron. Nier ces énoncés serait tout simplement affirmer quelque chose de faux. Dans de telles situations, il serait ainsi aisé de croire que la description exprimée par ces énoncés dépeint un certain fait : le fait qu'une relation d'identité existerait, objectivement, entre moi à t et moi-même à t' , tout comme c'est un fait qu'il existe ici une table ou là, le Soleil. Comme le révèlent toutefois certains cas plus problématiques, il ne peut en être ainsi. Considérons en effet l'exemple du Dr. Jekyll et de Mr. Hyde. Ces deux personnages sont en effet très semblables sous certains aspects, et très dissemblables à d'autres égards, tant physiques que psychologiques. Sont-ils le même individu ? Ce point semble indéterminé. Il n'y a à cette question aucune réponse qui soit simplement vraie ou fausse. En d'autres termes, il est possible que l'on connaisse tous les *faits* à propos de Jekyll et de Hyde, sans connaître *ipso facto* s'ils sont ou non identiques. Car précisément, cette identité n'est pas elle-même un fait, mais un certaine

manière de les décrire. Or il y a des cas où aucune des descriptions possibles ne s'impose. Parfit qualifie de « vides » (*empty*) les questions d'identité portant sur de tels cas : car une fois les faits connus, statuer sur la relation d'identité ou de distinction qui les régirait n'apporte pas la moindre information supplémentaire⁴. Juger que Jekyll et Hyde sont identiques ne décrit pas un nouveau fait : c'est, bien plutôt, une nouvelle manière de décrire les faits, que l'on connaissait pourtant déjà tous. C'est une décision linguistique, non une découverte portant sur le monde.

Cet avertissement étant donné quant au sens de la thèse, il est possible de restituer plus rigoureusement la formulation qu'en donne Parfit. Les réductionnistes se signalent selon lui à ce qu'ils affirment :

« (1) que le fait de l'identité d'une personne à travers le temps ne consiste que dans d'autres faits, plus particuliers...

(2) que ces faits peuvent être décrits sans présupposer l'identité de cette personne, ou sans affirmer explicitement que les expériences dans la vie de cette personne sont les expériences de cette personne, ou même sans affirmer explicitement que cette personne existe. Ces faits peuvent être décrits d'une manière *impersonnelle* »⁵.

La première clause a pour fonction d'établir que l'identité personnelle n'est pas un fait de plein droit (tout en suggérant apparemment qu'il s'agit encore d'un fait, ce qui rendait la précision précédente nécessaire). Dans un article publié onze ans après *Reasons and Persons*, et intitulé « *The Unimportance of Identity* »⁶, Parfit précise le sens de cette première clause. Il la glose ainsi :

« L'existence d'une personne consiste simplement dans l'existence d'un corps, et dans l'occurrence d'une série de pensées, d'expériences, et d'autres événements mentaux et physiques »⁷.

La nature des faits qui sont censés constituer l'identité est ainsi précisée : ils sont de nature physique et psychologique. Dans cet article, Parfit précise de surcroît la nature de la relation de réduction qui existe entre de tels faits et l'identité personnelle⁸. Tout d'abord,

4 Voir *RP*, p. 213.

5 *RP*, p. 210.

6 Voir note 3.

7 *TUOI*, p. 294.

8 *TUOI*, p. 294-295.

cette relation ne doit pas être comprise selon un « réductionnisme d'identification » (*identifying reductionism*), qui affirmerait qu'une personne n'est rien d'autre que son corps et son flux de pensées, au point qu'il deviendrait impossible de les distinguer. Selon Parfit, cette position est trop forte : une personne a un corps et des pensées, dont elle peut encore être distinguée. Cette relation, par ailleurs, ne doit pas non plus être comprise comme un « réductionnisme éliminativiste » (*eliminative reductionism*), selon lequel les personnes n'existeraient pas, parce qu'il n'y aurait que des corps et des pensées. La relation appropriée, aux yeux de Parfit, est bien plutôt celle d'un « réductionnisme constitutif » (*constitutive reductionism*) : les personnes existent, et sont distinctes de leur corps et de la série de leurs pensées, mais leur existence est constituée par celle de ce corps et de cette série de pensées. Tout comme une nation – qui est l'analogie préférée de Parfit pour concevoir les personnes – n'est pas simplement identique à un certain groupe de gens ni à un certain territoire, mais est constituée malgré tout de ces éléments, et n'est pas indépendante de leur existence.

La seconde clause, quant à elle, a pour fin d'éviter l'objection de circularité qui ne peut manquer d'être opposée au réductionnisme – et qui a été en effet opposée à Locke, peut-être le premier réductionniste au sens où l'entend Parfit, par l'évêque Butler. Ce dernier remarquait que la continuité psychologique, et notamment mémorielle, ne peut constituer l'identité d'une personne, parce que pour savoir si plusieurs pensées sont continues au sein d'un même flux de conscience, il est nécessaire de savoir déjà si ce sont ou non les pensées d'une même personne. C'est la continuité de la conscience qui présuppose donc l'identité de la personne, et non l'inverse⁹. Cette identité personnelle est à son tour fondée, selon Butler, sur la possession continue d'une âme, éternelle et inobservable. Selon Parfit, la position de Butler est au fond celle que doit accepter tout opposant au réductionnisme. Soit l'identité personnelle consiste en faits observables, physiques et/ou psychologiques, qui doivent alors pouvoir être décrits en termes impersonnels sous peine de circularité, comme le prescrit la deuxième clause (c'est la position réductionniste). Soit elle consiste en l'existence d'un « fait supplémentaire » (*further fact*) inobservable, telle que la possession d'une âme ou d'un *ego* cartésien, auquel cas elle constitue bien un fait de plein droit, primitif et irréductible à d'autres faits, physiques ou psychologiques.

9 J. Butler, *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*, Londres, George Bell & Sons, 1897, premier appendice (publication originale en 1736).

Or Parfit considère que la théorie du fait supplémentaire n'est tout simplement pas plausible. De manière remarquable, il donne pour cela un argument empirique¹⁰. Nous aurions pu avoir des raisons d'admettre la persistance d'une âme : cela aurait par exemple été le cas si des phénomènes de réminiscence de vies antérieures avaient été observés de manière massive et vérifiable (une telle réminiscence serait constatée, par exemple, si une japonaise disait avoir enterré, dans une vie de celte deux-mille ans auparavant, des outils au pied d'un certain menhir, ce qui serait relativement facile à vérifier). Le fait est que nous n'avons simplement pas de telles raisons empiriques de penser que la théorie du fait supplémentaire est vraie. Il est donc plus raisonnable de penser que notre identité consiste en des faits physiques ou psychologiques observables, qui doivent donc pouvoir être décrits, quant à eux, de manière impersonnelle¹¹. Le point crucial est alors qu'à ces faits, l'identité n'ajoute rien. C'est là le sens fondamental de la thèse réductionniste. Comme l'écrit Parfit à propos des questions vides, telles que celles de l'identité et de la distinction de Jekyll et de Hyde :

« Lorsque nous posons une question vide, il n'y a qu'un seul fait ou qu'un seul résultat qui est porté à notre considération. Différentes réponses à notre question sont simplement des descriptions différentes de ce fait ou de ce résultat. C'est pourquoi, sans avoir répondu à cette question vide, nous pouvons déjà savoir tout ce qu'il y a à savoir. [L'identité ou la distinction de Jekyll et Hyde] ne sont pas des possibilités différentes, dont l'une devrait être vraie »¹².

Cela, bien sûr, serait faux si l'identité de Jekyll consistait dans la possession continue d'une âme (que Hyde posséderait ou ne posséderait pas quant à lui, de manière déterminée). Mais c'est précisément ce que nie la position réductionniste.

I.B. Premier argument en faveur du réductionnisme (RP, §§84-86)

Comme cela vient d'être expliqué, Parfit considère qu'entre la théorie réductionniste et celle du fait supplémentaire, le tiers est exclu. Si l'une est vraie l'autre est fautive. Sa stra-

10 RP, §82.

11 Parfit construit, pour permettre une telle description, le célèbre concept de « quasi-souvenir » (*quasi-memory*), qui élimine la référence à la personne ayant eu l'expérience à l'origine du souvenir. Pour me souvenir d'avoir vu un feu d'artifice, il faut que je me rappelle l'avoir vu, et que je l'aie vu en effet ; mais pour m'en quasi-souvenir, il est seulement nécessaire que je me rappelle l'avoir vu, et que *quelqu'un* l'ait vu en effet. Voir RP, §80.

12 RP, p. 214. L'exemple de Jekyll et de Hyde est le mien, et non celui que donne Parfit.

tégie pour prouver la première est donc de réfuter la seconde. Or cette théorie du fait supplémentaire implique que notre identité est toujours clairement déterminée. En effet, je possède ou je ne possède pas une âme. Il ne paraît pas y avoir ici de place pour le vague. Si je continue de la posséder, je demeure moi-même, quels que soient les changements physiques ou psychologiques que j'endure ; et si je cesse de la posséder, je perds mon identité numérique, quelque semblable que je demeure par ailleurs à celui que j'étais. La théorie du fait supplémentaire impliquant donc la détermination de l'identité, il suffit pour la réfuter de raisonner par contraposition : si l'on peut prouver que, dans certains cas, l'identité personnelle est indéterminée, on aura prouvé que la théorie du fait supplémentaire est fautive (et donc que le réductionnisme est vrai). C'est la stratégie qu'adopte Parfit. Son premier argument consiste à examiner successivement plusieurs spectres continus de changement qualitatif que peut subir un individu, pour établir qu'au milieu de ce spectre, l'identité de cet individu est en effet indéterminée.

Parfit commence par envisager un spectre psychologique (RP, §84), qui présente un large éventail de degrés de connexion psychique entre les états mentaux d'un individu (la connexion psychique se définissant par la capacité d'accès direct qui caractérise notamment la mémoire : une telle connexion existe par exemple entre mon état mental actuel, et l'état mental dans lequel j'étais lorsque j'ai posé mes lunettes sur la commode il y a une heure, *car je suis capable de m'en souvenir*). Au début du spectre, la continuité psychologique est maximale : on considère deux personnes x et y qui partagent tous leurs souvenirs, toutes leurs intentions, tous leurs désirs, toutes leurs angoisses. À l'autre bout du spectre, la discontinuité psychique entre deux personnes x et y est à l'inverse maximale : celles-ci ne partagent aucun état mental commun. La description intuitive est bien sûr que dans le premier cas, x et y sont la même personne (il existe entre eux la relation qu'il y a entre deux « phases » de moi-même à quelques minutes d'intervalle) ; et que dans le second cas, x et y sont deux personnes distinctes. Mais qu'en est-il au milieu du spectre, quand la continuité et la discontinuité psychique s'équilibrent presque exactement ?

Si l'on pouvait conclure que, dans un tel cas, la relation d'identité (ou de distinction) entre x et y serait indéterminée, Parfit aurait atteint son objectif. Cette conclusion, toutefois, ne s'impose pas immédiatement. L'on pourrait aussi bien en conclure que le critère psychologique de l'identité personnelle n'est pas le bon. Ce n'est pas de conserver sa continuité psychique qui fait d'une personne ce qu'elle est, mais, pourrait-on dire aussi

bien, de conserver une certaine continuité physique. Telle était la position de Bernard Williams¹³, dont Parfit s'inspire ici en imaginant de tels spectres de changement continu :

« [Williams] préfère conclure que, si mon cerveau continue d'exister, et d'être le cerveau d'une personne vivante, je serai cette personne. Cela serait le cas même si, entre moi maintenant et moi-même plus tard, il n'y avait *aucune* connexion psychologique »¹⁴.

Mais si tel est le cas, la relation entre x et y au milieu du spectre est finalement pleinement déterminée, et l'argument échoue.

Pour éviter cet écueil, Parfit envisage un second spectre, physique cette fois (*RP*, §85). Au lieu de différer de plus en plus par leur degré de continuité psychologique, les individus envisagés diffèrent progressivement par leurs caractères physiques, et se ressemblent ainsi de moins en moins, jusqu'à ne plus avoir un seul atome en commun. Tout le problème est qu'en chaque lieu de ce spectre, le couple d'individus considéré est relié par une continuité psychologique totale (comme ils l'étaient, dans le cas précédent, par une similarité physique totale). Comme l'argument du spectre psychologique était vulnérable à l'objection de partisans d'un critère d'identité physique, l'argument du spectre physique est ainsi impuissant à convaincre le partisan d'un critère d'identité psychologique. Plutôt que d'admettre que les cas intermédiaires sont indéterminés, celui-ci préférera conclure que x et y sont bien identiques, sur toute l'ampleur du spectre.

Parfit, tirant la leçon de ce double échec, choisit alors d'envisager un « spectre complexe » (*combined spectrum*), mêlant des changements progressifs à la fois physiques et psychologiques (*RP*, §86). À mesure que l'on avance dans le spectre, les individus x et y imaginés sont de plus en plus dissemblables à ces deux égards. À une extrémité du spectre, l'individu considéré est qualitativement identique, à la fois physiquement et psychologiquement, à Derek Parfit ; à l'autre extrémité, il est qualitativement identique à Greta Garbo. Cette fois, il n'est plus possible de recourir à un critère physique pour conclure que les changements psychologiques n'importent pas, ni de recourir à un critère psychologique pour obtenir la conclusion symétrique à propos des changements physiques. Au milieu du spectre, cette personne tient autant de Parfit que de Garbo, tant

13 B. Williams, « The Self and the Future », in R. Matin & J. Barresi (dir.), *Personal Identity*, Malden, Blackwell Publishing, 2003 (publié pour la première fois en 1970).

14 *RP*, p. 233.

physiquement que psychologiquement. Est-elle Garbo ou Parfit ? Selon ce dernier, il s'agit là d'une question vide. Nous pouvons décrire la situation de deux manières différentes, mais c'est d'une seule et même situation qu'il s'agit. Si tel est bien le cas, alors la conception réductionniste de l'identité est vraie : celle-ci n'est pas un fait, elle n'est qu'une manière de décrire certains faits impersonnels.

I.C. Deuxième argument en faveur du réductionnisme (RP, §§87-88)

Le second argument de Parfit se fonde sur la description des conséquences d'opérations de commissurotomie cérébrale, menées dans les années 1950 pour tenter de guérir des patients atteints d'épilepsie. Ces opérations consistent à trancher le corps calleux reliant les deux hémisphères du cerveau. Les patients ainsi traités parvenaient à vivre normalement, mais développaient deux « sphères de conscience » distinctes, comme le révélèrent certaines expériences. Par exemple (selon la présentation d'un cas légèrement simplifié par Parfit), lorsque l'on montrait à l'un de ces patients un écran dont la moitié gauche était rouge et la moitié droite était bleue, et dont chacune de ces moitiés comportait l'inscription « quelle est l'unique couleur que vous voyez ? », le patient écrivait « rouge » avec l'une de ses mains et « bleu » avec l'autre.

Parfit se fonde alors sur la possibilité empiriquement avérée d'une telle dissociation de la conscience pour imaginer le scénario suivant (*RP*, § 87). Supposons qu'après avoir subi une telle opération, j'aie acquis la capacité de maîtriser la dissociation de ma conscience. Par exemple, lorsque je hausse les sourcils ma conscience se dissocie en deux flux distincts, et lorsque je les hausse à nouveau elle se réunifie. Supposons que je me trouve dans la situation de passer un examen de physique, et que j'envisage deux manières différentes de résoudre un problème, dont j'ignore laquelle s'avérera la plus efficace. Dans ce cas, il me serait utile de dissocier ma conscience, et de laisser chacun de mes deux « flux », contrôlant chacun une de mes mains, accomplir l'un des deux calculs pendant dix minutes, avant de les réunifier pour comparer leurs résultats. Pendant la durée du calcul, je n'aurais conscience dans aucun de mes deux flux de ce que fait l'autre (chacun est, vu de l'autre, comme une autre personne) ; après leur réunification je me rappellerais en revanche, dans mon unique flux de conscience, ce que j'aurais fait dans chacun d'eux.

Selon Parfit, la bonne description de cette situation est que, pendant dix minutes, je possède un « esprit divisé ». Il n'apparaît pas soudain deux personnes distinctes l'une de

l'autre, et distinctes de moi, de sorte que cette expérience de pensée impliquerait trois protagonistes au total. Étant donné la brièveté de la période de division, ainsi que la continuité de conscience qui existe entre chacune de ces deux sphères et ma conscience réunifiée au terme de l'expérience, il est bien plus plausible de dire que cette situation n'implique qu'une seule personne, qui sépare temporairement sa conscience en deux sphères¹⁵.

Mais si cette description est correcte, elle nous laisse face au défi d'expliquer ce qui fait l'unité de chacune de ces deux sphères de conscience pendant la période de division (RP, §88). Dans cette expérience de pensée, il y a en effet une unique personne possédant deux sphères de conscience : il est alors impossible de dire que ce qui unifie chacune de ces sphères de conscience, c'est d'être la conscience *d'une personne* (sans quoi ces deux sphères seraient unifiées entre elles).

Il paraît ainsi impossible d'expliquer l'unité de la conscience par l'identité d'une personne. Parfit remarque alors que l'option inverse demeure disponible. Cette expérience de pensée semble en effet indiquer que l'identité d'une personne s'explique bien plutôt par l'unité de son flux de conscience. Parce que la période de dissociation n'est qu'un bref épisode au sein d'un large flux de conscience continue, elle doit être décrite comme la dissociation temporaire de la conscience d'une unique personne. Or inverser ainsi l'ordre de l'explication, c'est nier la théorie du fait supplémentaire, et admettre la position réductionniste :

« [Le réductionnisme] est corroboré par le cas de l'esprit divisé. Il n'est pas seulement vrai, en effet, que l'unité des différentes expériences *n'a pas besoin* d'être expliquée en me les attribuant. L'unité de mes expériences, dans chaque flux, *ne peut pas* être expliquée ainsi. Il y a seulement deux possibilités. Nous pouvons attribuer les expériences dans chaque flux à un sujet d'expériences qui n'est *pas* moi, et donc, n'est pas une personne. Ou, si nous doutons de l'existence de telles entités, nous pouvons accepter l'explication réductionniste. Au moins dans ce cas, cela semble être la meilleure explication »¹⁶.

En d'autres termes, si l'on souhaite éviter d'admettre que, pendant la courte période de dissociation envisagée, chaque sphère de conscience est soutenue et unifiée par une âme

15 RP, p. 248.

16 RP, p. 251.

inobservable – ce qui semble peu plausible –, il faut bien accepter la position réductionniste, selon laquelle ce n'est pas l'identité de la personne qui fonde l'unité de sa conscience (mais l'inverse)¹⁷.

II. L'insignifiance de l'identité

II.A. « Le réalisme de l'importance »

À partir du réductionnisme en matière d'identité personnelle, il paraît facile de conclure que l'identité n'importe pas. L'inférence n'est toutefois pas immédiate. On ne peut pas tirer de conclusion à partir d'une seule prémisse : il en faut au moins deux. Or la deuxième de ces prémisses paraît si évidente à Parfit qu'il ne l'énonce pas explicitement dans *Reasons and Persons*. Cette lacune est toutefois comblée dans l'article « *The Unimportance of Identity* », où Parfit formule ce qu'il appelle le « réalisme de l'importance » (*realism about importance*)¹⁸.

Cette thèse énonce que, lorsqu'un fait de niveau supérieur est de nature purement conceptuelle, ce fait n'importe pas : seuls importent les faits de niveau inférieur. Tentons d'expliquer ce que cela signifie. Toute position réductionniste implique une distinction d'au moins deux niveaux : le niveau supérieur qui est réduit, et le niveau inférieur auquel il est réduit. Or il importe de distinguer ici deux types de réductionnisme. Selon le premier type, le niveau supérieur peut posséder des propriétés qui ne se retrouvent pas au niveau inférieur. C'est ce que l'on constaterait si le physicalisme était vrai, de sorte que tous les faits consisteraient ultimement en faits concernant les particules fondamentales (ou les champs, ou les entités, quelles qu'elles soient, que la physique pose comme fondamentales). Dans un tel cas, une relation amoureuse entre deux personnes se réduirait à des faits concernant les particules. Mais cela ne signifie pas que les particules elles-mêmes seraient amoureuses. Seules des personnes le seraient. Dans cette situation, il serait évidemment absurde de dire que rien de ce qui arrive au niveau supérieur ne pourrait avoir d'importance propre, parce que seul ce qui se produit au niveau inférieur en posséderait réellement.

17 Dans les cas où cette unité est totale, il serait faux de dire qu'il y a deux personnes. Dans les cas où cette unité est *presque* totale (comme celui de l'examen de physique, où je ne dissocie ma conscience que quelques minutes), il paraît également correct de conclure qu'il n'y a qu'une personne. À mesure que l'unité de la conscience s'estompe, toutefois, l'identité de la personne devient de plus en plus indéterminée, comme au milieu du spectre complexe. On retrouve bien l'idée que, si le réductionnisme est vrai, l'identité personnelle n'est qu'une manière de décrire des faits impersonnels.

18 TUOI, p. 308.

Il n'en va toutefois pas de même pour un autre type de réductionnisme, qui est précisément celui dont parle Parfit. Dans ce second cas, la relation de réduction est simplement conceptuelle : ce que l'on comprend comme les faits de niveau supérieur ne sont, en réalité, que les faits du niveau inférieur décrits d'une nouvelle manière. Dans ce cas, il paraît plausible de dire que seuls les faits du niveau inférieur importent, car eux seuls sont des faits du monde (le niveau supérieur ne leur ajoute que des faits concernant nos choix linguistiques). Comme l'écrit Parfit :

« Nous discutons des cas dans lesquels, relativement aux faits du niveau inférieur, le fait de niveau supérieur est, au sens esquissé, purement conceptuel. J'affirme que de tels faits conceptuels ne peuvent pas être rationnellement ou moralement importants. Ce qui importe est la réalité, non la manière dont elle est décrite. Cette thèse peut être appelée le *réalisme de l'importance* »¹⁹.

Ces prérequis étant en place, il est alors facile de restituer le raisonnement de Parfit. Il procède comme suit :

(i) L'identité personnelle n'est pas un fait primitif, mais n'est qu'une manière de décrire des faits impersonnels (réductionnisme).

(ii) Ce qui importe est la réalité, non la manière dont elle est décrite (réalisme de l'importance).

(C) Donc l'identité n'importe pas.

Ce raisonnement étant clairement valide, si l'on concède à Parfit ses deux prémisses, il faut alors admettre également sa conclusion. Encore importe-t-il de bien la comprendre. Si l'identité n'importe pas, qu'est-ce qui importe exactement ? Ce sont, bien sûr, les faits de premier ordre dont l'identité n'est qu'une description ; mais en quoi consistent précisément ces faits ?

Parfit répond à cette question en élaborant ce qu'il appelle la « relation *R* ». Pour cela, il commence par proposer quelques définitions conventionnelles²⁰. Parfit appelle ainsi « *connexion psychologique* » (*psychological connectedness*) l'existence, entre deux états mentaux, de relations psychologiques directes (comme dans le cas où je préserve certains souvenirs, où je conserve certaines intentions, etc.). Il appelle « *connexion forte* » la relation qui existe entre deux moments de la vie psychologique d'une personne si 50% au moins des relations de connexion psychologique sont conservées de l'un à l'autre. Il

19 *Ibid.*

20 *RP*, p. 206.

appelle enfin « *continuité psychologique* » la relation qui existe entre deux moments de la vie psychologique d'une personne s'ils sont reliés par une chaîne de connexions fortes qui se chevauchent au moins partiellement (par exemple une personne *a* est psychologiquement *continue* avec une personne *c*, même si elles ne sont pas fortement *connectées*, pourvu qu'il y ait une personne *b* telle que *a* et *c* soient toutes deux fortement connectées à *b*). La relation *R* se définit alors comme « la connexion et/ou la continuité psychologique », avec n'importe quelle cause. Selon Parfit, *c'est la relation R qui importe*. Que je sois ou non identique à une certaine personne n'est pas l'essentiel. Ce qui importe est que j'entretienne avec elle une relation de continuité psychique suffisante. Ainsi seulement sera-t-elle susceptible d'accomplir mes intentions et de réaliser mes désirs, de perpétuer mes souvenirs, etc. Pourvu que cela soit acquis, que cette personne soit moi ou non n'importe pas.

II.B. Un argument supplémentaire (RP, §§89-90)

N'est-il pas toutefois tentant de répondre que si une personne entretient une relation *R* avec moi, alors *ipso facto*, je suis cette personne ? Parfit congédie cette objection par une nouvelle expérience de pensée, qui vise cette fois à prouver que si l'identité et la relation *R* coïncident dans les situations ordinaires (ce qui explique selon lui notre croyance en l'importance de l'identité), il est néanmoins concevable qu'elles soient dissociées. Dans ce cas, c'est bien la relation *R* qui importerait, et non l'identité.

Pour le montrer, Parfit remanie un cas qu'il emprunte à David Wiggins²¹. Imaginons qu'au cours d'un accident mon corps soit gravement blessé, ainsi que le cerveau de chacun de mes deux frères jumeaux. Chaque hémisphère de mon cerveau est alors transféré dans le corps de l'un de mes frères, de sorte que chacun des deux individus résultant de l'opération (appelons-les *A* et *B*) se réveille avec mes souvenirs, mes intentions, et en pensant être moi. Ma conscience s'est ainsi divisée, comme dans le cas de l'examen de physique, mais chacune des deux sphères ainsi créée s'est en même temps logée dans un corps différent, sans espoir de réunification. Parfit demande : qui suis-je au terme de l'opération ? Il y a évidemment quatre possibilités : soit je ne suis aucun des deux ; soit je suis *A* ; soit je suis *B* ; soit je suis les deux à la fois²².

21 D. Wiggins, *Identity and Spatio-Temporal Continuity*, Oxford, Basil Blackwell, 1967, p. 50.

22 RP, p. 254-256.

Or Parfit montre qu'aucune de ces quatre possibilités ne paraît vraiment admissible. La première option est vouée à l'échec parce qu'il serait vraisemblablement correct de dire que je survivrais si un seul hémisphère de mon cerveau était transplanté dans le corps de l'un de mes frères, alors que l'autre aurait été détruit – pourvu que cela permette, comme le suppose l'expérience, une pleine et entière continuation de ma conscience. Mais comment le fait que mon autre hémisphère soit lui aussi transféré dans un corps pourrait-il changer quoi que ce soit à cette situation ? « Comment un double succès pourrait-il être un échec ?²³ »

Quant à la possibilité d'être soit *A*, soit *B*, le problème qu'elle pose est évident : pourquoi l'un plutôt que l'autre, puisque par hypothèse, leurs situations sont parfaitement symétriques ? Il faudrait qu'un fait puisse justifier que l'un des deux individus soit choisi de préférence à l'autre ; mais par les données même de l'expérience de pensée ici envisagée, ce fait n'existe pas.

Ne reste donc que la possibilité que je sois les deux individus à la fois. Mais dire cela requerrait de distordre les concepts de *personne* et d'*identité* bien au-delà de leur sens habituel. Cela requerrait en effet de dire que les deux individus *A* et *B* sont une unique personne ; mais chacun d'eux possède pourtant tous les caractères qui font habituellement d'un individu une personne unique. Ils sont physiquement et psychologiquement indépendants, ils peuvent vivre en des lieux différents et développer des goûts, des idées, des désirs différents. Comment ne seraient-ils pas deux personnes ? Si à l'inverse ils n'en étaient qu'une seule, à savoir moi-même, il deviendrait possible de dire des choses comme : « Ce que vous voyez là-bas est une seule personne jouant au tennis avec elle-même. Dans chaque moitié de son esprit, elle croit à tort qu'elle joue au tennis avec quelqu'un d'autre ». Cette situation est évidemment absurde : elle suppose de prendre le concept de *personne* en un sens qu'il ne possède clairement pas.

Parfit conclut que dans une telle situation, la relation d'identité est donc indéterminée – et impossible à déterminer autrement que par une stipulation arbitraire. Si l'identité était ce qui importe, alors ce qui importe se serait perdu. Se réveiller dans un lit d'hôpital serait préférable à la mort ; mais se réveiller dans deux lits d'hôpital ne le serait pas. C'est exactement ce que nie Parfit. Certes, dans la situation décrite, rien ne permet de dire que je suis identique à l'un des deux individus résultant de l'opération (ou aux deux) ; mais j'entretiens une relation *R* avec chacun des deux. Chacun d'eux se réveille

23 *RP*, p. 256.

en pensant être moi, en ayant tous mes souvenirs, en ayant le sentiment de continuer ma vie. Cela paraît assurément préférable à la situation où ma conscience aurait purement et simplement disparu dans l'accident. Parfit conclut que c'est bien cela qui importe – et non l'identité.

II.C. Conséquences morales : l'atténuation du moi

L'insignifiance de l'identité a bien sûr, aux yeux de Parfit, d'importantes conséquences morales, qui sont étudiées en détail principalement dans les chapitres 14 et 15 de *Reasons and Persons*, et dont il est possible de donner ici un aperçu. Il faut d'abord dire un mot de la conception du soi qui est ici défendue. Aux yeux de Parfit, l'identité des personnes n'est pas un fait, mais une simple manière de décrire des faits impersonnels. Cela revient à dire que le monde tel qu'il est, indépendamment de la manière dont nous le percevons, ne contient pas de personnes – de même qu'indépendamment de nos constructions intellectuelles ou de nos conventions, il ne contient pas de nations. Celles-ci sont des vues de l'esprit, des fictions : c'est précisément le sens dans lequel Parfit cite à plusieurs reprises certains textes bouddhistes²⁴. En conséquence, la frontière entre le moi et le non-moi, sans disparaître tout à fait, ne peut manquer de se brouiller.

Parfit dit considérer cette perspective avec soulagement. Répétant sa conclusion selon laquelle notre identité ne réside pas dans un « fait supplémentaire » profond, tel que la possession d'une âme, mais au contraire dans la continuité physique et psychologique, qui admet quant à elle des degrés, il écrit :

« Cette vérité est-elle déprimante ? Certains pourraient le penser. Mais je la trouve libératrice et consolatrice. Lorsque je croyais que mon existence était un tel fait supplémentaire, j'avais le sentiment d'être emprisonné en moi-même. Ma vie était semblable à un tunnel de verre, à travers lequel j'avancais plus vite chaque année, et au bout duquel il n'y avait que ténèbres. Lorsque j'ai changé d'opinion, les parois de mon tunnel ont disparu. Je vis maintenant à l'air libre. Il y a encore une différence entre ma vie et la vie des autres personnes. Mais cette différence est plus faible. Les autres sont plus proches. Je suis moins soucieux du reste de ma vie, et plus soucieux de la vie des autres »²⁵.

24 Voir notamment l'appendice J à *RP*, p. 502-503, ainsi que *RP*, p. 273.

25 *RP*, §95, p. 280.

Ainsi par exemple, il y a aujourd'hui des souvenirs qui sont étroitement liés à ma vie, à savoir ceux que j'entretiens ; et il y aura encore après ma mort des souvenirs liés à ma vie, quoique moins étroitement : ceux qu'entretiendront mes proches. La différence entre ces deux états semble être de degré plutôt que de nature. Congédier l'idée d'un « fait supplémentaire » conduit ainsi, aux yeux de Parfit, à ce qu'il faut bien appeler une atténuation du moi. Celui-ci n'a plus en effet de frontières nettes : ses limites dépendent d'une décision, qui d'ailleurs n'est pas contrainte par les faits au point qu'il n'y en aurait qu'une seule possible et légitime.

Cette atténuation a des conséquences importantes pour notre conception de la rationalité. Selon la conception classique, par exemple, être rationnel consiste à chercher à maximiser la satisfaction de son propre intérêt. Or un aspect particulièrement important de cette conception réside dans la clause de « souci égal » (*equal concern*) : les différents intérêts d'un agent rationnel ne doivent pas être pondérés en fonction du temps. Il serait ainsi rationnel d'endurer un faible mal présent afin de permettre un grand bien à venir, et irrationnel de s'y refuser. Or, remarque Parfit, cette conception suppose précisément que l'identité n'est pas une question de degrés, et que chacun demeure également lui-même tout au long de son existence. Mais c'est précisément ce qu'il nie. Je suis plus étroitement connecté à celui que je serai dans cinq minutes qu'à celui que je serai dans dix ans. Il est par conséquent rationnel de pondérer la satisfaction de ses intérêts par la proximité que l'on entretient avec la phase de soi-même qui en retirera les bénéfices, ce qui est en partie fonction du temps. La clause de souci égal doit donc être rejetée (*RP*, §105).

D'autres conséquences plus directement morales résident dans le fait qu'il faut, d'après Parfit mais pour reprendre le titre de Ricœur, se traiter « soi-même comme un autre ». Par exemple, en vertu de ce qui vient d'être dit, il n'est pas forcément irrationnel, c'est-à-dire contraire à son intérêt égoïste, de s'accorder un plaisir présent au risque d'un mal futur, tel que celui de fumer dans sa jeunesse au risque de développer plus tard un cancer. Cela n'est certes pas irrationnel, mais cela doit être critiqué comme immoral. Tout comme il est immoral de causer son propre plaisir au risque de la souffrance d'autrui, et pour la même raison, il est immoral de s'accorder un plaisir de jeunesse au risque d'une souffrance ultérieure. Nous devons à celui que nous serons plus tard, non point le calcul intéressé qui maximisera notre propre plaisir (ultérieur), mais le respect désintéressé par lequel nous traitons autrui moralement (*RP*, §§ 106).

Si la position « bouddhiste » que défend Parfit doit nous rendre plus sensibles au sort des autres, c'est donc en réalité par une subtile dialectique. Il nous faut nous soucier des autres parce qu'ils sont d'autres nous-mêmes, c'est-à-dire, parce que la frontière entre les autres et le soi s'estompe. Mais c'est seulement pour autant que nous sommes à nous-mêmes comme des autres : non point des substances monolithiques vouées à exister totalement (ou pas du tout), mais des êtres de degrés, reliés plus ou moins étroitement à ceux que nous continuons d'être.

III. Perspective critique

Après avoir exposé pour elle-même la thèse de Parfit, assortie des raisons qui l'étayent, on se propose ici de l'éprouver. Il y a assurément – et Parfit le reconnaît volontiers – quelque chose d'extrêmement contre-intuitif dans sa position. La différence entre le moi et le non-moi paraît ne pas être de degré : la mort est habituellement conçue comme une rupture abrupte entre l'existence et l'inexistence, plutôt qu'entre deux états d'existence plus ou moins intenses ; et si nous pouvons nous soucier du malheur d'autrui, ce n'est pas de la manière dont nous nous soucions du nôtre, c'est-à-dire en première personne (l'incapacité à faire cette différence étant clairement diagnostiquée comme pathologique). N'y a-t-il pas un soubassement rationnel à ces intuitions ? Et si la position de Parfit aboutit à les nier, ne devrait-elle pas du moins en rendre compte, en nous expliquant l'origine de notre erreur ?

Nous adopterons dans cette partie la stratégie suivante. Parfit, on l'a dit, déduit sa thèse selon laquelle l'identité n'importe pas de deux prémisses distinctes. La première consiste dans la thèse réductionniste, selon laquelle l'identité personnelle n'est pas un fait, mais une simple description de faits impersonnels. La seconde réside dans le « réalisme de l'importance », selon laquelle « ce qui importe est la réalité, non la manière dont elle est décrite ». On s'attachera ici à concéder la première prémisse, mais à contester la seconde. Même si le réductionnisme est vrai²⁶, la manière dont nous décrivons la réalité nous importe au plus haut point, et c'est précisément pourquoi notre identité nous importe.

Cette stratégie argumentative repose, dans son esprit, sur l'idée suivante : le réductionnisme de Parfit, loin d'impliquer par lui-même que l'identité est insignifiante, nous fournit au contraire les outils nécessaires pour comprendre pourquoi elle nous importe.

26 Je suis pour ma part convaincu de la vérité de cette thèse.

Car le réductionnisme affirme que notre identité n'est pas quelque chose qui nous arrive, mais au contraire quelque chose que nous faisons : elle n'est qu'une description. Elle n'est pas simplement un fait que nous trouverions déjà constitué dans le monde, mais résulte du sens que nous donnons aux événements de notre vie – sens à partir duquel nous pouvons finalement prendre une décision, ou émettre un jugement, concernant les bornes de notre identité. C'est justement pourquoi elle ne nous est pas indifférente. D'un côté, Parfit concède que nous sommes actifs vis-à-vis de notre identité, puisqu'elle est une description que nous donnons des faits qui nous concernent, elle réside dans le sens que nous leur insufflons, dans le choix que nous faisons de les reprendre à notre compte ou au contraire de nous en dissocier (c'est le réductionnisme). D'un autre côté, il affirme que nous sommes passifs vis-à-vis d'elle, puisque ce sens que nous donnons aux faits ne devrait faire pour nous aucune différence quant à ce qui importe – seuls compteraient ces faits eux-mêmes, muets et subis (c'est le réalisme de l'importance). Ces deux prémisses peuvent ainsi sembler menacées d'inconsistance, et c'est ce risque que l'on se propose de souligner et d'explorer dans cette section.

III.A. L'irréalisme de l'importance

Si la seconde prémisse de Parfit, exprimant « le réalisme de l'importance », paraît plausible au premier abord, sa vraisemblance s'estompe à mesure qu'on l'examine plus attentivement. Est-il vrai que « ce qui importe est la réalité, non la manière dont elle est décrite » ? Amie Thomasson, qui examine dans un article récent ce qui est en jeu dans les débats portant non sur les faits qui existent (ou non), mais sur la manière dont nous devons décrire ces faits, formule ainsi les raisons d'en douter :

« La manière dont nous utilisons les mots importe, étant donné leur relation à d'autres aspects de notre schème conceptuel, ainsi qu'à notre comportement non-verbal. Traiter un individu comme un "athlète", par exemple, est lié à tout un champ d'honneurs sociaux et de récompenses – ainsi, en demandant si des animaux non-humains [comme des chevaux] peuvent légitimement recevoir ce titre, nous négocions en même temps la manière dont il faut les considérer et les traiter, et dont nous devons distribuer les récompenses concernées. Ce cas n'est en aucune façon isolé. Dans de très nombreuses situations nous nous engageons dans des débats animés qui ne se fondent pas sur des thèses différentes concernant ce qui existe empiriquement, mais plu-

tôt sur des désaccords concernant la manière dont nos termes doivent être employés dans un certain contexte. Considérons notre usage du mot “torture”. Dans les débats demandant si le simulacre de noyade [*waterboarding*] est une torture, les protagonistes ne diffèrent pas quant à leur connaissance de cette pratique, de ses effets sur la personne interrogée, etc. Pourtant, on manque la profondeur et le sérieux du débat si on le qualifie de purement verbal »²⁷.

Notre manière de décrire la réalité possède ainsi une importance intrinsèque, irréductible à celle des faits décrits : décider ou non de caractériser le simulacre de noyade comme une torture pourra ainsi nous déterminer à l'utiliser, ou au contraire à l'interdire.

Exactement de la même manière, dans les zones indécises des « spectres » envisagés par Parfit, ou de manière plus réaliste, dans les situations où une personne subit des changements considérables dans sa vie ou sa personnalité (comme lors de cas de conversion, d'amnésie, de démence, etc.), que cette personne choisisse de s'identifier à celle qu'elle était et de concevoir sa nouvelle vie dans la continuité de l'ancienne, ou à l'inverse de s'en dissocier et de la désavouer, ne pourra sans doute manquer de faire une différence. Il y a certes un fait objectif (celui de l'amnésie consécutive à un accident, par exemple) ; mais le *sens* donné par la personne à ce fait est lui-même crucial. Si ce fait est conçu par la personne qu'il concerne comme un obstacle à surmonter dans la trame continue d'une seule vie, il est à gager que cette personne s'efforcera de renouer le fil de son existence en continuant, si elle le peut, de fréquenter les mêmes personnes, en s'efforçant de recouvrer ses souvenirs, en continuant de cultiver les goûts dont ses proches lui disent qu'ils étaient les siens. Si cette personne conçoit à l'inverse cet accident comme une rupture irréparable, elle se résoudra à ne pas rechercher vainement son passé à jamais perdu, et à se ménager un nouvel avenir en cultivant de nouvelles relations, en déménageant, en construisant en un mot une nouvelle vie. Quelle différence entre ces deux manières de décrire son existence, de donner un sens aux mêmes faits !

Il importe du reste de souligner que, même si de telles différences dans la description que nous donnons de notre vie – et dans notre manière de nous identifier à une personne ou de nous en désidentifier – n'avaient aucune conséquence *pratique*, et ne chan-

27 A. Thomasson, « *Metaphysical Disputes and Metalinguistic Negotiation* », *Analytic Philosophy* (vol. 0, n° 0), 2016, p. 11.

geaient rien à notre comportement, elles pourraient néanmoins avoir une importance symbolique considérable. Il y a sans aucun doute une différence cruciale entre le meurtrier qui regrette son crime de toutes ses forces, et celui qui se le rappelle avec délectation, et ce même s'il n'a pas pour projet de recommencer. De la même manière, que nous traitions deux individus *a* et *b* comme la même personne, ou comme deux personnes distinctes, est d'une importance cruciale. Les conséquences en sont certes pratiques : selon la réponse donnée, Jekyll pourra ou non être tenu responsable des crimes de Hyde. Mais elles sont également symboliques : Utterson ne ressentira assurément pas la même chose vis-à-vis de Hyde selon qu'il s'agira de son ami Jekyll ou d'un tiers. Jekyll lui-même est pris dans cette ambiguïté : il ne peut jouir d'avoir commis les crimes de Hyde que s'il s'identifie à lui, mais ne peut échapper à la culpabilité que s'il considère que c'est un *autre* qui les a commis. Il lui faut donc choisir. Ce choix, assurément, ne porte que sur deux manières différentes de décrire les mêmes faits. Mais une importance capitale s'y attache sans aucun doute.

On le voit, la prémisse du réalisme de l'importance paraît éminemment contestable. Il faudrait, semble-t-il, la remplacer par ce que l'on pourrait appeler par souci de symétrie « l'irréalisme de l'importance » : *la réalité importe évidemment, mais la manière dont elle est décrite possède elle aussi son importance*, à la fois symbolique et pratique. Une fois cette thèse admise, l'argument de Parfit devient un *non sequitur*. On peut certes admettre la thèse réductionniste, mais la conclusion selon laquelle l'identité n'importe pas n'en suit plus.

III.B. La description de notre vie prend la forme d'un récit

Ce résultat négatif étant acquis, il est à présent possible de défendre un résultat positif : la prémisse réductionniste, si on lui ajoute l'irréalisme de l'importance, implique et explique au contraire que l'identité importe. C'est ce que l'on se propose de montrer dans la prochaine sous-section. Pour cela, il est toutefois nécessaire d'avoir clarifié un point d'importance : si l'identité personnelle n'est qu'une manière de décrire des faits, comment caractériser plus rigoureusement la nature de cette description ?

La littérature récente répond, de manière presque consensuelle²⁸, qu'elle prend la forme d'un récit. Notre identité est ainsi, fondamentalement, une *identité narrative*. Cette

28 On pourra par exemple se référer à C. Mackenzie & K. Atkins (éds.), *Practical Identity and Narrative Agency*, New York, Routledge, 2008 ; H. Lindemann, *Holding and Letting Go*, Oxford, Oxford University Press, 2014 ; A. Kind, *Persons and Personal Identity*, Cambridge, Polity Press, 2015.

notion seule permet en effet de capturer l'idée que les événements de notre vie ne sont pas simplement des choses qui nous arrivent, mais bien des choses que nous nous approprions en les décrivant d'une certaine manière, et ainsi, en leur donnant un sens. Certains épisodes sont ainsi perçus par nous comme périphériques – et sont bientôt oubliés –, tandis que d'autres sont vus comme capitaux ; parmi ces derniers, certains sont tristes ou joyeux, solennels ou légers ; certains marquent une rupture, d'autres sont au contraire le point d'aboutissement de longues entreprises. Même quelqu'un qui ne se soucierait que fort peu de son passé et de son avenir²⁹ donnerait encore par là un sens (insignifiant) aux événements de sa vie. Le récit de notre existence n'est, en ce sens, rien de plus que la description que nous nous en faisons, c'est-à-dire le sens que nous lui donnons – ce qui distingue une vie proprement *personnelle* d'une succession neutre et muette d'événements impersonnels.

Le récit qui donne corps à la notion d'identité narrative peut encore être caractérisé plus précisément par quatre caractères, mis au jour par Hilde Lindemann³⁰. Ce récit est tout d'abord *figuratif* (« *depictive* »), au sens où il représente des faits – même si toute figuration, par exemple picturale, ne constitue pas un récit. Il est ensuite *sélectif*, puisqu'il ne se contente pas d'exhiber un à un tous les faits concernant notre vie, mais en met au contraire certains en exergue tout en reléguant les autres dans l'ombre. Il est par ailleurs *interprétatif*, dans la mesure où toute sa fonction est de donner un certain sens aux faits ainsi mis en lumière. Il est, enfin, *connectif* (« *connective* »), puisqu'il enchâsse tous ces faits dans une trame (sinon une trame unique, au moins plusieurs trames éparses), en les expliquant les uns par les autres, et en leur donnant sens les uns par les autres. On voit ainsi que grâce à cette caractérisation, certes incomplète³¹, le concept d'identité narrative permet de fournir un contenu plus précis à l'idée d'une *description* que nous donnons de notre vie, et donc à la thèse réductionniste selon laquelle ce que nous jugeons être les limites de notre identité, loin de constituer un fait de plein droit, dépend au contraire de la manière dont nous choisissons de caractériser un certain ensemble de faits.

29 Pour reprendre une objection célèbre de Galen Strawson au concept même d'identité narrative. Voir G. Strawson, « Against Narrativity », *Ratio* 17, 2004, p. 428-52.

30 H. Lindemann, *Damaged Identities, Narrative Repair*, Ithaca, Cornell University Press, 2001, p. 11-15.

31 Il faudrait par exemple se demander quel est le rôle d'autrui dans la construction de ma narration, ou encore s'il existe des contraintes (par exemple en termes de correspondance avec la réalité) pour qu'un récit puisse être considéré comme constitutif de mon identité. Résoudre ces questions n'est évidemment pas l'objet de cet article.

S'agit-il bien toutefois du même type de description ? Dans la théorie de Parfit, les *faits* décrits sont de nature physique (e.g. le fait que tel corps mesure 1m81, qu'il soit de sexe masculin, etc.) ou psychologique (e.g. le fait que tel flux de conscience comporte tel souvenir, le fait qu'il contient telle intention, etc.). La *description* consiste alors à caractériser certains de ces faits impersonnels comme faisant partie de ma vie, ou comme n'en faisant pas partie au contraire. On pourra ainsi dire, vers le milieu du spectre complexe, que l'individu examiné « est le même que celui du début du spectre » (ou au contraire qu'il n'est pas le même). S'agit-il bien là du type de description que capture le concept d'identité narrative ? La réponse est assurément positive. Plus précisément, la description ramassée « *a* est identique à *b* » découle de la description plus vaste qui constitue le récit de cet individu, et le résume. Par exemple, en l'absence de tout récit permettant de caractériser l'individu *a* au début du spectre, qui possède l'ensemble de propriétés *P*, il est impossible de juger (sinon de manière purement arbitraire) si un individu *b* vers le milieu du spectre, possédant un ensemble de propriétés *P'*, est ou non identique à *a*. En effet, pour procéder à un tel jugement, il faudrait savoir (i) quelles propriétés dans *P* sont particulièrement importantes pour l'identité de *a* et le caractérisent en propre, et (ii) si ces propriétés se retrouvent encore dans *P'*, ou si au contraire il ne s'y trouve plus que des propriétés qui n'ont aucune importance pour l'identité de *a*. Or en l'absence de récit, (i) manque. Comment savoir si les souvenirs qui importent sont, par exemple, ceux de la douleur d'un accouchement ou plutôt d'un tatouage, si les joies qui importent sont celles de l'obtention d'un diplôme de fin d'études, ou d'avoir appris à pêcher avec un proche ? Le point de savoir ce qui *nous* importe ne peut, par principe, être déjà déterminé indépendamment de notre avis. En l'absence de (i), nous n'avons donc aucun moyen de juger si *b* est ou non identique à *a*, c'est-à-dire aucun moyen de donner une description de la situation au sens du réductionnisme de Parfit. C'est parce que *a* décrit sa vie au moyen d'un récit qu'il donne un sens aux propriétés dans *P* ; et c'est à cette condition seulement qu'il est possible de juger, en se fondant sur ses propriétés, si *b* est ou non identique à *a*. Autrement ce ne sont, précisément, que des propriétés impersonnelles.

On voit donc que le concept de récit permet de donner à ce réductionnisme un sens plus rigoureux. Ayant à présent établi ces prérequis, il devient ainsi possible de montrer pourquoi le réductionnisme de Parfit, loin d'impliquer par lui-même que l'identité n'importe pas, est au contraire une thèse qu'il est nécessaire d'admettre pour pouvoir comprendre que l'identité importe.

III.C. L'identité importe parce que le réductionnisme est vrai

D'après ce qui précède, notre identité n'est pas un fait de plein droit, mais dépend de la manière dont nous décrivons certains faits objectifs (réductionnisme) ; cette description prend plus particulièrement la forme d'une mise en récit de ces faits, par laquelle ils se trouvent agencés dans la forme d'une vie (identité narrative) ; enfin, la manière dont nous décrivons les faits objectifs possède une importance intrinsèque (irréalisme de l'importance). Ici, la thèse de l'identité narrative ne constitue pas elle-même une prémisses autonome, mais précise seulement le sens de la thèse réductionniste. Nous disposons donc de deux prémisses – le réductionnisme et l'irréalisme de l'importance –, desquelles il découle immédiatement que l'identité importe. On se propose dans cette dernière sous-section d'explicitier la signification de cette inférence.

Remarquons tout d'abord que la littérature contemporaine présente une forte tendance à rejeter, au moins implicitement, le réductionnisme de Parfit. On constate en effet dans les débats récents un accord presque général pour dissocier la question de notre identité numérique et littérale de la question de notre *caractérisation* – celle de savoir quelles sont les propriétés qui sont vraiment *nôtres*. Alors que la première serait fixée de manière *toujours déterminée* par des caractères objectifs, comme la continuité dans la possession d'un métabolisme et d'un système immunitaire (disons simplement la persistance d'une *vie animale*), c'est seulement à la seconde que l'identité narrative apporterait une réponse. Celle-ci ne permettrait pas de déterminer où commence et où finit mon identité numérique, mais seulement de savoir « qui je suis vraiment » ou « ce qui me caractérise en tant qu'individu », ces deux questions étant comprises comme indépendantes l'une de l'autre. L'identité narrative ne constituerait ainsi une identité qu'en un sens métaphorique et atténué³². Mais c'est évidemment faire bon marché des arguments de Parfit. C'est, en fait, rejeter le réductionnisme, en faisant de notre identité numérique un fait objectif, déjà constitué, et indépendant de ce que nous nous jugeons être, ainsi que de la manière dont nous choisissons de décrire les faits de notre vie. Toute la signification du réductionnisme de Parfit est en effet de rejeter la dissociation entre la ques-

32 Sur cette dissociation, on pourra consulter E. Olson, *The Human Animal*, Oxford, Oxford University Press, 1997 ; M. Schechtman, *The Constitution of Selves*, Ithaca, Cornell University Press, 1996 ; C. West, « Personal Identity: Practical or Metaphysical », in C. Mackenzie & Kim Atkins (éds.), *op. cit.* Les deux premiers auteurs admettent cette séparation, mais C. West en donne une rare critique.

tion « métaphysique » de ce qu'est réellement ou factuellement notre identité, et la question « pratique » de ce que nous la jugeons être.

Or si ce réductionnisme n'était pas vrai, et si ces deux questions devaient en effet être nettement dissociées, on voit mal pourquoi notre identité importerait. S'il fallait en effet distinguer la question de l'identité à proprement parler de celle de notre simple « caractérisation », de sorte que notre identité serait fixée sans indétermination par des caractères objectifs (tels que la persistance d'une vie animale), alors il serait en effet possible de construire de nombreux cas imaginaires où ce qui importe pour nous, c'est-à-dire ce que notre récit caractérise comme vraiment crucial, comme vraiment *important*, serait préservé alors même que notre identité se serait perdue. On pourrait par exemple concevoir une situation où une personne obsédée par son travail serait clonée tous les ans, justement pour pouvoir mener à bien sa tâche : la continuité de sa vie animale serait ainsi à chaque fois rompue, et son identité littérale perdue, mais cela ne serait d'aucune importance pour cette personne, parce que ce que son récit caractérise comme cruciallement important dans son existence serait préservé. On voit donc que l'identité narrative et l'irréalisme de l'importance *sans* le réductionnisme (qui constituent la position dominante dans la littérature contemporaine) ont le même résultat que le réductionnisme *sans* l'identité narrative et l'irréalisme de l'importance (c'est-à-dire la position de Parfit) : celui de nous obliger à conclure que l'identité n'importe pas.

Il faut au contraire le réductionnisme de Parfit, assorti de l'irréalisme de l'importance et de la notion d'identité narrative, pour comprendre que l'identité importe. Pour le montrer, reconsidérons sous ces nouvelles prémisses l'expérience de pensée envisagée par Parfit aux paragraphes 89-90 de *Reasons and Persons*. Chaque hémisphère de mon cerveau est transféré dans le corps de l'un de mes deux frères jumeaux, de sorte que des personnes ayant une conscience continue avec la mienne se réveillent après l'opération dans deux lits au lieu d'un seul. Selon Parfit, dans ce cas, mon identité est simplement indéterminée, et pourtant ce qui importe est clairement préservé. Mais dire que dans un tel cas mon identité est *de facto* indéterminée, c'est peut-être trop concéder à l'adversaire du réductionnisme. C'est congédier trop vite l'idée que mon identité dépend de la description que je choisis de donner des faits me concernant. Tout se passe, en un mot, comme si Parfit hésitait à suivre jusqu'au bout les implications de sa thèse réductionniste, en rechignant à conclure que, si ce réductionnisme est vrai, alors les jumeaux A et B ont eux-mêmes un mot à dire sur leur identité. Supposons alors qu'après l'opération,

chacun de ces jumeaux développe au fil du temps un récit différent, qui implique dans les deux cas un rapport différent à α , l'individu unique dont les deux hémisphères cérébraux ont été séparés pendant l'opération.

Le jumeau A en vient peu à peu à se souvenir de l'accident auquel l'opération a fait suite comme d'un traumatisme profond. Pour cette raison, il ne veut plus conserver aucun lien avec sa vie antérieure à cet événement. Il veut « aller de l'avant ». Dans ce but, il cesse peu à peu de fréquenter les amis de la personne α . Il cesse même de voir sa femme. Il change de métier, il déménage dans un autre pays, et apprend une nouvelle langue. Il se fait de nouveaux amis et développe à leur contact de nouveaux goûts. α pratiquait le tennis, mais A préfère les échecs ; α votait systématiquement pour le candidat d'un certain parti politique, mais A soutient le bord opposé. L'accident, suivi d'une opération aussi importante et inhabituelle, marque tout simplement une rupture dans sa vie. A considère qu'il n'est pas la même personne qu' α . *Si l'on prend la thèse réductionniste au sérieux, alors il faut conclure que A et α sont numériquement distincts.*

Le jumeau B, à l'inverse, développe une angoisse de plus en plus profonde à l'idée que la trame de sa vie puisse avoir été rompue par l'accident. Pour cette raison, il consacre un soin presque pathologique à ne rien changer dans son existence, et à continuer au contraire à vivre comme le faisait α . Il continue à habiter dans la même maison avec la femme qui était mariée à α (et se trouve pour cette raison ravie que A ait décidé de s'éloigner) ; il continue à fréquenter les mêmes amis, à poursuivre les mêmes projets, à pratiquer les mêmes passe-temps. Il considère qu'il est α , il en est même intimement convaincu. *Si l'on prend la thèse réductionniste au sérieux, alors il faut conclure que B et α sont numériquement identiques.*

Plusieurs leçons peuvent être tirées de cette réinterprétation de l'expérience de pensée de Parfit. À un niveau superficiel, on voit que ses prémisses mêmes lui interdisaient de conclure sans nuance que la situation résultant de l'opération est simplement indéterminée. Si le réductionnisme est vrai, alors il est possible (quoique non nécessaire) que les relations d'identité et de distinction entre A, B et α soient clairement établies. À un niveau plus profond toutefois, une conclusion plus significative se dessine. Cette conclusion est que, si le réductionnisme est vrai (et si l'on admet l'irréalisme de l'importance), alors l'identité importe. En effet, tout le ressort de la conclusion de Parfit était que, lors de la séparation de mon cerveau dans deux corps distincts, quelles que soient les relations d'identité résultantes, cela ne peut rien changer à ce qui importe pour moi,

à savoir la continuité de ma vie psychologique, puisque celle-ci est nécessairement conservée. Or le cas qui vient d'être présenté montre que cette conclusion est erronée. *A* et *B* sont certes psychologiquement continus avec α , mais tous deux ne poursuivent pas pour autant sa vie. Ils ne mènent pas tous les deux ses projets à bien, ils ne prennent pas tous les deux soin de sa famille, ils ne veillent pas tous les deux à sa réussite professionnelle. Ce qui importe aux yeux d' α n'est donc pas préservé dans les deux cas. Seul *B*, qui s'identifie à α et souhaite perpétuer son existence, s'assure de cela. Mais c'est pour autant que *B* est α . C'est bien l'identité qui importe.

On voit donc que si ces prémisses permettent de dériver la conclusion *que* l'identité importe, elles permettent également de comprendre *pourquoi*. Si le réductionnisme est vrai, alors les limites de mon identité sont fixées par la manière dont je décris ma vie et choisis de lui donner sens (c'est précisément pourquoi les limites de l'identité de *B*, mais pas celle de *A*, incluent la vie d' α). Mais si, de surcroît, l'irréalisme de l'importance est vrai lui aussi, alors cette manière dont je choisis de donner un sens aux événements de ma vie détermine ce qui importe ou non. L'identité d'une personne n'est pas quelque chose qui lui arriverait passivement, mais découle au contraire de ce qui importe à ses yeux. C'est parce que c'est une propriété à laquelle elle n'accorde pas d'importance qui se perd, que ce changement qualitatif n'affecte en rien son identité numérique – comme lors du passage d' α à *B*, qui ne sont que deux parties d'une unique existence. C'est à l'inverse parce que tel bouleversement transforme et perd ce qui était crucial à ses yeux que cet événement majeur fait d'elle une autre personne – comme *A* est une autre personne qu' α . Ce qui importe à α , à travers la transplantation de son cerveau, est assurément que sa vie continue. Or cette continuation suppose davantage que le maintien factuel et objectif d'une relation *R*. Elle suppose que *A*, ou *B*, assume activement l'identité d' α , la reprenne à son compte, et entreprenne de la continuer – car continuer une vie n'est rien d'autre que cela. C'est ce rapport actif qui importe. Mais si le réductionnisme est vrai, alors cette *identification* n'est rien d'autre que *l'identité* elle-même. À l'inverse, si le réductionnisme n'était pas vrai, alors ce à quoi nous nous identifions, et de manière indissociable, ce à quoi nous accordons de l'importance, ne détermineraient en rien les limites de notre identité. On voit mal alors pourquoi celle-ci importerait.

En admettant le réductionnisme, Parfit affirmait avec force que notre identité n'est pas un fait objectif dont nous aurions à nous accommoder. En tant que personnes, nous avons un rapport actif à ce que nous sommes. Mais en admettant le réalisme de l'impor-

tance, il semblait en même temps renverser ce rapport en passivité, en posant que seuls importent les événements qui nous arrivent, indépendamment du sens que nous leur donnons. Une telle position peut donner le sentiment frustrant que Parfit manque de suivre jusqu'au bout les conséquences auxquelles auraient pu le porter ses principes audacieux. La troisième section de cet article n'est bien sûr qu'une ébauche ; son ambition est néanmoins de faire entrevoir les possibilités philosophiques considérables que recèle la thèse réductionniste, et qui ne s'épuisent nullement dans la conclusion – certainement pas inévitable – que notre identité n'importe pas.

Bibliographie

- Butler J., *The Analogy of Religion, Natural and Revealed, to the Constitution and Course of Nature*, Londres, George Bell & Sons, 1897.
- Kind A., *Persons and Personal Identity*, Cambridge, Polity Press, 2015.
- Lindemann H., *Damaged Identities, Narrative Repair*, Ithaca, Cornell University Press, 2001.
- Lindemann H., *Holding and Letting Go*, Oxford, Oxford University Press, 2014.
- Mackenzie C. & Atkins K. (éds.), *Practical Identity and Narrative Agency*, New York, Routledge, 2008.
- Olson E., *The Human Animal*, Oxford, Oxford University Press, 1997.
- Parfit D., *Reasons and Persons*, Oxford, Oxford University Press, 1984.
- Parfit D., « The Unimportance of Identity », in R. Matin & J. Barresi (dir.), *Personal Identity*, Malden, Blackwell Publishing, 2003.
- Schechtman M., *The Constitution of Selves*, Ithaca, Cornell University Press, 1996.
- G. Strawson, « Against Narrativity », *Ratio* 17, 2004.
- Thomasson A., « *Metaphysical Disputes and Metalinguistic Negotiation* », *Analytic Philosophy* (vol. 0, n° 0), 2016, p. 1-28.
- West C., « Personal Identity: Practical or Metaphysical », in C. Mackenzie & Kim Atkins (éds.), *Practical Identity and Narrative Agency*, New York, Routledge, 2008.
- Wiggins D., *Identity and Spatio-Temporal Continuity*, Oxford, Basil Blackwell, 1967.
- Williams B., « The Self and the Future », in R. Matin & J. Barresi (dir.), *Personal Identity*, Malden, Blackwell Publishing, 2003.