

Comparer les *qualia*. La thèse Frege/Schlick et ses prolongements

Le projet de Goodman dans *La structure de l'apparence*¹, qui est d'étudier les différentes façons dont le monde sensible peut être construit, se situe dans la lignée des travaux de Bertrand Russell, Jean Nicod et Rudolf Carnap². Il présente cependant l'originalité de proposer une construction des individus sensibles à partir d'un ensemble de qualités élémentaires, les *qualia*. Goodman est en effet l'un des premiers philosophes (après C. I. Lewis, à qui il emprunte l'expression³), à parler des *qualia* à propos des caractères, ou des propriétés purement phénoménales, des présentations sensibles. Or la possibilité même de la construction logique d'un « système des apparences » repose sur la possibilité d'une comparaison entre les *qualia* : il doit être possible de reconnaître l'instanciation d'une même qualité à deux moments différents dans deux expériences d'un même sujet, mais également d'établir des relations entre différentes qualités, relations susceptibles de permettre la construction d'un ordre. Ces comparaisons entre les *qualia* conduisent à l'établissement de ce que l'on nomme aujourd'hui un espace de qualités, ou pour utiliser le vocabulaire de Goodman lui-même, d'une « carte » : « Pour le dire en gros : le problème central consiste à construire pour chaque catégorie de *qualia*, une carte qui assignera à chaque *quale* dans la catégorie une position unique et représentera la similitude relative des *qualia* par une proximité relative dans la position » (p. 240/194).

Je m'interrogerai, dans le présent article, sur les conditions de possibilité de telles comparaisons entre les *qualia*. Je commencerai par rappeler les conditions d'identité proposées par Goodman pour les types qualitatifs, et je soutiendrai que sa conception tout à fait originale de ces conditions d'identité a des conséquences quant à la connaissance que des sujets peuvent acquérir des ressemblances qualitatives. Je discuterai ensuite le lien que l'on peut établir entre les relations de ressemblance et de dissemblance qualitatives d'une part, et les jugements de discrimination des sujets de l'autre, afin de présenter les différentes interprétations possibles de la notion d'espace de qualités.

¹ Cf. (Goodman, 1977 [1951]).

² Cf. (Russell, 1917), (Nicod, 1924), (Carnap, 1928).

³ Cf. (Lewis, 1929).

L'idée d'un espace de qualités

Goodman nomme « *qualia* » les propriétés des choses telles qu'elles nous sont présentées dans l'expérience. Ce terme permet de distinguer entre les propriétés perceptibles que les choses possèdent effectivement, et celles qu'elles nous *semblent* posséder, qu'elles nous *apparaissent* comme possédant : « dire qu'une chose a l'air verte consiste en gros à formuler un énoncé à propos de la qualité présentée, une qualité de couleur d'une présentation de la chose, alors qu'en disant d'une chose qu'elle est verte on formule un énoncé plus complexe à propos de qualités de couleur manifestées par diverses présentations de la chose » (p. 126/96). Goodman ajoute que le vocabulaire des propriétés perceptibles est systématiquement ambigu. Les termes de couleur permettent ainsi dans certains usages de désigner des propriétés des choses colorées, et dans d'autres usages de désigner les présentations des couleurs dans l'expérience : « il est évident que les noms de couleur sont employés de ces deux façons dans le langage ordinaire : dans un cas pour des caractères présentés, que j'appellerai dorénavant des *qualia* ; dans l'autre, pour des propriétés des choses » (*ibid.*)⁴.

Nous nous concentrerons, dans la suite de la discussion, sur le premier usage distingué par Goodman, c'est-à-dire sur l'usage dans lequel les expressions appartenant au vocabulaire sensoriel désignent les qualités ressenties, présentées dans des expériences. Il peut sembler de prime abord parfaitement naturel, et même assez simple, de classer les qualités de nos expériences sensorielles. Il semble évident que le rouge ressemble plus au rose et à l'orange qu'au vert ou même qu'au bleu ; de façon semblable, un arôme de cassis ressemble à un arôme de framboise, mais pas à un arôme d'eucalyptus ni à un arôme de gingembre. Lorsque deux présentations sont subjectivement indiscernables quant à la qualité qu'elles instancient, on peut considérer qu'il existe entre elle une relation d'appariement (*matching*).

Il faut distinguer la notion d'appariement, qui repose sur la connaissance qu'un sujet possède de la façon dont les choses lui apparaissent, de celle d'indiscriminabilité⁵. L'indiscriminabilité entre deux sortes de *stimuli* renvoie au *comportement* du sujet, et il s'agit d'une relation entre des classes de *stimuli*, et non d'une relation entre des présentations particulières. Les *stimuli* appartenant à la classe A sont indiscernables de ceux qui appartiennent à la classe B s'il n'existe pas de différence significative,

⁴ Cf. également (Clark, 2000, chap. 1).

⁵ Cf. (Clark, 1993), pp. 45-50.

statistiquement, dans le traitement cognitif des *stimuli* appartenant à la classe A et dans celui des *stimuli* appartenant à la classe B. L'indiscriminabilité entre deux classes de *stimuli* se manifeste normalement au travers des jugements des sujets, et il existe donc un lien entre l'appariement et l'indiscriminabilité : si deux *stimuli* apparaissent comme indiscernables à un sujet, on peut imaginer que c'est parce qu'ils appartiennent à des classes indiscernables. Cependant, ce n'est pas nécessairement vrai : on peut parfaitement concevoir qu'un sujet soit en réalité capable de discriminer, dans son comportement, des *stimuli* qui lui apparaissent subjectivement indiscernables.

Par ailleurs, les relations d'appariement et d'indiscriminabilité ont des statuts épistémologiques fort différents. Par définition, seul le sujet peut juger si deux présentations sont appariées. En revanche, il n'y a pas d'accès privilégié à l'indiscriminabilité : un observateur peut en théorie déterminer, en étudiant les réactions d'un sujet et le fonctionnement de son système cognitif, si deux classes sont indiscernables ou non. L'indiscriminabilité, contrairement à l'appariement, peut donc être vérifiée en troisième personne.

Peut-on définir une notion d'identité qualitative en utilisant le concept d'appariement ? Suffit-il, autrement dit, que deux présentations soient appariées (indiscernables subjectivement quant à leur qualité), pour qu'on puisse considérer les qualités instanciées comme identiques ? Si c'était le cas, le principe (1) devrait être accepté :

(1) Si a s'apparie à b, alors $a = b$

Goodman soulève une importante difficulté pour le principe (1). Il peut arriver que deux présentations A, B, et C soient telles que A soient indiscernables de B pour le sujet (relativement à une certaine qualité), B de C, mais que A et C n'apparaissent pas comme ayant la même qualité. Dans une telle situation, on a s'apparie à B, B s'apparie à C. Si l'on accepte (1), cela implique que $A = B$ et que $B = C$. Par transitivité de l'identité, on a alors $A = C$, ce qui est contradictoire avec le fait qu'il soit possible de distinguer A de C. Ce point est particulièrement intéressant. Il conduit en effet à penser que l'indiscernabilité subjective de deux *qualia* n'est pas une condition suffisante de leur identité. Cela peut sembler paradoxal : faut-il en effet en conclure que l'identité des apparences des *qualia* ne suffit pas à leur identité *tout court* ? Voici ce qu'en dit Goodman :

Même si deux *qualia* q et r s'apparient exactement, il peut y avoir un troisième *quale* qui s'apparie à l'un mais pas à l'autre. Donc des *qualia* appariés ne sont pas toujours identiques. Mais il y a là quelque chose de paradoxal : puisque les *qualia* sont des individus phénoménaux, on peut difficilement affirmer que des

qualia apparemment identiques peuvent être objectivement distincts. Par exemple, il semble à première vue que des *qualia* de couleur qui ont l'air d'être les mêmes doivent être les mêmes. (...) Devrons-nous nous résoudre à affirmer que l'apparence d'identité n'est pas une condition suffisante pour l'identité des apparences et tenter d'expliquer comment une différence entre des phénomènes peut être non phénoménale ? On peut en vérité rendre compte du fait que des *qualia* appariés sont distincts sans aller derrière les apparences. Il suffit d'admettre que deux *qualia* sont identiques si et seulement s'ils s'apparient exactement aux mêmes *qualia*. Bien que des *qualia* distincts doivent être phénoménalement distincts, dire qu'ils sont phénoménalement distincts ne revient pas à dire qu'ils ne s'apparient pas mais qu'il existe un *quale* auquel l'un s'apparie mais pas l'autre (Goodman, 1977 [1951] p. 196, trad. fr. p. 243).

La conclusion que l'on peut tirer de ce texte, c'est que l'appariement, à l'instant t , entre les *qualia* a et b n'est pas une condition suffisante pour que l'on puisse identifier a à b . Ce n'est que sur les bases d'une comparaison entre a , b , et tous les autres *qualia*, que l'on peut inférer une identité qualitative. Pour le dire autrement, il ne suffit pas qu'un sujet fixe son attention sur deux exemplaires de rouge, à un moment t , et ne puisse les distinguer quant à la couleur pour qu'il puisse conclure que les *qualia* instanciés sont identiques, puisque de toutes petites différences qualitatives se révèlent imperceptibles pour l'attention consciente⁶.

On voit donc que la question de la comparaison des *qualia* est centrale lorsqu'on s'intéresse à la construction d'un espace qualitatif. Pour pouvoir donner un sens aux conditions d'identité des *qualia* tels que Goodman les définit, il faut en effet qu'un sujet puisse décider si tel exemplaire au moment t instancie la même qualité que tel autre exemplaire au moment t' . On ne voit pas, en effet, comment un sujet pourrait au même moment t fixer son attention sur toutes les nuances de couleur du spectre ; or, selon la définition que nous venons de discuter, seule une comparaison faisant intervenir *l'ensemble des qualia* d'un espace donné peut justifier un énoncé d'identité qualitatif.

⁶ Le problème posé par l'existence de différences qualitatives imperceptibles pour l'attention consciente avait été identifié par Bertrand Russell. Au chapitre VI de *Signification et vérité*, celui-ci notait déjà que l'indiscernabilité entre des nuances de couleur n'était pas une relation transitive, et que l'on ne pouvait donc pas se fonder sur cette relation pour définir les conditions d'identité de types qualitatifs. La solution que préconise Russell est exactement identique à celle de Goodman : « La nuance précise de couleur de x peut (...) se définir comme la couleur commune à toutes les tâches y , telles que tout ce qui est indiscernable de x quant à la couleur soit également indiscernable de y quant à la couleur, et réciproquement ; de sorte que chaque tâche est indiscernable à la fois de x ou de y , ou bien ne l'est ni de l'une ni de l'autre » (Russell, 1969, p. 119).

Nous allons donc maintenant nous pencher sur les conditions dans lesquelles des comparaisons entre les *qualia* peuvent être effectuées.

L'inversion de spectre

Les ressemblances entre les aspects de vécus d'expériences ressenties subjectivement se reflètent normalement dans des jugements et dans les comportements qui accompagnent ces jugements : il y a un lien étroit, comme je l'ai souligné plus haut, entre les appariements de *qualia* et l'indiscriminabilité entre des *stimuli*, telle qu'elle se manifeste normalement dans les jugements des sujets. Pour autant, les philosophes ont été fascinés depuis fort longtemps par l'apparente concevabilité d'une dissociation entre les caractères qualitatifs intrinsèques de nos expériences et les jugements de perception. Nous allons décrire maintenant un tel cas imaginaire⁷.

Supposons qu'un beau matin vous vous levez, et qu'à la suite d'une manipulation chirurgicale effectuée par un savant fou pendant la nuit sur votre rétine et votre cortex visuel, vous vous aperceviez que tout le spectre des couleurs s'est inversé : le ciel vous apparaît dorénavant comme étant jaune, les tomates comme étant vertes, l'herbe comme étant rouge, le soleil comme étant bleu, ...etc. Notons immédiatement qu'à cette étape de l'expérience de pensée de l'inversion de spectre, celle-ci est parfaitement détectable au travers de votre comportement. Vous pouvez en effet, une fois l'inversion effectuée, comparer les *qualia* de vos nouvelles expériences avec celles de vos anciennes expériences, du moins pour autant que vous vous souveniez de celles-ci.

Supposons cependant que vous décidiez de ne parler de cette mésaventure à quiconque, et de reprendre le cours de votre vie normale. Au début, il est un peu difficile d'appeler « rouges » des choses qui vous apparaissent vertes, « vertes » des choses qui vous apparaissent rouges, « jaunes » des choses qui vous apparaissent bleues, et « bleues » des choses qui vous apparaissent jaunes. Mais peu à peu, vous vous habituez complètement à votre nouvel état, et votre comportement de catégorisation des couleurs devient parfaitement indiscernable de ce qu'il était auparavant. Personne ne s'aperçoit de quoi que ce soit, et vous reprenez une vie tout à fait normale. Supposons de plus qu'à la suite d'un accident de voiture, vous perdiez complètement le souvenir de votre mésaventure, et uniquement ce souvenir, sans qu'aucune fonction cognitive ne soit atteinte. Votre comportement apparaîtra normal ; vous aurez vous-même l'impression de

⁷ On attribue en général à John Locke la première description d'une situation de spectre inversé. Cf. John Locke, *Essai philosophique concernant l'entendement humain*, II, chap. 32, § 15. Voir, pour une présentation presque exhaustive, (Byrne, 2004).

voir le monde exactement de la même façon que tous vos semblables ; vous n'aurez absolument aucune difficulté pour communiquer de l'information à propos des couleurs : vous vous étiez en effet habitué(e) à appeler « bleu » ce qui vous semblait jaune, et qui apparaît bleu aux sujets normaux, «vert » ce qui vous semblait rouge, etc ... ; et du coup, l'inversion de vos sensations n'est plus détectable à partir de votre comportement linguistique. Mais pourtant, le ciel vous apparaît jaune et l'herbe rouge, malgré toutes vos dénégations !

Si tout cela vous semble concevable, une énigme philosophique s'en suit. Comment pouvez-vous en effet être certain que le scénario que nous venons de décrire ne correspond pas à votre situation cognitive réelle, ou à la situation cognitive réelle de la prochaine personne avec laquelle vous discuterez après avoir lu cet article ? Il semble que rien ne permette de distinguer le comportement extérieur d'un agent au spectre inversé — j'appellerai dorénavant un tel agent un « inversé » — de celui d'un agent normal. Le fait que votre comportement de catégorisation des couleurs vous paraisse normal n'est donc pas un bon argument contre la thèse selon laquelle vous seriez un (ou une) Inversé(e). Par ailleurs, protester du fait que vous savez tout de même bien quelle est la couleur du ciel, et que vous savez que vous le voyez en bleu, n'est pas non plus un bon argument. Selon le scénario que nous avons décrit, c'est le rouge que vous appelez, à votre insu, le bleu. Ou plus précisément, car la terminologie devient ici importante, c'est une sensation de rouge que vous associez au concept d'une surface colorée en bleu. Sans que vous le sachiez, voir le ciel ne vous fait pas le même effet qu'aux êtres humains normaux : cela vous fait l'effet que leur fait, à eux, la vision d'une banane uniformément jaune.

Un tel scénario, dans lequel l'espace qualitatif d'un agent se trouve permuté relativement à celui d'un autre agent, sans que rien ne soit modifié à son comportement, est connu dans la philosophie contemporaine de l'esprit sous l'appellation d'« inversion de spectre ». Or, il existe un lien conceptuel entre la concevabilité de telles situations et la possibilité de comparer les *qualia* non pas simplement à différents moments du vécu d'une unique conscience, mais entre les vécus de conscience de sujets différents. Appelons « inter-personnelles » de telles comparaisons, par exemple entre les *qualia* d'un sujet normal et ceux d'un sujet inversé. Il est bien évident que l'on ne peut parler d'inversion de spectre qu'à condition que cela ait un sens, en premier lieu, de comparer les traits qualitatifs des expériences du sujet normal et avec ceux du sujet inversé. Cela doit donc avoir un sens de soutenir, par exemple, que le sujet inversé a une sensation de rouge en présence d'une pelouse causant chez le sujet normal une sensation de vert ; ou à

l'inverse que les deux sujets ont des expériences de couleur identiques qualitativement en présence d'objets différents.

Or, il ne va absolument pas de soi que de telles comparaisons inter-personnelles soient conceptuellement possibles. Supposons en effet qu'un observateur extérieur soutienne qu'en dépit d'un comportement identique, le sujet normal et le sujet inversé aient deux sensations de couleur différentes en regardant un champ de tournesols. Comment vérifier l'assertion de l'observateur ? Quelles raisons avancer en sa faveur ou en sa défaveur ? Si seule compte la nature intrinsèque de la sensation de couleur des deux sujets, n'est-il pas tout simplement impossible d'avoir accès au même moment à ces deux contenus qualitatifs situés pour ainsi dire dans deux esprits séparés, et donc de les comparer ? C'est en tout cas la conclusion de Frege lorsqu'il discute de cette question dans « La pensée » :

La prairie et ses grenouilles, le soleil qui les éclaire, sont là, peu importe que je les regarde ou non ; mais quand j'ai une impression sensible de vert, elle n'existe que pour moi ; je suis son porteur. Il nous semblerait incongru qu'une douleur, un état d'âme, un désir vagabondent dans le monde indépendamment d'un porteur. Une sensation n'est pas possible sans quelqu'un qui la ressent. Le monde intérieur suppose un individu dont il soit le monde intérieur. (...) Mon compagnon et moi sommes persuadés que nous voyons tous les deux la même prairie ; mais chacun de nous a une impression sensible particulière du vert. J'aperçois une fraise parmi les feuilles vertes du fraisier. Mon compagnon qui est daltonien ne l'y trouve pas. L'impression de couleur qu'il reçoit de la fraise ne se distingue pas de manière sensible de celle qu'il reçoit de la feuille. Mon compagnon voit-il la feuille verte rouge ou voit-il la fraise rouge verte ? Ou bien voit-il l'une et l'autre dans une couleur que je ne connais pas ? *Toutes questions auxquelles on ne peut répondre, elles sont même proprement insensées.* Car le mot « rouge », s'il doit non pas indiquer une propriété des choses, mais caractériser une impression sensible appartenant à ma conscience, n'a d'usage que dans le domaine de ma conscience ; *il est alors impossible de comparer mon impression sensible avec celle d'un autre. Il faudrait pour cela réunir en une seule conscience une impression sensible appartenant à une conscience et une impression sensible appartenant à une autre conscience* (nous soulignons). Et même s'il était possible de faire à la fois disparaître une représentation d'une conscience et émerger une représentation dans une autre conscience, il serait encore toujours impossible de répondre à

la question de l'identité de ces représentations. Il appartient si étroitement à l'essence de chacune de mes représentations d'être le contenu de ma conscience que toute représentation d'un autre homme, en tant que telle, est différente de la mienne. (Frege, 1918-19, trad. fr. pp. 181-182).

Dans ce texte célèbre, Frege s'attaque à ce que j'appellerai, en suivant Shoemaker, la *conception de sens commun*. Selon cette conception, il va de soi que l'on peut comparer les contenus intrinsèques des états qualitatifs de deux sujets.

Or, soutient Frege, il est « proprement insensé » de vouloir opérer de telles comparaisons. Afin de bien saisir l'argument qu'il avance à l'appui de son attaque contre la conception de sens commun, commençons par clarifier la terminologie. Les termes de couleur comme « rouge » peuvent, selon Frege, être employés pour désigner des propriétés radicalement distinctes. En premier lieu, on peut utiliser le mot « rouge » pour « indiquer une propriété des choses » plus précisément une propriété de leurs surfaces — disons, par exemple, la disposition d'une surface à refléter un certain pourcentage de la lumière relativement à chaque longueur d'onde. Mais il existe un autre usage du mot « rouge », dans lequel il caractérise « une impression sensible appartenant à ma conscience ». Dans cet usage, que l'on qualifierait aujourd'hui de « phénoménal » ou de « qualitatif », le terme ne peut s'appliquer correctement qu'à ma propre expérience. Il est du coup impossible de formuler une hypothèse mettant en relation un aspect de ma propre expérience avec un aspect de celle d'un autre sujet, puisque je ne peux penser qu'à *mes propres expériences* en utilisant des concepts phénoménaux.

Soulignons que la thèse de Frege n'a pas une portée métaphysique : il ne nie pas que les autres sujets possèdent des expériences. Ce qu'il refuse, c'est que l'on puisse comparer les expériences de différents sujets entre elles. On trouve une idée très semblable à celle de Frege dans différents travaux de Moritz Schlick. Comme Frege, Schlick considère qu'il n'est pas simplement impossible, mais même dénué de sens de prétendre comparer les *qualia* de différents sujets — les *contenus* de leurs expériences, pour utiliser sa terminologie. Nous appellerons donc dorénavant cette thèse négative la *thèse Frege-Schlick*, en suivant un usage initié par Sidney Shoemaker⁸.

À l'inverse de Frege, Schlick ne se contente cependant pas de formuler une thèse négative : il présente une conception détaillée des états qualitatifs et des concepts permettant de faire référence à ces états, conception qui est supposée expliquer comment la thèse négative n'est pas aussi incompatible avec la conception de sens commun qu'elle

⁸ Cf. (Shoemaker, 1981).

pourrait le sembler de prime abord. Étant donné la grande richesse des réflexions de Schlick sur ces questions, il me paraît utile de les présenter. Je discuterai ensuite de la plausibilité qu'il convient d'accorder à la thèse négative, et des conséquences qu'il faut en tirer relativement au projet goodmanien de construction d'espaces qualitatifs.

Schlick et la conception structuraliste des *qualia*

Il convient de commencer par situer la thèse négative selon laquelle comparer les *qualia* est dénué de sens dans le cadre du projet schlickien de critique de la connaissance par intuition. Contrairement à Bertrand Russell, pour lequel la connaissance directe des contenus phénoménaux est première, et constitue même le fondement de toute connaissance empirique, Schlick affirme très tôt dans son oeuvre qu'il ne peut y avoir de connaissance du contenu purement qualitatif des expériences. Ce n'est cependant pas à Russell qu'il s'oppose, mais plutôt à des auteurs comme Bergson et Husserl. À Bergson, qui dans *La pensée et le mouvant* affirme que c'est l'intuition, conçue comme une « vision directe de l'esprit par l'esprit »⁹, qui constitue la source de connaissances permettant à la philosophie de se différencier des sciences, Schlick répond de façon radicale qu'il n'existe purement et simplement pas de connaissance des contenus d'expériences : « Je crois au contraire, écrit-il, que par l'intuition et la contemplation nous ne pouvons obtenir nulle connaissance de quelque sorte que ce soit, et qu'elle ne constitue donc non seulement pas une méthode pour la science rigoureuse, mais même pas une méthode scientifique *du tout* »¹⁰.

L'argument que Schlick propose en 1913 pour appuyer sa thèse négative est encore un peu sommaire, et possède schématiquement la forme suivante :

Prémisse 1 : Toute connaissance est de nature discursive.

Prémisse 2 : Or, par définition, l'intuition n'est pas discursive, mais immédiate.

Conclusion : « Une connaissance intuitive, une connaissance qui serait simultanément une intuition, serait donc une contradiction dans les termes ».

Cependant, l'idée fondamentale qui nous intéresse dans cet article est déjà présentée très clairement. L'intuition, selon Schlick, s'apparente à une pure passivité, qui ne concerne que ce qui est « donné » ; en revanche, la connaissance suppose toujours deux éléments qui sont *comparés* : l'élément qui est connu, l'objet de la connaissance ; et ce *comme quoi* il est connu, le concept qu'on lui applique au travers d'un jugement. Nous

⁹ Cf. (Bergson, 1985 [1938], p. 27).

¹⁰ Cf. (Schlick, 1913, trad. anglaise p. 144).

dirions aujourd'hui que la connaissance est forcément de nature *propositionnelle* : elle est connaissance *que quelque chose est le cas*, y compris dans les jugements de perception. Même un jugement de reconnaissance dépasse l'intuition, puisqu'il implique la possibilité, justement, de reconnaître l'objet comme tel ou tel.

Dans la suite de ses travaux en philosophie de la connaissance, Schlick va considérablement préciser sa position. Dans « Erleben, Erkennen, Metaphysik », il clarifie ainsi les relations qui existent, selon lui, entre le contenu phénoménal et le langage¹¹. De nouveau, la thèse centrale est négative : le contenu subjectif de nos expériences ne peut pas être communiqué, nous ne pouvons strictement rien en dire. À l'appui de cette thèse, Schlick utilise l'argument bien connu aujourd'hui de l'ineffabilité de l'expérience subjective. Si le contenu de nos expériences subjectives était communicable, on pourrait le transmettre y compris à des agents n'ayant jamais vécu eux-mêmes ces expériences ; or, ce n'est pas le cas, puisqu'on ne peut par exemple pas communiquer à une personne non-voyante l'effet que cela fait de voir du rouge. En l'absence d'expériences visuelles, la personne non-voyante ne peut en effet pas comprendre le sens associé au concept de rouge.

C'est donc encore à l'impossibilité des comparaisons inter-subjectives de *qualia* que nous sommes ramenés. Il n'existe aucune méthode permettant de vérifier si deux sujets possèdent ou non des *qualia* similaires. En conséquence, tenter d'établir des relations de ressemblance entre les *qualia* d'une personne et ceux d'une autre est tout bonnement dénué de signification. Nous ne pouvons ni parler du contenu de l'expérience, ni connaître quoique ce soit à son propos. Cela n'implique pas simplement que nous ne pouvons pas *savoir* si les contenus des expériences de nos semblables sont ou non les mêmes que ceux de nos propres expériences : de façon beaucoup plus radicale, Schlick soutient que l'idée même d'une comparaison des contenus est dénuée de sens.

Il faut insister sur l'importance de l'idée de comparaison. Pour Schlick, comme nous l'avons déjà souligné, toute connaissance suppose une comparaison. Pour pouvoir déterminer si le contenu de l'expérience d'une autre personne est ou non semblable au contenu de ma propre expérience, il faudrait donc qu'il existe une méthode permettant de comparer ces deux contenus d'expérience. Mais même si l'on admet que l'intuition me permet d'entrer directement en contact avec le contenu de *mon* expérience, on voit mal par quel moyen je pourrai entrer également en contact *avec le contenu de l'expérience d'autrui* afin de le comparer avec celui de la mienne. Par définition, un sujet ne peut en

¹¹ Cf. (Schlick, 1926).

effet avoir d'accès privilégié en première personne qu'à ses propres expériences, pas à celles des autres personnes. La proximité avec Frege est frappante, puisque celui-ci écrivait, dans le texte déjà cité plus haut :

Il est (...) impossible de comparer mon impression sensible avec celle d'un autre. Il faudrait pour cela réunir en une seule conscience une impression sensible appartenant à une conscience et une impression sensible appartenant à une autre conscience. Et même s'il était possible de faire à la fois disparaître une représentation d'une conscience et émerger une représentation dans une autre conscience, il serait encore toujours impossible de répondre à la question de l'identité de ces représentations.

Néanmoins, Frege et Schlick tirent des conclusions assez différentes de l'impossibilité de la comparaison inter-personnelle des contenus d'expérience. Pour Frege, elle doit conduire à un rejet de la conception de sens commun selon laquelle les expériences de différentes personnes sont commensurables entre elles. La position de Schlick apparaît plus complexe. Ce qu'il rejette, ce n'est en effet pas l'attitude du sens commun, mais plutôt une certaine interprétation philosophique de cette attitude.

Certes, si l'on considère que le vocabulaire de l'expérience sensorielle qu'utilise l'homme de la rue réfère à des états intuitifs purement subjectifs, alors les contenus de ces états sont incommunicables et incommensurables entre les personnes. Mais l'on peut rejeter l'antécédent de ce conditionnel : on peut, autrement dit, considérer que le vocabulaire de l'expérience ne fait précisément *pas* réellement référence à de tels états subjectifs incommunicables. Le *modus ponens* frégeen se transforme alors en *modus tollens* :

(*Modus ponens* de Frege) Nos expériences consistent en un accès purement subjectif à des contenus intuitionnés ; or, de tels contenus subjectifs sont absolument incommunicables, et incommensurables entre les personnes ; le sens commun a donc tort de croire qu'on peut comparer les *qualia* entre les personnes.

(*Modus tollens* de Schlick) : Le sens commun a raison de croire que l'on peut comparer les contenus d'expérience entre les personnes ; si nos expériences consistent en un accès purement subjectif à des contenus intuitionnés, alors on ne peut pas comparer leurs contenus entre les personnes ; donc, nos

expériences ne consistent pas en un accès privilégié à des contenus subjectifs et incommunicables.

Cette interprétation de la thèse de l'incommensurabilité possède le grand avantage d'être conservatrice vis-à-vis des intuitions de sens commun, qu'il faut non pas rejeter, mais bien plutôt analyser sous un nouveau jour. Schlick propose ainsi une théorie *structuraliste* du vocabulaire des sensations utilisé dans la psychologie populaire. Selon cette théorie, nous pouvons bel et bien comparer nos sensations d'une façon parfaitement douée de sens, à condition de considérer que ces états psychologiques ne sont pas des intuitions purement subjectives de contenus incommunicables, mais plutôt des états dont on peut entièrement caractériser la nature par leur place dans une structure complexe. Voici ce qu'écrit Schlick à ce propos dans *Forme et contenu* :

Il est parfaitement légitime, je pense, de dire que deux individus éprouvent « les mêmes » sensations ou qualités sensorielles, ou bien qu'ils éprouvent des sensations « différentes », dans la mesure où l'on peut réellement tester la vérité ou la fausseté de tels énoncés. Le physiologiste effectue des tests de ce genre pour examiner et comparer les facultés perceptives de différentes personnes. Il découvre par exemple que, quand on leur présente deux couleurs différentes, la plupart des gens montrent une différence dans leurs réponses (par exemple dans leurs déclarations verbales), mais qu'un certain pourcentage d'individus ne peut être amené à réagir différemment face à ces deux couleurs. Le physiologiste appelle ces derniers « daltoniens », et il affirme que la qualité de leurs perceptions sensorielles n'est pas la même que celle des gens dotés d'une vue normale. Il a parfaitement raison d'affirmer cela (...). Ce jugement affirme qu'il existe des structures déterminées chez les individus en question. L'éventail des réactions d'un daltonien diffère de celui de quelqu'un qui voit normalement : il y a une diversité plus grande dans les perceptions d'un individu normal. C'est là évidemment une différence formelle. Voilà tout ce qu'on peut dire, et la proposition « Les qualités sensorielles diffèrent de telle et telle manière dans chacun des deux cas » n'affirme rien d'autre. Le système des couleurs est plus compliqué chez la personne normale que chez un daltonien : les relations internes sont moins simples. C'est là une différence de structure¹².

¹² Cf. (Schlick, 1932, trad. fr. pp. 68-69).

Deux sensations peuvent ainsi être dites instancier le même type qualitatif pour autant qu'elles possèdent la même place dans une structure, c'est-à-dire, d'après le texte que nous venons de citer, pour autant qu'elles jouent exactement le même rôle dans un réseau causal. Nous ne pouvons pas « entrer » dans l'esprit d'une autre personne pour accéder directement aux propriétés intrinsèques de ses sensations — pour autant que cela ait un sens de parler de propriétés intrinsèques des sensations (ce qui n'est pas le cas pour Schlick). En revanche, nous pouvons parfaitement observer le lien d'un type de sensation avec les *stimuli*, avec les comportements ; nous pouvons étudier la physiologie associée à la sensation, ce qui revient à examiner les circonstances physiques dans lesquelles elle peut avoir des occurrences ; enfin, nous pouvons comparer, sur cette base, les sensations de personnes différentes entre elles. Nous pourrions ainsi légitimement considérer comme « les mêmes » deux sensations de couleur déclenchées par des stimulations semblables, chez deux sujets possédant un système normal de perception de la couleur. Ces deux sensations posséderont des rôles causaux similaires, qui se manifesteront en particulier dans des jugements de ressemblance et de discrimination : dans des conditions de perception semblables, les deux sujets considéreront que les mêmes objets se ressemblent ou peuvent être distingués quant à la couleur.

Il ne fait pas de doute que la conception structuraliste des *qualia* est parfaitement cohérente. On peut la qualifier de fonctionnaliste, puisqu'une sensation s'y trouve réduite à n'être rien d'autre qu'un certain rôle dans une structure causale complexe. Est-elle cependant réellement plausible ? Afin de tenter de répondre à cette question, je voudrais examiner comment la thèse Frege/Schlick peut rendre compte de certaines de nos intuitions sur les *qualia*, en particulier des intuitions portant sur la ressemblance et la dissemblance des sensations qui constituent le fondement de l'approche goodmanienne des espaces de qualités.

Selon la conception standard de l'expérience, que l'on peut attribuer sans trop de risque à Goodman, ce sont des propriétés intrinsèques des *tokens* d'expériences qui expliquent les relations d'appariement, de ressemblance, et de dissemblance qui existent entre elles. Nous appellerons *qualia** ces propriétés purement intrinsèques des expériences — à supposer bien entendu qu'elles existent. La thèse Frege/Schlick semble nous obliger ou bien à abandonner cette conception standard, ou bien à renoncer à l'idée de sens commun selon laquelle les qualités des expériences peuvent être comparées entre les personnes. La conception structuraliste renverse, en un certain sens, la conception standard : ce ne sont pas les *qualia** qui expliquent les relations que les expériences entretiennent entre elles, mais au contraire ces relations qui définissent la

nature des *qualia*. Dans ce cadre, les *qualia* ne sont donc pas conçues comme des propriétés intrinsèques, mais comme des propriétés relationnelles. C'est par exemple en vertu de sa place dans une certaine structure, et donc en vertu des relations qui la lient à d'autres expériences, qu'une sensation de couleur possède telle ou telle qualité. Ce n'est pas parce que la sensation de rouge possède un mystérieux *quale**, accessible au seul sujet, qu'il ressemble plus à l'orange qu'au vert, mais c'est au contraire parce que cette sensation de couleur est associée à des jugements de ressemblance qui la rapprochent de l'orange et du rose, mais qui l'éloignent du vert, qu'elle possède une certaine nature qualitative.

Deux façons de construire les espaces de qualités

Goodman lui-même fonde la construction des espaces de qualités sur des jugements d'appariement et de dissemblance. Comme nous l'avons souligné plus haut, de tels jugements sont fondés sur l'introspection : seul le sujet peut déterminer si deux couleurs sont appariées ou lui apparaissent différemment.

En revanche, les comportements de discrimination entre les *stimuli* peuvent être observés en troisième personne et étudiés de façon parfaitement objective. On suppose en général qu'il existe un lien étroit entre les jugements d'appariement, fondés sur l'introspection, et les comportements de discrimination. Pour cette raison, la comparaison intrasubjective entre les *qualia* a paru relativement peu problématique à de nombreux philosophes. Si le rouge de telle tomate et celui de telle Ferrari sont subjectivement indistinguables pour un sujet, celui-ci peut tout simplement en témoigner, au travers d'un jugement de discrimination. Il pourrait sembler, de ce point de vue, qu'il existe une dissymétrie entre la comparaison intra-personnelle et la comparaison inter-personnelle des *qualia*. De fait, les discussions philosophiques se soient beaucoup focalisées sur la seconde. Wittgenstein envisage ainsi la possibilité de l'inversion intrasubjective :

Considérez cette situation : quelqu'un dit « je n'y comprend rien, mais aujourd'hui je vois en rouge tous les objets bleus, et inversement ». Nous répondons : « cela doit vous faire un drôle d'effet (*look queer*) ! » Il dit que c'est bien le cas, et continue en s'étonnant que les charbons ardents apparaissent froids et que le ciel bleu, dans toute sa clarté, apparaisse chaud. Je pense que dans ces circonstances, ou dans d'autres circonstances similaires, nous devrions incliner à dire qu'il a vu en rouge ce que nous voyons en bleu. De plus nous devrions dire que nous savons qu'il veut dire la même chose que nous par

les mots « bleu » et « rouge », puisqu'il les a toujours utilisés comme nous (Wittgenstein, 1968, p. 284).

Que l'on adopte l'approche traditionnelle des *qualia* comme propriétés intrinsèques des expériences (les *qualia** donc, dans notre terminologie) ou l'approche structuraliste qui en fait des propriétés relationnelles, les comparaisons intrasubjectives pourraient sembler posséder une forme de priorité. Si les classe de *stimuli* A et B apparaissent de la même façon au sujet, il ne pourra dans la plupart des cas pas discriminer entre ces classes.

Pourtant, il convient d'insister, à la suite d'Austen Clark, sur la *différence* entre l'appariement et l'indiscriminabilité :

L'indiscriminabilité n'est autre qu'un échec de la discrimination. L'expression ne signifie pas la même chose qu'« être apparié avec » ou qu'« apparaître de la même façon ». Le fait qu'une personne déclare, même après une inspection attentive et scrupuleuse, que x et y « apparaissent identiques » ne suffit pas à justifier l'application du terme. Le sujet peut insister sur le fait que deux *stimuli* lui semblent identiques même s'il peut d'une façon ou d'une autre, dans certains contextes, discriminer entre eux, en détectant une différence (Clark, 1993, p. 47).

Dans l'approche traditionnelle de l'expérience, on considère qu'un sujet possède un accès privilégié aux relations de ressemblance et d'appariement, puisque ces relations dépendent directement des *qualia** instanciées par l'expérience, et que le sujet peut prendre connaissance de ces *qualia** par l'introspection. Mais selon la conception structuraliste, les *qualia* sont définies par leur place dans des structures complexes de relations causales. Il n'y a donc pas de raison que le sujet puisse déterminer de façon incorrigible, sur la base de l'introspection, si deux expériences instancient ou non la même qualité. C'est donc l'étude du comportement de discrimination du sujet, étude qui peut être menée selon une méthode d'investigation scientifique en troisième personne, qui peut permettre de déterminer si des types de *stimuli* peuvent être ou non discriminés.

Nous voyons donc qu'il existe deux méthodes très différentes conduisant à la construction d'un espace de qualités goodmanien. Selon la première méthode, que l'on trouve chez Goodman lui-même, les relations entre les expériences dépendent de leurs propriétés qualitatives intrinsèques, et elles ne sont accessibles que pour le sujet, dans l'introspection. Selon la seconde méthode en revanche, les relations de proximité entre les types d'expérience sont inférées à partir des comportements de discrimination des agents, comportements qui ne sont pas simplement accessibles au sujet par introspection, mais également à des observateurs extérieurs.

Il importe de souligner que si ces méthodes de construction diffèrent, c'est également le cas de leurs résultats, c'est-à-dire des espaces qualitatifs obtenus au terme de la construction. Dans un espace qualitatif construit à l'aide de la *première méthode*, les relations de proximité entre les points de l'espace représentent des relations de ressemblance, subjectivement ressentie, entre les expériences. Dans le vocabulaire technique philosophique, on nomme « internes » de telles relations, car la nature même des expériences serait modifiée si les relations de ressemblance qu'elles entretenaient les unes avec les autres se trouvaient changées. Ainsi, dans un espace de couleurs, le rouge est plus proche du rose que du vert, parce qu'il ressemble plus au rose qu'au vert ; et le fait de plus ressembler au rose qu'au vert fait partie de son essence, en tant que qualité : une sensation de rouge n'aurait pas la même nature si elle ressemblait plus à une sensation de vert qu'à une sensation de rose. En revanche, dans un espace qualitatif construit à l'aide de la *seconde méthode*, les distances représentent la facilité relative d'une discrimination : plus les sensations sont proches dans l'espace qualitatif, plus elles sont difficiles à discriminer. Les relations de ressemblance relient des sensations particulières : on peut dire d'une présentation particulière de rouge qu'elle ressemble plus ou moins à une présentation d'une autre couleur, la limite de la ressemblance étant constituée par l'appariement. En revanche, comme le souligne Austen Clarck, la discriminabilité porte non pas sur des présentations particulières, mais sur des *classes* de *stimuli*. Ce n'est en effet qu'en considérant le comportement d'un sujet dans un grand nombre de situations que l'on peut déterminer s'il est ou non capable de discriminer un *stimulus* de type A d'un *stimulus* de type B. En psychologie par exemple, on utilise souvent des protocoles de choix forcé pour déterminer les relations de discriminabilité : on demande à un sujet d'identifier une certaine cible C présentée en présence d'autres *stimuli*, Si. Si le sujet ne parvient pas à identifier la cible plus souvent qu'au hasard, on peut conclure que C n'est pas discriminable des Si pour le sujet. La discriminabilité se révèle donc être une notion statistique, qui relie des types de *stimuli* et non des occurrences d'expériences¹³.

Mais le point le plus important est ailleurs : il réside dans la différence des domaines respectifs relativement auxquels les relations de ressemblance et de discriminabilité sont définies. Si l'on accepte la thèse Frege/Schlick, la relation de ressemblance ne vaut que relativement à la subjectivité particulière d'un sujet. Chaque espace de qualités devra donc être relativisé à une subjectivité donnée, et les espaces ne

¹³ Cf. (Clark, 1993, pp. 47-48).

pourront pas être comparés entre eux : une telle comparaison n'aurait à première vue aucun sens¹⁴. En revanche, rien n'empêche de comparer les comportements de discrimination de différents sujets, et de faire des statistiques non seulement sur les jugements d'un unique sujet, mais également sur ceux d'un grand nombre de sujets. On voit donc que ce qui pourrait au départ ne sembler être qu'un détail de méthode — construire les espaces de qualités sur la base des jugements de similarité des sujets, ou sur la base de leurs comportements de discrimination — s'avère avoir des conséquences philosophiques cruciales. Si Frege et Schlick ont raison, les *qualia** construits par la première méthode sont incommensurables entre les personnes, ce qui est incompatible avec notre façon habituelle de parler des sensations ; en revanche, les qualités construites selon la seconde méthode peuvent être comparées, mais elles ne correspondent pas à la conception traditionnelle que l'on se fait en général des *qualia*, puisqu'elles ne peuvent pas être considérées comme des propriétés *intrinsèques* des expériences.

On voit donc que l'idée d'une structure de l'apparence sensible, fondée sur une comparaison systématique de ces apparences, pose d'épineux problèmes d'interprétation. Pour cette raison, je me propose maintenant de présenter systématiquement les différentes façons d'interpréter la thèse Frege/Schlick, et éventuellement d'y répondre, dans le cadre de la construction d'espaces de qualités goodmaniens.

Interpréter la structure de l'apparence

Toutes les théories que nous allons présenter et discuter maintenant ont un point commun : elles partagent la prémisse selon laquelle il est essentiel, pour comprendre la nature des qualités sensibles, de partir de nos jugements de ressemblance et/ou de discrimination. Toutes ces théories sont donc compatibles avec l'idée goodmanienne d'un espace de qualités représentant la structure des apparences, dans une modalité sensorielle donnée. On peut toutes les qualifier, en un certain sens, de fonctionnalistes ou de structuralistes, puisque selon elles la nature d'une qualité sensible est reflétée par sa position dans un espace goodmanien. Néanmoins, la façon dont elles conçoivent les relations entre les qualités sensibles permettent de les différencier nettement les unes des autres, comme nous allons le voir.

¹⁴ Voir cependant plus bas la discussion de la théorie de Shoemaker.

La conception structuraliste fondée sur les capacités de discrimination

À un extrême du spectre des positions, on trouve la suggestion selon laquelle les espaces de qualités doivent être construits en se fondant sur les *comportements de discrimination* des sujets. Nous avons vu que l'on peut trouver l'origine de cette suggestion chez Schlick¹⁵. Plus récemment, elle a été développée en détail par Austen Clark¹⁶, ainsi que par David Hilbert, Mark Calderon et Alexander Byrne¹⁷.

Comme nous l'avons souligné plus haut, cette position conduit à réinterpréter le concept traditionnel de *qualia*, et même à inverser la perspective traditionnelle. Une qualité sensorielle n'est plus comprise comme une propriété intrinsèque d'une expérience susceptible d'expliquer ses relations de ressemblance avec les autres expériences ; au contraire, c'est la position d'une qualité dans un espace construit à partir des comportements de discrimination qui en définit la nature. Du point de vue métaphysique, une qualité doit donc être considérée comme une propriété relationnelle, et non comme une propriété intrinsèque.

Il faut souligner qu'une telle position peut s'articuler de façon élégante à une conception représentationnaliste de l'expérience¹⁸. Il est en effet naturel de supposer que les comportements de discrimination des agents se fondent eux-mêmes dans les dispositions des systèmes sensoriels (vision, audition, olfaction, etc...) à grouper les *stimuli* en classes d'équivalence. Ainsi, les comportements de discrimination relativement à deux surfaces S1 et S2 seront d'autant plus semblables que les deux surfaces seront considérées comme équivalentes par le système relativement à une certaine propriété perceptible — par exemple, relativement à la couleur. Or, c'est certainement lorsque le système sensoriel détecte la même propriété qu'il classe deux *stimuli* dans la même catégorie. Austen Clark a défendu de façon détaillée une telle conception des qualités sensorielles, qui articule la conception structuraliste héritée de Goodman au représentationnalisme. Selon lui, les psychophysiciens qui décrivent la structure d'espaces qualitatifs goodmaniens ne font rien d'autre que de décrire des systèmes représentationnels complexes, dont la fonction est de détecter et de traiter les informations disponibles dans l'environnement, et portant sur des propriétés perceptibles — par exemple, sur des propriétés de *réflectance* dans le cas du système de détection des

¹⁵ Voir en particulier (Schlick, 1932).

¹⁶ Voir (Clark, 1993) et (Clark, 2000).

¹⁷ Voir (Hilbert et Calderon, 2000), ainsi que (Byrne et Hilbert, à paraître).

¹⁸ Sur la conception représentationnaliste de l'expérience, voir (Dretske, 1995) et (Tye, 1995).

couleurs¹⁹. En ce sens, il serait incorrect de croire que c'est la position d'une qualité sensible dans un espace qualitatif qui détermine sa nature. Il existe un très grand nombre de relations objectives de ressemblance entre les propriétés détectées par les systèmes sensoriels, et ces relations sont complètement indépendantes de l'existence des systèmes de détection. Mais seules certaines de ces propriétés sont effectivement classifiées et représentées par les systèmes de détection, et c'est de cette sélection que rend compte la structure des espaces de qualités²⁰.

Par ailleurs, nous avons vu également plus haut que partir des comportements de discrimination pour construire les espaces de qualités et définir la nature de ces dernières permettrait de répondre à la thèse Frege/Schlick. Selon Schlick lui-même, ce n'est que si l'on considère les sensations comme des états privés, dont le contenu se révèle au sujet par introspection, que l'on doit rejeter comme dénuée de sens l'idée d'une comparaison inter-subjective des *qualia*. On peut en revanche reconnaître comme légitime la conviction du sens commun selon laquelle les *qualia* peuvent être comparés entre les personnes, à condition d'analyser ces dernières de façon relationnelle, en partant des comportements de discrimination des agents et de la physiologie des systèmes sensoriels de détection.

Quoique la position que nous venons de décrire soit incompatible avec la thèse Frege/Schlick *comprise de façon stricte*, il est important de souligner qu'elle tire toutes les conséquences de cette thèse et de ses motivations. En particulier, elle implique que les situations d'inversion de spectre ne peuvent être conçues sans contradiction. Les propriétés phénoménales d'une sensation doivent en effet entièrement dépendre de sa fonction psychologique, et en particulier de la façon dont elle contribue à la détection des propriétés perceptibles existant objectivement dans l'environnement du sujet. Il n'est donc pas concevable, selon ce point de vue, qu'une sensation voit son rôle psychologique rester constant et son contenu sensoriel se modifier : deux agents cognitifs dont les capacités de discrimination et de représentation sensorielle fonctionnent de la même manière possèdent nécessairement les mêmes sensations subjectives.

La conception structuraliste fondée sur les relations de ressemblance entre sensations privées

La première position structuraliste répond aux difficultés soulevées par la thèse Frege/Schlick en abandonnant la conception traditionnelle des qualités sensibles comme propriétés intrinsèques des expériences. Il reste cependant possible d'adopter un point de

¹⁹ Cf. (Hilbert, 1987).

²⁰ Cf. (Hilbert et Kalderon, 2000).

vue structuraliste tout en conservant la conception traditionnelle. C'est par exemple la position que semble adopter Frege, comme nous l'avons suggéré plus haut²¹. Frege semble en effet considérer que les sensations ne peuvent être classées en types qu'en partant des ressemblances qu'elles entretiennent *pour une subjectivité donnée*. Si les qualités sensibles ne peuvent être comparées entre les personnes, c'est parce qu'il s'agit de propriétés intrinsèques d'expériences privées, qui ne sont accessibles qu'au sujet qui possède les expériences. Frege conserve donc un aspect de la conception naïve de l'expérience, tout en abandonnant un autre aspect de cette conception : il maintient que les qualités sensibles sont des propriétés intrinsèques, mais rejette comme incohérente l'idée selon laquelle ces qualités pourraient être comparées entre les personnes.

La conception de sens commun peut-elle être entièrement réconciliée avec une conception structuraliste fondée sur les relations de ressemblance entre sensations privées ? C'est la thèse que Sidney Shoemaker défend dans une série d'articles importants²². Shoemaker pense pouvoir faire valoir un argument décisif contre la thèse Frege/Schlick, qu'il tire de la concevabilité des expériences d'inversion de spectre. Nous avons déjà insisté sur le fait qu'il existe une incompatibilité entre la thèse Frege/Schlick et la concevabilité d'un agent « inversé ».

Or selon Shoemaker, la conception de sens commun implique la concevabilité de tels agents. Nous n'avons certes pas *immédiatement* d'intuitions très fermes sur les situations d'inversion *inter-subjective*. En revanche, comme le souligne Shoemaker, nous avons la conviction que l'inversion *intra-subjective* est possible, et même parfaitement non problématique. Nous pourrions nous réveiller un matin et ne plus voir les choses qui nous entourent colorées comme avant. Il n'y a non seulement là rien d'inconcevable, mais même un philosophe vérificationniste ne devrait pas contester la possibilité de telles situations, à partir du moment où l'inversion se reflète bien, dans le cas intra-subjectif le plus simple, dans des dispositions comportementales. Nous n'avons pas en général de raisons de douter du témoignage des agents sur la façon dont les objets leur apparaissent : si un agent se comporte exactement comme s'il voyait en vert ce que nous voyons en rouge, pourquoi donc refuserions-nous de considérer qu'il s'agit d'un cas d'inversion intra-subjective, c'est-à-dire d'un cas dans lequel les qualités des expériences du sujet se sont, pour une raison quelconque, modifiées ? Shoemaker cite, à l'appui de

²¹ Cf. (Frege, 1918-1919).

²² Cf. (Shoemaker, 1975), (Shoemaker, 1981) et (Shoemaker, 2006).

cette thèse, un texte de Wittgenstein, qui, décrivant un cas d'inversion intra-subjectif, admet que « nous devrions dire que la personne voir vert ce que je vois rouge »²³.

A partir de ce point de départ susceptible d'être admis par tous, Shoemaker tente de montrer qu'on peut dériver la concevabilité des inversions inter-subjectives. En premier lieu, il soutient que la concevabilité de l'inversion intra-subjective la plus simple implique immédiatement celle de l'inversion inter-subjective. Supposons en effet que Pierre et Jean aient des expériences visuelles similaires en voyant des tomates à l'instant t1, et que Jean subisse une inversion intra-subjective entre t1 et t2. Ne peut-on conclure que Jean est forcément inversé relativement à Pierre ? Si les tomates ont conservé leur apparence visuelle pour Pierre entre les deux moments, mais qu'elles ont changé d'apparence pour Jean, n'est-on pas nécessairement conduit à penser que Pierre et Jean ont bien des expériences distinctes ? Mais si c'est le cas, cela n'implique-t-il pas que Jean voit dorénavant vert ce que Pierre voit rouge, et inversement ?

Comme le souligne Stalnaker, on peut résister à la conclusion que Shoemaker voudrait tirer de ce premier argument²⁴. Il n'est en effet pas évident que les prémisses mêmes de l'argument ne présupposent pas la thèse qui se trouve en question. Certes, si l'on suppose que l'on peut donner un sens à la proposition selon laquelle les sensations de Pierre et de Jean à l'instant t sont *les mêmes*, alors on peut conclure qu'elles sont *devenues différentes* et donc qu'une inversion s'est produite. Mais l'intelligibilité même d'une telle proposition identifiant les contenus qualitatifs d'expériences de sujets distincts est précisément ce que la thèse Frege/Schlick remet en question.

Par ailleurs, la conception structuraliste fondée sur les capacités de discrimination, que nous avons esquissée dans le paragraphe précédent, peut rendre compte de la concevabilité d'au moins *certaines* situations d'inversion. Dans la situation décrite par Wittgenstein, l'inversion est détectable par le comportement. On pourrait être tenté, dès lors, d'adopter un point de vue représentationnel sur la situation, et de soutenir que le système visuel de Jean représente incorrectement les couleurs après l'inversion. Ce point de vue, comme nous l'avons souligné précédemment, est compatible avec une interprétation faible de la thèse Frege/Schlick.

Shoemaker est évidemment conscient de toutes ces difficultés, et c'est la raison pour laquelle il développe et complexifie la situation décrite par Wittgenstein de deux façons. En premier lieu, on peut selon lui passer facilement de la description d'une

²³ Cf. (Wittgenstein 1968, p. 316).

²⁴ Cf. (Stalnaker, 2003).

situation *déTECTABLE* d'inversion à celle d'une situation *INDÉTECTABLE*. Supposons que Jean s'habitue aux nouvelles caractéristiques de son système visuel, et qu'il prenne l'habitude d'appeler vert ce qui lui semble rouge, et rouge ce qui lui semble vert. Supposons de plus qu'il oublie complètement que son système visuel ne fonctionne pas normalement relativement aux couleurs. Dans cette nouvelle situation imaginaire, il utilisera les concepts de couleur exactement comme Pierre, et leurs comportements de discrimination seront devenus absolument indiscernables. Or selon Shoemaker, nous avons *l'intuition* que Jean voit les surfaces rouges en vert et les surfaces vertes en rouge.

À ce stade, un défenseur de la thèse Frege/Schlick fera évidemment remarquer que nos intuitions cessent précisément d'avoir quelque solidité que ce soit lorsque nous les mobilisons dans de telles situations imaginaires. Comment savoir en effet que les sensations de Jean possèdent un contenu qualitatif différent de celles de Pierre, dès lors que nous ne disposons d'aucune méthode de vérification ?

Afin de renforcer nos intuitions en faveur de l'intelligibilité des inversions indétectables, Shoemaker montre que celles-ci peuvent se produire de façon graduelle. Imaginons ainsi que le spectre visuel de Jean, représenté par un disque pour les besoins de l'argument, soit divisé en « tranches » suffisamment nombreuses pour qu'à chaque tranche corresponde une nuance de couleur tout juste discernable visuellement des tranches adjacentes. Supposons qu'une série d'inversions détectables se produisent. À chaque inversion, Jean voit les objets d'une façon légèrement différente : si l'objet était perçu comme ayant la couleur N sur le disque, il est dorénavant perçu comme ayant la couleur adjacente N+1. Comme chaque inversion a un effet relativement mince sur la structure de l'espace de couleurs de Jean, on peut supposer qu'il s'habitue facilement à chaque fois, et qu'il continue donc à utiliser les termes de couleur comme avant, à quelques petites différences détectables près. Pourtant il est clair qu'au bout d'un certain nombre de petites inversions, le spectre tout entier sera permuté : Jean percevra rouge ce que nous percevons vert, et inversement, sans toutefois que son comportement de discrimination diffère du nôtre.

Le cas des inversions graduelles est intéressant, car chaque petite inversion est détectable, alors que leur somme ne l'est plus. Un défenseur de la thèse Frege/Schlick est donc contraint de soutenir que dans cette situation, les sensations de Jean diffèrent du point de vue qualitatif entre chaque petite inversion, mais qu'au bout du compte leurs propriétés qualitatives redeviennent exactement ce qu'elles étaient au départ, en raison de la permutation du spectre.

Nous reviendrons plus bas sur ce dernier argument, qui est certainement le plus élaboré qui ait été avancé en faveur de la concevabilité des cas d'inversion indétectable de spectre. Qu'on soit ou non convaincu que la situation décrite est intelligible, il importe de tenir compte de la position qui en découle selon Shoemaker, car il s'agit d'une conception structuraliste fondée sur les relations de ressemblance entre les sensations privées, mais qui diffère de façon marquée de la théorie de Frege. Shoemaker rejoint Frege sur un point central : seules les relations d'appariement, de ressemblance et de dissemblance entre les expériences vécues subjectivement peuvent nous permettre de saisir les propriétés essentielles des expériences. La concevabilité des situations d'inversion indétectable montre en effet selon lui qu'il faut distinguer l'objet intentionnel d'une expérience sensorielle — ce que l'expérience représente, selon le point de vue représentationnaliste — de son contenu phénoménal, accessible uniquement en première personne. En revanche, il rejette l'idée d'une incommensurabilité entre les expériences d'un sujet donné et celles d'un autre sujet : précisément parce que les situations d'inversion indétectable sont concevables, il doit toujours être possible de comparer les *qualia* entre les personnes.

Afin de concilier ces exigences, qui pourraient sembler contradictoires, Shoemaker propose de définir fonctionnellement non pas les états qualitatifs des sujets, mais simplement les conditions d'appartenance de ces états qualitatifs à des classes d'équivalence. N'étant pas définie fonctionnellement, une propriété qualitative donnée doit selon lui être identifiée à une certaine *réalisation physique*, vraisemblablement cérébrale. Du point de vue du sujet cependant, deux états cérébraux correspondant aux *mêmes* vécus d'expérience, c'est-à-dire à des vécus susceptibles d'être appariés, appartiennent à la même classe d'équivalence. Néanmoins, l'identification des états qualitatifs à des réalisations cérébrales (ou au moins à des classes d'équivalence de telles réalisations) permet de donner un sens aux comparaisons inter-subjectives. On pourra ainsi affirmer que l'état qualitatif correspondant, pour Pierre, à ce que cela fait de voir une tomate rouge, est le même que l'état qualitatif de Jean si, et seulement si (i) la position de l'état de Jean dans l'espace qualitatif défini par les ressemblances entre ses sensations est la même que celle de l'état de Pierre dans l'espace qualitatif défini pour Pierre et (ii) la réalisation cérébrale de cet état est une des réalisations cérébrales possibles pour l'état correspondant de Jean.

Le structuralisme est-il vraiment incompatible avec le sens commun ?

Avant de conclure, je voudrais revenir sur la partie critique de l'oeuvre de Shoemaker vis-à-vis du structuralisme et de la thèse Frege/Schlick comprise en son sens le plus fort. Cet auteur semble, à certains moments, attiré par la thèse Frege/Schlick. Il la rejette pourtant, essentiellement en raison de son incompatibilité avec la conception naïve, de sens commun, que nous sommes censés associer à l'expérience subjective. Selon Shoemaker, le caractère intrinsèque de l'expérience découle *a priori* de la conception naïve. C'est précisément ce que les expériences de pensée d'inversion de spectre ont pour fonction de montrer : si les caractères qualitatifs des expériences étaient définis fonctionnellement, de façon relationnelle, les situations d'inversion indétectable devraient être inconcevables.

Il n'est cependant pas évident que notre conception naïve implique quoi que ce soit quant au caractère relationnel ou intrinsèque des propriétés qualitatives. Nous pourrions en effet multiplier à loisir les exemples de propriétés relationnelles qui sont conçues, dans nos théories naïves, comme des propriétés intrinsèques. Ainsi, nous avons du mal, 375 ans après les travaux de Galilée, à rejeter l'idée selon laquelle un corps est ou bien au repos, ou bien en mouvement, en un sens absolu : le référentiel physique constitué par notre corps recèle en effet une telle importance écologique que les mouvements relatifs à ce référentiel sont représentés par notre système perceptif comme absolus. On pourrait en dire autant d'autres propriétés physiques, comme le poids d'un corps ou la simultanéité de deux événements.

Un argument, comme celui des *qualia* inversés, qui repose sur nos intuitions naïves quant au caractère intrinsèque ou relationnel des propriétés phénoménales, peut donc sembler fragile. Comme le souligne Robert Stalnaker, il serait facile de réfuter la théorie relationnelle de l'espace si l'on faisait appel de façon incontrôlée à nos intuitions physiques naïves²⁵. Supposons en effet que tous les objets présents dans l'univers soient alignés. Bougeons le premier objet vers la droite. Il a évidemment changé de position. Bougeons maintenant le second objet d'une même distance vers la droite, puis le troisième ... le quatrième, et enfin le Nième objet. N'est-il pas évident que tous les objets se soient déplacés vers la droite ? Si c'est le cas, bien entendu, il faut conclure à l'existence d'un espace absolu : les relations spatiales entre les objets sont en effet exactement les mêmes à la fin de l'opération de déplacement qu'à son début, puisque tous les objets ont été, par hypothèse, déplacés d'une même distance vers la droite. Mais l'argument repose évidemment sur une pétition de principe. Un défenseur de la

²⁵ Cf. (Stalnaker, 2003, pp. 226-227).

conception relationnelle de l'espace peut accepter qu'il existe bien un déplacement dans les N premières étapes, puisque l'on peut définir à chaque fois un repère relativement auquel le déplacement peut être décrit ; en revanche, on ne peut conclure que l'ensemble des objets se soient déplacés, sauf à présupposer ce qui est en question, c'est-à-dire l'existence d'un espace absolu. L'intuition selon laquelle l'ensemble des objets se sont déplacés peut facilement être expliquée : elle provient de notre tendance, d'origine perceptive, à représenter les mouvements comme absolus lorsque le repère relativement auquel le mouvement est décrit est particulièrement saillant. Cette intuition n'a du coup guère de poids dans le débat sur la nature absolue ou relationnelle de l'espace.

Lorsqu'on réfléchit à cet argument fallacieux contre la théorie relationnelle de l'espace, l'argument des inversions graduelles présenté plus haut perd de sa force. Rappelons la conclusion de cet argument : il peut exister des différences dans les propriétés phénoménales des expériences d'un sujet même lorsque toutes les relations de ressemblance entre ces expériences sont conservées, c'est-à-dire même lorsque la structure de l'espace qualitatif construit à partir des relations de ressemblance et de dissemblance entre expériences demeure invariante. Mais exactement comme dans le cas de l'argument contre la théorie relationnelle de l'espace, l'argument tire sa force d'une pétition de principe : rien, sinon la supposition de ce qui est en question, à savoir la supposition du caractère intrinsèque des propriétés phénoménales, ne permet de tirer la conclusion à l'issue des différentes étapes du raisonnement.

Conclusion

Nous avons essayé, dans cet article, de présenter le concept goodmanien d'espace qualitatif et de mettre en évidence sa fécondité dans les discussions contemporaines sur les propriétés qualitatives des expériences subjectives. Nous avons présenté les deux grandes interprétations actuellement dominantes de cette notion.

Selon la première interprétation, les espaces de qualités peuvent être construits à partir des jugements de discrimination des sujets. Nous avons insisté sur le fait que cette première interprétation était compatible avec les approches représentationnalistes des propriétés phénoménales. Puisque deux états phénoménaux possédant exactement la même fonction psychologique, c'est-à-dire jouant le même rôle dans les comportements de discriminations des sujets, ont également selon cette approche le même caractère qualitatif, les situations d'inversion indétectable de spectre sont considérées comme métaphysiquement impossibles.

Selon la seconde interprétation, les espaces de qualités sont construits à partir des relations de ressemblance, d'appariement et de dissemblance entre les sensations, telles qu'elles peuvent être évaluées par le sujet *via* l'introspection à un moment donné. La seconde interprétation est compatible la conception naïve de l'expérience, qui conçoit les propriétés phénoménales comme des propriétés intrinsèques. Shoemaker a également montré qu'elle était compatible avec un rejet de la thèse Frege/Schlick de l'incommensurabilité des *qualia* entre les personnes, et donc qu'elle était compatible avec l'intelligibilité des situations d'inversion indétectable de spectre.

Les expériences de pensée d'inversion de spectre jouent ainsi un rôle stratégique lorsqu'il s'agit de choisir entre les deux interprétations que nous avons esquissées. Quoique notre but n'ait pas été de trancher entre les deux interprétations dans le cadre de cet article, nous avons souligné ce qui nous paraît être une faiblesse importante de ces expériences de pensée : elles reposent sur des intuitions issues de notre conception naïve de l'expérience, dont il n'est pas certain qu'elles reflètent correctement la structure ontologique des propriétés phénoménales.

Il donc est probable que les discussions futures sur le concept d'espace qualitatif vont se focaliser sur la question de la comparabilité des expériences à la fois entre les sujets, et à différents moments du temps pour un même sujet, et donc du même coup sur l'intelligibilité même des situations d'inversion de spectre.

Références :

Bergson H. (1985)[1938], *La pensée et le mouvant*, Paris : Presses universitaires de France.

Byrne A. (2004), "Inverted Qualia", Stanford Encyclopedia of Philosophy, <http://plato.stanford.edu/entries/qualia-inverted/>.

Byrne A. et Hilbert D. (à paraître), « Basic sensible qualities and the structure of appearance », à paraître dans *Philosophical Issues*.

Carnap R. (1928), *Der logische Aufbau der Welt*, Berlin : Weltreis-Verlag. Trad. fr. par T. Rivain, *La construction logique du monde*, Paris : Vrin, 2002.

Clark A. (1993), *Sensory Qualities*, Oxford : Clarendon Press, 1993.

Clark A. (2000), *A Theory of Sentience*, Oxford : Oxford University Press.

Dretske F. (1995), *Naturalizing the Mind*, Cambridge, Mass. : MIT Press.

- Frege G. (1918-19), « Der Gedanke », *Beiträge zur Philosophie des deutschen Idealismus*, I, 58-77. Trad. fr. par Claude Imbert, « La pensée », in *Ecrits logiques et philosophiques*, Paris : Seuil, 170-195, 1971.
- Goodman N. (1977)[1951], *The Structure of Appearance*, Dordrecht : Reidel. Trad. fr. par P. Livet, P. Minh, N. Murzili, M. Pavlopoulos, J. B. Rauzy, N. Yunez-Naude, *La structure de l'apparence*, Paris : Vrin, 2004.
- Hilbert D. (1987), *Color and Color Perception. A Study in Anthropocentric Realism*, CSLI Lectures Notes N°9, Stanford : CSLI.
- Hilbert D. et Kalderon M. (2000), « Color and the Inverted Spectrum, » in *Color Perception: Philosophical, Psychological, Artistic and Computational Perspectives*, Steven Davis (dir.), *Vancouver Studies in Cognitive Science*, Vol. IX, New York : Oxford University Press.
- Lewis C. I. (1929), *Mind and the World Order: Outline of a Theory of Knowledge*, .
- Nicod, J. (1924), *La géométrie dans le monde sensible*, Paris : Alcan.
- Russell B. (1969) [, *Signification et vérité*, 1969, Paris : Flammarion.
- Russell B., (1917), “The Relation of *Sense-Data* to Physics”, in *Mysticism and Logic*, Londres et New-York : Routledge, 140-172.
- Schlick M. (1913), « Gibt es intuitive Erkenntnis? », *Vierteljahrsschrift für wissenschaftliche Philosophie und Sociologie*, 37, 472-488. Trad. anglaise in Moritz Schlick, *Philosophical Papers: 1909-1922*, vol. 1, Berlin : Springer, 141-152, 1978.
- Schlick M. (1926), « Erleben, Erkennen, Metaphysik », *Kant-Studien*, 31, 146-158. Trad. anglaise in Moritz Schlick, *Philosophical Papers: 1925-1936*, vol. 2, Dordrecht : Kluwer, 99-111, 1980.
- Schlick M. (1932), « Form and Content. An Introduction to Philosophical Thinking », in Moritz Schlick, *Philosophical Papers: 1925-1936*, vol. 2, Dordrecht : Kluwer, 285-369 , 1980. Trad. fr. par D. Chapuis-Schmitz, *Forme et contenu. Une introduction à la pensée philosophique*, Marseille : Agone, 2003.
- Shoemaker S. (1975), « Functionalism and *Qualia* », *Philosophical Studies*, 27, 291-315.
- Shoemaker S. (1981), « The Inverted Spectrum », *Journal of Philosophy*, 79, 357-381.
- Shoemaker S. (2006), « The Frege—Schlick View », in J. Thomson et A. Byrne (dir.), *Content And Modality: Themes from the Philosophy of Robert Stalnaker*, New York : Oxford University Press, 18-33.
- Stalnaker R. (2003) , « Comparing *Qualia* Accross Persons », in *Ways a World Might Be. Metaphysical and Anti-Metaphysical Essays*, Oxford : Clarendon Press, 219-238.
- Tye M. (1995), *Ten Problems of Consciousness. A Representational Theory of the Phenomenal Mind*, Cambridge, Mass. : MIT Press.
- Wittgenstein L. (1968), « Notes for Lectures on ‘Private Experience’ and ‘Sense Data’ » in R. Rhees (ed.), *The Philosophical Review* 77, 275-320.