

Compositionnalite, composition conceptuelle et combinaison prototypique

Edouard Machery

► **To cite this version:**

Edouard Machery. Compositionnalite, composition conceptuelle et combinaison prototypique. In *cognito: revue internationale francophone en sciences cognitives* éditée par l'association In cognito, Grenoble : In cognito, 2001, 22, pp.5-26. <ijn_00000401>

HAL Id: ijn_00000401

https://jeannicod.ccsd.cnrs.fr/ijn_00000401

Submitted on 1 Oct 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

COMPOSITIONNALITÉ, COMPOSITION CONCEPTUELLE ET COMBINAISON PROTOTYPIQUE

Édouard MACHERY

Université de Paris-Sorbonne, département de philosophie
1, rue Victor Cousin
75230 PARIS cedex 05 France
Mél : edouardmachery@hotmail.com

Résumé

Les concepts sont des représentations mentales (RM) qui sont supposées être activées dans des processus cognitifs de nature variée. Les théories psychologiques de la catégorisation les identifient fréquemment aux prototypes et/ou aux théories mentales. Néanmoins, certains philosophes et psychologues soutiennent que la compositionnalité est incompatible avec ces deux identifications. Cet article discute la première prémisse de cet argument (la compositionnalité est une propriété des concepts). Le premier but de cet article est en effet de clarifier la notion ambiguë de compositionnalité en distinguant la propriété de compositionnalité stricto sensu, la composition des concepts et la combinaison des prototypes et des théories. La composition des concepts et la combinaison des prototypes et des théories sont deux processus mentaux qui produisent des RM complexes à partir d'autres RM, appelées « composants » : dans le premier cas, mais non dans le second, les RM composantes sont des constituants des RM complexes, c'est-à-dire sont nécessairement coactivées quand les complexes le sont. La compositionnalité est une propriété possible de ces processus : un processus est compositionnel s'il en résulte une relation entre certaines propriétés des composants et certaines propriétés des RM complexes. Trois formes différentes de compositionnalité doivent être distinguées, la compositionnalité sémantique, la compositionnalité des conditions de possession et la compositionnalité inverse. Le deuxième but de cet article est de montrer par quelques données empiriques que ces notions décrivent bien la production de concepts complexes d'une part, de prototypes et de théories complexes de l'autre. Finalement, le troisième but de cet article est de fournir un nouvel argument contre l'identification des concepts aux prototypes ou aux théories sur la base de la distinction entre la combinaison et la composition.

Abstract

Concepts are mental representations (MR) supposed to be activated in many types of cognitive processes. Psychological theories of categorization identify them frequently with prototypes and/or mental theories. However, some philosophers and psychologists claim that the compositionality is inconsistent with these identifications. In this article we discuss the premise of that argument (the compositionality is a property of concepts). The first goal of this article is indeed to clarify the ambiguous notion of compositionality by distinguishing the property of compositionality stricto sensu, the composition of concepts and the combination of prototypes and mental theories. The composition of concepts and the combination of prototypes and theories are two mental processes which produce complex MRs out of other MRs, called "components": in the first case, but not in the second, the component MRs are the constituents of the complex MRs, i.e., are necessarily coactivated when the complex MRs are activated. The compositionality is a possible property of these processes: a process is compositional if there is a relation between some properties of the component MRs and some of the complex ones. Three types of compositionality should be distinguished, the semantic compositionality, the compositionality of possession conditions and the reverse compositionality. The second goal of this article is to show on the basis of empirical evidence that these notions really describe the production of complex concepts on the one hand and complex prototypes and theories on the other. Finally, the third goal of this article is to provide a new argument against the identification of concepts with prototypes and theories based on the distinction between composition and combination.

Introduction

La notion de concept désigne une espèce particulière de représentations mentales (RM) stockées en mémoire sémantique et douées de propriétés sémantiques, en particulier d'une extension¹. Selon de nombreux

psychologues (par exemple Rips, 1995), les concepts sont activés dans des tâches cognitives nombreuses et de nature variée. Les concepts généraux interviennent d'abord dans la catégorisation (affirmation de l'appartenance d'une entité à un ensemble ou de l'inclusion d'un ensemble dans un autre²) et les concepts singuliers dans la reconnaissance (identification d'une entité avec une entité antérieurement

¹ Il y a en fait deux notions de concept qu'il est parfois convenu d'appeler la notion psychologique et la notion philosophique. Selon la notion psychologique, les concepts sont une espèce particulière de RM douées de propriétés sémantiques, alors que selon la notion philosophique les concepts sont des entités abstraites exprimées par les RM (Rey, 1983 ; Peacocke, 1992). Il convient toutefois de remarquer que de nombreux philosophes emploient le terme « concept » à la manière des

psychologues (par exemple, Fodor, 1998). J'adopte la notion psychologique.

² Les psychologues distinguent rarement ces deux types de jugements.

rencontrée³. En outre, leurs propriétés déterminent en partie le cours des raisonnements inductifs et déductifs (Osherson *et al.*, 1990). Les concepts sont aussi exprimés par les expressions linguistiques : le nom « chien » exprime le concept CHIEN⁴. Le point essentiel de cette brève caractérisation est le suivant :

(CONCEPT) *Les mêmes RM sont activées dans différentes tâches et activités cognitives. Ces RM sont appelées « concepts ».*

Si l'on devait admettre que des représentations mentales de nature différente sont activées dans ces différentes tâches, alors la notion de concept perdrait beaucoup de son intérêt.

Depuis les travaux de Rosch et de ses collègues, de nombreux psychologues, et particulièrement ceux qui étudient la catégorisation, identifient les concepts avec d'autres espèces de RM. On trouve en particulier les deux assimilations suivantes :

- (A) *Les concepts sont des prototypes : le concept TABLE est le prototype correspondant à la classe des tables.*
- (B) *Les concepts sont des ensembles de croyances, parfois appelés mini-théories (Rips, 1995 ; Hampton, 1997b), qui expliquent certaines propriétés des entités qui appartiennent à leur extension. Le concept ANIMAL chez les enfants de 4 ans est ainsi un ensemble tout à fait spécifique de croyances relatives aux animaux, différent des croyances qui composent le concept ANIMAL chez l'enfant de 10 ans (Carey, 1985 ; Carey et Johnson 2000).*

Même s'il existe différentes théories des prototypes, elles s'accordent pour dire qu'un prototype est une RM qui représente les propriétés les plus fréquentes d'une classe d'objets. Le prototype CHIEN représente les propriétés les plus fréquentes possédées par les chiens, c'est-à-dire leurs propriétés typiques (par exemple, les chiens aboient, ont une queue etc.). Le point essentiel est que ces propriétés ne sont pas des conditions nécessaires de catégorisation dans la classe d'objets représentée : un animal peut être catégorisé comme un chien même s'il n'aboie pas. Généralement, mais non nécessairement, les propriétés représentées ne sont pas relationnelles et sont relativement superficielles : elles concernent l'apparence des objets. Les différentes théories des prototypes caractérisent différemment la manière dont ces propriétés sont représentées. La notion de théorie que l'on trouve dans la littérature psychologique est ambiguë : elle désigne tantôt un ensemble de croyances relatives à un domaine d'entités (par exemple, les états mentaux d'autrui pour la théorie de l'esprit, ou encore les objets en mouvement pour l'ontologie naïve), tantôt une RM qui représente

certaines propriétés d'une classe d'objets (souvent appelée mini-théorie). Dans le premier cas (Gopnik et Meltzoff, 1998), les concepts sont définis par leur place dans cet ensemble de croyances, à la manière dont les termes théoriques scientifiques sont souvent supposés être définis par leur place et leur rôle dans la théorie scientifique à laquelle ils appartiennent. Par exemple, le concept CROYANCE est défini par sa place dans la théorie de l'esprit. Dans ce cas, la référence du concept n'est pas l'objet de la théorie : CROYANCE réfère à un certain type d'états mentaux, à savoir les croyances, alors que la théorie de l'esprit a pour objets les états mentaux en général (croyances, mais aussi désirs, espoirs...). Dans le second cas, les concepts ne sont pas partie d'une théorie, comme CROYANCE peut être partie de la théorie de l'esprit, mais sont eux-mêmes des théories, autrement dit, ils sont identiques à des ensembles de propriétés représentées. Les théories en ce second sens se distinguent des prototypes par les propriétés qu'elles représentent. Les théories contiennent des informations de nature variée à propos de la classe représentée : elles peuvent concerner aussi bien l'apparence des objets de cette classe que leurs propriétés essentielles. Les propriétés représentées peuvent aussi être relationnelles. Les deux points suivants distinguent toutefois plus nettement les mini-théories des prototypes. Les théories peuvent représenter des propriétés des membres atypiques de la classe d'objets représentée. En outre, le degré de variation des objets de la classe (c'est-à-dire leur écart par rapport aux membres typiques) peut être représenté, alors qu'un prototype fait abstraction de cette dimension. Enfin les relations entre les propriétés représentées peuvent aussi être représentées : par exemple, la théorie représentant la classe des oiseaux peut contenir l'information que les oiseaux ont des ailes pour voler au lieu de contenir les deux informations séparées que les oiseaux volent et qu'ils ont des ailes. J'utiliserai le mot « théorie » pour désigner ce second sens. Néanmoins, l'existence de théories en ce second sens n'exclue en rien l'existence de théories dans le premier sens.

Or, selon de nombreux philosophes de l'esprit (en particulier Fodor, 1998), mais aussi selon certains psychologues (Osherson et Smith, 1981 ; Rips, 1995), la compositionnalité est une propriété des concepts incompatible avec cette double assimilation. De manière schématique, leur argument repose sur les deux prémisses suivantes :

- La compositionnalité est une propriété des concepts.
- Cette propriété est incompatible avec la double assimilation (A) et (B).

Je me propose d'analyser la première prémisse de ce raisonnement⁵. En effet, la notion de compositionnalité qui intervient dans la première prémisse est ambiguë : elle correspond à différentes notions sous la plume de différents auteurs. Mon premier but est donc d'éclaircir la notion de compositionnalité en procédant à un certain nombre de distinctions. Je propose ainsi de distinguer la

³ J'admets donc l'existence de concepts d'individus. Par exemple, j'admets l'existence d'un concept de de Gaulle.

⁴ Les noms des concepts sont écrits en majuscules. Il importe de remarquer que le caractère composé d'un nom ne dit rien sur la structure du concept nommé : VIN ROUGE n'est pas nécessairement un concept composé. Les propriétés sont écrites en italiques. Les noms de prototypes sont écrits en majuscules italiques.

⁵ Je ne vais donc pas faire une nouvelle recension des arguments de Fodor, même si je mentionnerai certains d'entre eux. (Pour une évaluation des arguments de Fodor, cf. Horwich, 1998, 1999 ; Horgan, 1998 ; Millikan, 1998 ; Peacocke, 2000).

propriété de compositionnalité *stricto sensu*, la composition des concepts et la combinaison des prototypes et des théories. Mon deuxième but est de montrer que ces notions sont utiles pour décrire le processus de production de RM complexes : en particulier, les concepts se composent tandis que les prototypes se combinent. Finalement, ces deux résultats me permettront de reformuler l'argument précédent, en ayant éliminé toute ambiguïté et en ayant montré que les notions proposées s'accordent avec les données empiriques.

Conformément à ce projet, les deux premières parties de cet article sont essentiellement théoriques. J'analyse d'abord la notion de compositionnalité puis, dans un deuxième temps, l'idée selon laquelle les concepts se composent en RM conceptuelles complexes. La troisième et la quatrième parties sont plus empiriques : elles examinent différentes données factuelles qui permettent d'affirmer que de fait les concepts se composent et possèdent la propriété de compositionnalité. La cinquième partie est de nouveau théorique : elle analyse la notion de combinaison. La sixième examine les données empiriques actuelles relatives à la combinaison des prototypes et des théories. Enfin, dans la dernière partie, je tire les conséquences de ces analyses et arguments empiriques en rejetant les assimilations (A) et (B).

Précisons dès maintenant que rejeter ces assimilations, ce n'est rejeter ni la notion de prototype, ni la notion de théorie, mais c'est simplement reconnaître que les concepts forment une espèce distincte de RM, irréductible à ces deux types de structures cognitives.

1. Comment caractériser la compositionnalité ?

Je caractérise d'abord la notion de compositionnalité de manière abstraite, sans référence particulière aux concepts. J'applique ensuite nos analyses aux concepts et au système conceptuel.

La compositionnalité concerne des entités douées de propriétés sémantiques (des symboles) qui, d'une manière ou d'une autre, peuvent s'associer pour produire d'autres entités de la même espèce : c'est le cas par exemple des morphèmes des langues naturelles⁶. Il y a différents processus par lesquels un symbole d'une espèce donnée est produit à partir d'autres symboles de cette même espèce. « R » désigne la classe ou le type de ces processus. Par exemple, la concaténation des mots au sein de phrases selon des règles syntaxiques est un processus de type R. De même, la composition des concepts (cf. plus loin) permet de produire une RM conceptuelle à partir d'autres RM conceptuelles. La composition des concepts est donc du type R (c'est aussi le cas de la combinaison des prototypes, cf. plus loin). Appelons « complexes » les symboles qui résultent des processus de type R et « composants » ceux auxquels ces processus sont appliqués. Les termes « composant » et « complexe » n'indiquent rien sur la nature de la relation entre les composants et les complexes : comme nous le

⁶ Les différentes espèces de RM sont bien sûr des espèces de symboles.

verrons plus loin, on peut distinguer différents processus de type R en fonction des relations entre composants et complexes qui en résultent. Ceci posé, il est possible de caractériser ainsi la compositionnalité :

Définition générale de la compositionnalité — De manière tout à fait générale, on parle de compositionnalité pour désigner une relation entre les propriétés des composants et celles des complexes. Différentes notions de compositionnalité correspondent aux propriétés prises en compte.

Distinguons les trois formes suivantes.

1.1. Compositionnalité sémantique (CS)

La notion de compositionnalité porte *stricto sensu* sur les relations entre les propriétés sémantiques des complexes (leur *contenu*⁷) et celles des composants : le complexe AB est CS-compositionnel si et seulement si son contenu est une fonction du contenu des composants A et B et de la règle de composition employée.

1.1.1. Application de cette première notion aux concepts

(CS) concerne alors les relations entre les propriétés sémantiques des représentations conceptuelles complexes⁸ et celles des concepts composants : le concept AB est CS-compositionnel si et seulement si son contenu est une fonction du contenu des concepts composants A et B et de la règle de composition employée⁹. Par exemple, le concept CHIEN NOIR est sémantiquement compositionnel si et seulement si a/ CHIEN NOIR est complexe, c'est-à-dire obtenu par un processus de type R et b/ son contenu (par exemple, son extension) est une fonction du contenu du concept CHIEN (par exemple, de l'extension de CHIEN), du contenu de NOIR et de la manière dont ces deux concepts sont composés.

Le système conceptuel est CS-compositionnel si et seulement si toute RM conceptuelle complexe est CS-compositionnelle¹⁰.

⁷ Le terme « contenu » a souvent un sens différent en philosophie et en psychologie : dans son acception philosophique, il désigne les propriétés sémantiques des RM, alors que dans son acception psychologique, il désigne les connaissances qui constituent une structure de connaissance. J'adopte ici le sens philosophique de ce terme.

⁸ J'admets parmi les RM conceptuelles complexes les concepts complexes et les attitudes propositionnelles, en particulier les croyances.

⁹ La compositionnalité ainsi caractérisée est une propriété des RM conceptuelles complexes. Par extension, c'est une propriété des systèmes conceptuels. Soulignons qu'en logique, la compositionnalité est une propriété de la sémantique, c'est-à-dire de l'interprétation des formules d'un langage dont la syntaxe est indépendamment spécifiée (Hodges, 1998). Pareillement, dire qu'elle est une propriété d'un langage, c'est dire qu'elle est une propriété de la sémantique de ce langage.

¹⁰ Un nombre fini d'exceptions peuvent être admis : les langues naturelles admettent ainsi un nombre fini d'exceptions évidentes à la compositionnalité sémantique, à savoir les idiomes. Par exemple, le sens de l'expression « faire long feu » n'est pas déterminé par le sens de « feu ». De même, « Saint

« $\mu(x)$ » désigne les propriétés sémantiques de x et α est la règle de composition du concept complexe AB , composé de A et de B selon α . Le système conceptuel est compositionnel en ce premier sens si et seulement si il existe une fonction f telle que l'on ait pour tout A , B et α :

$$(1) \quad \mu(AB) = f(\alpha, \mu(A), \mu(B))$$

1.1.2. Commentaires

1. Si un système de symboles est sémantiquement compositionnel, (CS) s'applique **de manière récursive** jusqu'à ce que l'on atteigne des symboles non composés (atomiques). Par exemple, si les parenthèses indiquent l'ordre d'application d'un processus de type R , on a pour un complexe $((AB)C)$:

$$(2) \quad \mu((AB)C) = f(\alpha_2, \mu(AB), \mu(C)) = f(\alpha_2, f(\alpha_1, \mu(A), \mu(B)), \mu(C)).$$

2. Les langages habituels de la logique des propositions sont un exemple simple de systèmes de symboles satisfaisant la compositionnalité sémantique. En effet, soit un langage L dont les constantes logiques sont la conjonction et la négation. La syntaxe et la sémantique de ce langage sont habituelles.

La sémantique de L est compositionnelle puisqu'il existe une fonction f telle que la valeur de vérité d'une formule complexe quelconque est déterminée par f à partir de la valeur de vérité de ses sous-formules immédiates et de la manière dont celles-ci sont composées :

$$(3) \quad f = \{((\neg, 0), 1), ((\neg, 1), 0), ((\wedge, 1, 1), 1), ((\wedge, 1, 0), 0), ((\wedge, 0, 1), 0), ((\wedge, 0, 0), 0)\}.$$

Par exemple, la valeur de vérité de $(\neg(p \wedge q))$ si $V(p)=1$ et $V(q)=0$ est 1. Or $f(\neg, V((p \wedge q))) = f(\neg, f(\wedge, V(p), V(q))) = f(\neg, f(\wedge, 1, 0)) = f(\neg, 0) = 1$. En conséquence, la valeur de vérité de cette formule est une fonction de la valeur de vérité de ses composants et de la manière dont ceux-ci sont composés, et cela de manière récursive.

3. L'application de (CS) aux concepts suppose que l'on définisse de manière indépendante ce qu'est un composant de RM conceptuelle complexe et quelles sont les règles de composition¹¹. Par exemple, il faut décider si la conjonction de coordination « et » exprime ou non un concept : selon la réponse à cette question, la fonction f déterminant le contenu d'un concept complexe obtenu à partir de CHAT et de NOIR par conjonction n'aura pas le même nombre d'arguments (trois ou quatre) ; pareillement, le mode de combinaison ne sera pas le même (concaténation ou composition). Ainsi, si « et » n'exprime pas un concept, mais le mode de composition des concepts exprimés par « chat » et « noir », la fonction f aura trois arguments, à savoir deux concepts, CHAT et NOIR, et la conjonction comme mode de composition ; au contraire si « et » exprime le concept ET, la fonction aura quatre arguments, trois

concepts, CHAT, NOIR, ET, et un mode de composition, peut-être la concaténation. Plus généralement, qu'elle soit appliquée ou non aux concepts, la compositionnalité sémantique n'a donc de sens que relativement à une détermination indépendante des entités composées et des entités complexes.

4. Les propriétés sémantiques d'une RM conceptuelle complexe doivent être en outre fonction de ses **seuls** composants et de leur mode de composition : elles ne peuvent dépendre d'éléments qui ne la composent pas.

5. La **notion de contenu** employée pour caractériser (CS) permet d'éviter une définition trop étroite de la compositionnalité sémantique : une propriété sémantique donnée d'une RM conceptuelle complexe est fonction des propriétés sémantiques *de même nature* que possèdent les concepts composants et de leur mode de composition. Cette définition est trop étroite puisque des propriétés sémantiques d'une certaine nature peuvent être fonction de propriétés sémantiques d'une autre nature. Par exemple, la valeur de vérité des croyances n'est pas fonction de la valeur de vérité des concepts qui les composent : un concept n'a pas de valeur de vérité. En revanche, la notion de contenu n'exige nullement que les propriétés sémantiques des RM conceptuelles complexes soient de même nature que les propriétés sémantiques des concepts composants. En outre, cette notion s'applique aussi bien aux sémantiques extensionnelles qu'aux sémantiques intensionnelles, puisqu'elle ne précise pas la nature des propriétés sémantiques envisagées. Selon le premier type de sémantiques (qui est, par exemple, celui de la logique des prédicats du premier ordre), les symboles sont doués uniquement d'une référence ou d'une extension, alors que les secondes attribuent aux symboles d'autres propriétés sémantiques en sus de leur référence (généralement appelées « sens » ou « intensions »).

6. Un résultat important, dont je me servirai à plusieurs reprises par la suite, doit être mentionné maintenant. Ce résultat correspond aux deux questions suivantes :

a/ Étant donné la syntaxe d'un langage, existe-t-il une sémantique pour ce langage qui soit compositionnelle ? La réponse est toujours positive, d'après un résultat de Zadorzny (Zadorzny, 1994), Janssen (Janssen, 1997) et Hodges (Hodges, 1998) : un langage doué d'une grammaire récursive admet toujours une sémantique compositionnelle¹². C'est le cas par exemple des langages faits pour l'indépendance (logique IF) développés par Hintikka et son école, contrairement aux affirmations initiales de Hintikka. Autrement dit, si l'on ne contraint pas la nature des propriétés sémantiques utilisées dans l'interprétation du langage, la compositionnalité est une propriété triviale du langage.

b/ Étant donné la syntaxe d'un langage, une sémantique donnée est-elle compositionnelle ? La réponse n'est pas toujours positive. C'est le cas de la sémantique des jeux proposée par Hintikka pour les langages IF : en n'admettant qu'un certain type de propriétés sémantiques dans l'interprétation des formules la sémantique proposée n'est pas compositionnelle (Hodges, 1998). De manière générale,

Empire Romain Germanique » désigne une entité qui n'a peut-être jamais été sainte !

¹¹ Cette remarque vaut pour toutes les espèces de symboles auxquels des processus de type R s'appliquent. Elle s'applique même aux langues naturelles si l'on admet que les composants ultimes ne sont pas les mots, mais plutôt les morphèmes. Il faut alors déterminer quels sont les morphèmes d'une langue — ce qui est un travail empirique difficile.

¹² Ce résultat affirme l'existence d'une sémantique compositionnelle sans déterminer laquelle.

si l'on restreint le type d'entités auxquelles la sémantique d'un langage formel peut faire appel, certains langages n'ont pas de sémantique compositionnelle¹³.

1.2. La compositionnalité des conditions de possession (CCP)

La notion de compositionnalité peut être étendue aux conditions de possession des entités douées de propriétés sémantiques, c'est-à-dire aux conditions que doit satisfaire un individu pour posséder effectivement ces entités. Par exemple, les conditions de possession d'un mot d'une langue naturelle sont les capacités qu'un individu doit posséder pour être un locuteur compétent relativement à ce mot. Un complexe AB est CCP-compositionnel si et seulement si pour posséder ce complexe (satisfaire ses conditions de possession), il suffit de posséder ses composants A et B (satisfaire les conditions de possession des composants) et de savoir comment les composer¹⁴.

1.2.1. Application de cette deuxième notion aux concepts

J'étends donc la compositionnalité aux conditions de possession des concepts. Un concept complexe AB est CCP-compositionnel si et seulement si pour posséder AB (satisfaire ses conditions de possession), il **suffit** de posséder ses composants A et B (satisfaire les conditions de possession des concepts composants) et de savoir comment les composer. Par exemple, si les concepts sont compositionnels en ce deuxième sens, quiconque possède les concepts CHAT et NOIR, et sait composer par conjonction deux concepts, possède par là même la capacité d'avoir le concept CHAT NOIR.

Le système conceptuel est CCP-compositionnel si et seulement si toute RM conceptuelle complexe est CCP-compositionnelle (*modulo* un nombre fini d'exceptions).

« Sat(A,x) » signifie que x satisfait les conditions de possession de A et « Sat(α ,x) » que x sait composer les concepts selon la règle α . Le système conceptuel est compositionnel en ce deuxième sens si et seulement si on a pour tout A, B et α tels qu'il est possible de composer A et B selon α :

$$(4) \text{ Sat}(A,x) \ \& \ \text{Sat}(B,x) \ \& \ \text{Sat}(\alpha,x) \ \rightarrow \ \text{Sat}(AB,x)$$

1.2.2. Commentaires

1. Une propriété p est **constitutive** d'un concept Γ si et seulement si une RM est le concept Γ à condition de posséder la propriété p .

Une propriété peut être constitutive des concepts en général, d'un certain type de concepts, par exemple des

¹³ Je reviens plus loin sur la signification de ce résultat pour la théorie des concepts.

¹⁴ Une précision s'impose (Fodor et Lepore, 2001). Dire qu'il est suffisant pour posséder un complexe de posséder les composants et de savoir comment les composer signifie que l'on a la compétence pour posséder le complexe. Bien sûr, la possession d'une compétence ne garantit nullement que nous soyons en mesure de l'appliquer (performance). En particulier, la production effective de certains complexes peut être trop coûteuse en mémoire de travail.

concepts mathématiques, ou bien d'un concept singulier, par exemple NOIR. Quelle est la relation entre les propriétés constitutives et les conditions de possession ? Premièrement, la distinction entre les propriétés constitutives et les conditions de possession correspond à deux approches distinctes de l'individuation des concepts. On cherche à déterminer les propriétés constitutives quand on caractérise les concepts en termes de conditions d'individuation : il s'agit alors de répondre à la question « qu'est-ce que tel concept (ou tel type de concepts) ? ». En revanche, on cherche à déterminer les conditions de possession si on considère que caractériser les (un type de) concepts, c'est répondre à la question suivante : qu'est-ce que c'est que d'avoir tel concept (ou tel type de concept) ?¹⁵ Deuxièmement, les propriétés constitutives des concepts déterminent certaines de leurs conditions de possession : on peut donc toujours passer des premières aux secondes. En effet, il est possible d'associer à chaque propriété constitutive p du concept A, la condition de possession suivante : un individu x possède A si et seulement s'il possède une RM qui a p pour propriété constitutive. Si *avoir pour composants les concepts ADULTE et NON MARIE* est une propriété constitutive du concept CELIBATAIRE, alors CELIBATAIRE a pour condition de possession la possession du concept complexe ADULTE NON MARIE. L'inverse n'est sans doute pas vrai : les conditions de possession des concepts ne correspondent pas toujours à leurs propriétés constitutives. Admettons, par exemple, que ROUGE ait pour condition de possession la capacité à reconnaître la couleur rouge : cette condition de possession ne correspond pas à une propriété que doit avoir une RM pour être le concept ROUGE, c'est-à-dire à une propriété constitutive du concept ROUGE.

Ceci posé, (CCP) concerne toutes les conditions de possession, qu'elles soient ou non dérivées de propriétés constitutives et que celles-ci soient ou non sémantiques. C'est ce dernier point qui distingue cette analyse de (CS).

2. Quelles sont les conséquences de cette propriété pour la nature des conditions de possession ? Puisqu'il suffit pour avoir le concept AB d'avoir les concepts A et B et de savoir les composer d'une certaine manière, les conditions de possession de AB sont réductibles aux conditions de possession de A et de B : précisément, si x est une condition de possession de AB, alors x est une condition de possession de A, ou x est une condition de possession de B ou x résulte de l'association de conditions de possession de A et de B. Sans quoi, un individu pourrait satisfaire les conditions de possession de A et de B tout en sachant comment les composer sans satisfaire les conditions de possession de AB. Par exemple, si la capacité à reconnaître la couleur noire est une condition de possession du concept CHAT NOIR, alors c'est une condition de possession du concept CHAT ou du concept NOIR ou bien cette capacité dérive des conditions de possession de ces deux concepts. On peut dire que **les concepts complexes héritent intégralement leurs conditions de possession de leurs composants**.

¹⁵ Comme le fait par exemple Peacocke (1992).

1.3. Compositionnalité inverse (CI)

Un complexe AB est CI-compositionnel si et seulement si la satisfaction des conditions de possession des composants A et B est une condition nécessaire de la satisfaction des conditions de possession du complexe AB¹⁶. Dire qu'il est nécessaire de satisfaire les conditions de possession des concepts composants pour satisfaire les conditions de possession des concepts complexes, c'est dire que les conditions de possession des composants sont des conditions de possession des complexes. On peut dire alors que **les composants apportent leurs conditions de possession aux complexes**.

1.3.1. Application aux concepts (Fodor, 1999a, b)

Un concept AB est CI-compositionnel si et seulement si la satisfaction des conditions de possession des concepts composants A et B est une **condition nécessaire** de la satisfaction des conditions de possession du complexe AB, donc si et seulement si les conditions de possession des concepts composants sont des conditions de possession des concepts complexes. Le système conceptuel est CI-compositionnel si et seulement si toute RM conceptuelle complexe est CI-compositionnelle (*modulo* un nombre fini d'exceptions).

Le système conceptuel est compositionnel en ce troisième sens si et seulement si on a pour tout A et B :

$$(5) \text{ Sat}(AB,x) \rightarrow \text{Sat}(A,x) \& \text{Sat}(B,x)$$

Le concept CHAT NOIR est compositionnel en ce troisième sens si et seulement si quiconque satisfait les conditions de possession de CHAT NOIR satisfait aussi les conditions de possession de CHAT et de NOIR.

1.3.2. Commentaires

1. Si l'on admet à la fois (CCP) et (CI), on a donc :

$$(6) \text{ Sat}(AB,x) \leftrightarrow \text{Sat}(A,x) \& \text{Sat}(B,x) \& \text{Sat}(\alpha,x).$$

2. *Quelles sont les conséquences de cette propriété pour la nature des conditions de possession des concepts ?* Si les concepts satisfont ce troisième type de compositionnalité, x est une condition de possession de possession d'un concept F seulement si c'est une condition de possession de tous les concepts dont x est un composant. Sans quoi, un individu pourrait posséder le concept complexe sans satisfaire les conditions de possession des concepts composants. Par exemple, la capacité de reconnaître la couleur noir est une condition de possession du concept NOIR seulement si c'est aussi une condition de possession du concept complexe CHIEN NOIR (et de CHAT NOIR...). Ce point a une conséquence importante : **si x n'est pas une condition de possession des complexes dans lesquels A figure, alors si (CI) est vrai des concepts, x n'est pas une condition de possession de A**. En conséquence, si les concepts satisfont (CI), ce principe fournit une contrainte rigoureuse sur les conditions de possession possibles des concepts.

3. En conséquence de quoi, si l'on admet (CI) et (CCP), x est une condition de possession d'un concept

¹⁶ Je présente plus loin des processus de type R qui n'impliquent pas (CI) : c'est en particulier le cas de la combinaison des prototypes.

complexe AB si et seulement si x est une condition de possession de A ou de B ou peut être dérivée des conditions de possession de ces deux concepts. Autrement dit, les concepts complexes héritent toutes leurs conditions de possession des concepts composants et les concepts composants apportent toutes leurs conditions de possession aux concepts complexes.

4. « cons(A,x) » signifie que la propriété x est une propriété constitutive du concept A. Quel est le lien entre (CI) et le principe suivant :

$$(7) \text{ Cons}(A,x) \rightarrow \text{Cons}(AB,x) ?$$

De prime abord, ce principe semble expliquer (CI)¹⁷ : un individu qui satisfait les conditions de possession d'un concept complexe satisfait les conditions de possession des concepts composants, parce que les propriétés des concepts composants sont aussi des propriétés des concepts complexes. Néanmoins, sous cette forme, ce principe doit être rejeté, car les propriétés sémantiques ne le satisfont pas. Même si les propriétés sémantiques des concepts complexes sont une fonction des propriétés sémantiques des concepts composants selon (CS), les propriétés sémantiques des concepts composants sont différentes des propriétés sémantiques des concepts complexes (et inversement). Par exemple, CHAT a pour extension la classe des chats, alors que CHAT BOTTE a pour extension la classe des chats bottés. CHAT BOTTE n'a donc pas la propriété d'avoir pour extension la classe des chats et, inversement, CHAT n'a pas pour propriété d'avoir pour extension la classe des chats bottés¹⁸.

En résumé, j'ai distingué dans cette première partie trois types de compositionnalité, la compositionnalité sémantique, la compositionnalité des conditions de possession, la compositionnalité inverse.

2. Notion de composition des concepts

(CONST)

α/ Les concepts se composent, selon des règles de composition, pour former d'autres RM conceptuelles, en particulier des croyances et des concepts.

β/ Les concepts composés sont des constituants des concepts complexes.

D'après ce principe, un concept complexe, par exemple CHAT NOIR, résulte de la composition de concepts plus simples, CHAT et NOIR, conformément à une règle de composition (α). Parmi ces règles, on trouve les fonctions booléennes, comme la conjonction ou la disjonction, mais d'autres règles sont aussi possibles.

¹⁷ Je présente plus loin une explication de (CI) qui ne repose pas sur ce principe.

¹⁸ Notons l'analogie avec la relation physique partie/tout. Les propriétés des parties, par exemple d'un bras, ne sont pas nécessairement celles du tout, par exemple d'un corps. En revanche, les conditions de possession d'un bras sont aussi des conditions de possession d'un corps intègre. On ne peut posséder un corps intègre sans posséder ses deux bras intègres.

En outre, CHAT et NOIR sont des constituants de CHAT NOIR (β)¹⁹.

(CONST) n'affirme pas une propriété analytique des concepts, comme si le terme « concept » était réservé pour les RM satisfaisant ce principe. (CONST) est au contraire une thèse empirique et contingente : s'il existe des concepts, c'est-à-dire s'il existe des RM activées dans diverses activités cognitives (cf. (CONCEPT)), alors ces RM possèdent *de fait* la propriété décrite par (CONST). Il est aussi facile d'imaginer un esprit qui serait doué de plusieurs milliers de RM satisfaisant (CONCEPT), en étant par exemple capable de les exprimer par des milliers d'onomatopées et de les appliquer aux objets, mais qui serait incapable de les composer.

En vertu de α , (CONST) est du type R. (CONST) se distingue des autres processus de type R en raison de β . C'est donc β qu'il convient d'expliquer²⁰. Je réserve le terme « constituant » pour les composants qui entretiennent cette relation spécifique avec le complexe.

2.1. En quel sens les concepts se composent-ils ?

2.1.1. Analyse de la notion de composition

La notion de composition des concepts est bien moins claire qu'il n'y paraît, en particulier dès que l'on renonce à suivre de trop près l'analogie avec les expressions des langues naturelles (cf. (CM) en dessous) et les métaphores spatiales de l'inclusion des concepts composés dans le concept composant (« contenir », « inclure », « avoir pour partie »).

Distinguons quatre manières de concevoir la composition des concepts. La *première* est méréologique :

(CM) — Les concepts sont les parties physiques des RM conceptuelles complexes.

(CM) n'a de sens que rapportée aux états neuraux qui réalisent les RM : en tant que telles, les RM n'ont pas de parties physiques. En outre, (CM) est conçue par analogie avec les langues naturelles : les mots se composent en effet au sein de phrases au sens où ils sont les parties physiques de celles-ci. Mais il est peu probable que la composition des concepts soit de cette

nature (même si la question est empirique) : même si l'activation des concepts CHAT et NOIR correspond à l'activation de deux patterns neuraux, il est peu crédible qu'un pattern neural correspondant à CHAT NOIR ait pour parties physiques les deux patterns neuraux initiaux.

La *deuxième* fait appel aux propriétés physiques des états neuraux qui réalisent les concepts :

(CPh) — Les propriétés physiques des états neuraux qui réalisent des concepts complexes ou des croyances sont déterminées par les propriétés physiques des états neuraux qui réalisent les concepts composants, et ce en fonction des règles de composition du concept complexe.

(CPh) est bien plus plausible que (CM), même si le rejet de l'identité type-type rend complexe la dépendance des propriétés physiques des concepts complexes à l'égard des propriétés physiques des diverses réalisations possibles des concepts composants (thèse de la réalisation multiple²¹). Néanmoins, même si la thèse (CPh) était vraie, elle ne dit rien sur la composition des concepts complexes et des croyances. Soutenir que les concepts se composent, ce n'est en effet pas soutenir une thèse sur les réalisations physiques des états mentaux, mais sur un certain type de relations, entre les concepts composants et les concepts complexes : les premiers constituent les seconds.

La *troisième* manière de concevoir la composition repose sur le contenu des RM : elle est donc sémantique.

(CIn) — Γ est une RM conceptuelle complexe si et seulement si son contenu est complexe.

POISSON ROUGE est un concept complexe, parce qu'il a pour extension l'ensemble des poissons rouges, c'est-à-dire l'intersection de l'ensemble des poissons et de l'ensemble des choses rouges (ou bien exprime la propriété complexe *être un poisson rouge*). Pas plus que les deux premières explications de la notion de composition, la thèse (CIn) ne peut être retenue²².

¹⁹ Quel est le lien entre (CONST) et l'hypothèse du langage de la pensée ?

(LP) Le système conceptuel est de nature linguistique. On peut soutenir raisonnablement qu'un ensemble F de symboles est linguistique si a/ parmi F, certains symboles sont primitifs tandis que les autres résultent de la composition des symboles primitifs selon certaines règles ; et b/ si parmi les règles de composition des symboles, on peut distinguer des règles syntaxiques d'autres types de règles : *pas de langage sans syntaxe*. Une règle est syntaxique si elle détermine les compositions possibles des symboles indépendamment de leurs propriétés sémantiques. (CONST) implique assurément a/, mais sans doute pas b/. D'où il est possible que (CONST) soit vraie et que (LP) soit fausse, autrement dit que la pensée conceptuelle soit composée par des constituants sans être pour autant linguistique : $\diamond ((CONST) \wedge \neg (LP))$.

²⁰ (CONST) est une application aux concepts du problème classique en sciences cognitives de la « constitution » (constituency) des représentations mentales.

²¹ L'identité type-type et la thèse de la réalisation multiple concernent les relations entre deux classes de propriétés, ici les propriétés mentales (par exemple, être la croyance que la France est en Europe...) et des propriétés physiques (par exemple, être telle activation neuronale...). L'identité type-type affirme qu'une entité possède une propriété A de la première classe ssi elle possède une propriété A de la seconde, donc qu'un état mental donné, par exemple la croyance que la France est en Europe, n'est rien d'autre qu'un état physique donné, par exemple telle activation neuronale (que *toutes* les occurrences de cette croyance sont des occurrences de cet état neural). La thèse de la réalisation multiple affirme qu'une entité possède une propriété A de la première classe ssi elle possède une propriété de la seconde classe parmi un ensemble de propriétés possibles {A, B, D}. Donc différentes occurrences d'un *même* état mental, par exemple la croyance que la France est en Europe, correspondent aux occurrences de *différents* états physiques, par exemple *différents* patterns neuraux.

²² Ce principe semble avoir pourtant été admis par les empiristes, en particulier par Hume : un concept simple exprime une propriété simple, c'est-à-dire sensorielle. Toute propriété complexe — donc toutes les propriétés qui ne sont pas sensorielles — est exprimée par un concept lui-même complexe, résultant de la composition des concepts simples. Par ailleurs, un tel principe semble devoir être admis par les

La notion de contenu complexe prise absolument (tout court) est obscure. La complexité d'un contenu de symbole semble en effet relative au langage utilisé pour exprimer ce contenu : le contenu d'un symbole est complexe si son contenu est exprimé conventionnellement dans une autre langue, ou éventuellement dans la même, par un symbole complexe (d'après des critères indépendants, par exemple syntaxiques) dont les parties sont douées d'un contenu propre. L'expérience de pensée suivante semble établir ce point. Supposons en effet une langue qui ne distingue que deux couleurs, le clair et le foncé. En outre, cette langue ne possède pas de termes spécifiques pour ces deux couleurs. Chaque objet coloré y est désigné par deux mots : le premier désigne l'objet s'il est clair et le second s'il est foncé. Ces deux mots ne sont pas des variantes obtenues à partir d'une racine commune au moyen de morphèmes exprimant les deux teintes. Pour nous, un mot désignant un objet foncé posséderait un contenu complexe. Pourquoi ? Parce que nous exprimerions le contenu de ce mot au moyen d'une expression composée, à savoir d'un adjectif spécifique (« fonc ») qui se composerait avec le nom de l'objet pour désigner une sous-classe de la classe des objets désignés par le nom. En revanche, pour les locuteurs de ce langage fictif, un tel mot n'exprimerait sans doute pas un contenu complexe. Pourquoi ? Nous pouvons affirmer cela par analogie avec la signification que nous attribuons à certains mots français. Pour un locuteur du français, le nom « coq » n'exprime pas le contenu complexe « gallinacé mâle » (et pareillement pour tous les autres noms d'animaux distingués en fonction de leur genre), dans la mesure où « coq » n'est pas conventionnellement équivalent à « gallinacé mâle ». En revanche, pour le locuteur d'une langue qui ne distinguerait pas les noms des animaux en fonction de leur genre, mais assignerait un nom à chaque espèce, modifié par deux adjectifs signifiant l'un « mâle » et l'autre « femelle », le nom « coq » aurait sans doute un contenu complexe.

En conséquence, la question « le mot « *chabada* » a-t-il un contenu complexe ? » n'a sans doute pas de sens. Il faut préciser le langage relativement auquel la complexité du contenu de « *chabada* » est évaluée. Ce langage peut être le langage auquel appartient ce mot, ou bien une langue différente.

En outre, dans certains systèmes de signes, la complexité d'un symbole, déterminée de manière syntaxique (application de règles récursives), n'est pas corrélée à la complexité de son contenu (déterminée par rapport aux contenus possibles dans ce système de signes). C'est le cas du calcul des propositions : la valeur de vérité d'une formule complexe, par exemple la tautologie $\Phi = p \rightarrow (q \rightarrow p)$, n'a pas pour partie les valeurs de vérité des formules qui la composent. La valeur de vérité de Φ est fonction des valeurs de vérité attribuées à ses composants, p et q , comme je l'ai montré plus

philosophes qui défendent la composition des états mentaux, mais rejettent toute idée de règles de composition analogues aux règles syntaxiques. C'est le cas de Millikan (1990 ; 1993b) : même si les états mentaux des systèmes cognitifs complexes (appelés « représentations ») sont complexes (« articulés ») et entrent dans des relations inférentielles, cette complexité est selon elle de nature sémantique (Millikan, 1993b, p. 109). De même, la composition des états mentaux reconnue par Barwise et Perry doit être sémantique.

haut. Mais cette détermination fonctionnelle ne signifie nullement que les valeurs de vérité attribuées à p et à q sont des parties de la valeur de vérité attribuée à Φ . La valeur de vérité de Φ est le vrai, *simpliciter*, c'est-à-dire dans le même sens que la valeur de vérité de p ou de q peut être le vrai.

(CIn) exclut qu'un concept simple puisse exprimer un ensemble que l'on pourrait par ailleurs définir comme l'intersection d'autres ensembles (ou une conjonction de propriétés) : un tel concept serait par définition composé. (CIn) ne peut donc faire la différence entre les deux situations suivantes :

- un individu possède un concept ayant pour extension la classe des poissons rouges sans l'avoir produit par composition de deux autres concepts parce qu'il ne possède pas de concept ayant pour extension la classe des poissons, ni de concept ayant pour extension la classe des choses rouges (pensons par exemple à un aveugle, qui ne connaîtrait pas les noms de couleurs et n'aurait été en contact qu'avec des poissons rouges à l'exclusion d'autres poissons — et ce par l'intermédiaire d'un détecteur de poissons rouges...);
- un individu possède un concept ayant pour extension la classe des poissons rouges par composition du concept ROUGE et du concept POISSON.

Outre le caractère apparemment confus de la notion de complexité absolue du contenu, on voit donc que la complexité d'un symbole n'implique pas la complexité de son contenu et qu'inversement la complexité d'un contenu n'implique pas la complexité d'un symbole. (CIn) semble donc inadéquat pour décrire la composition de concepts au sein de RM conceptuelles complexes.

La morale semble être la suivante : la notion de constitution ne porte pas en premier lieu sur le *contenu* des RM, c'est-à-dire sur leurs propriétés sémantiques. **Elle concerne donc avant tout leur véhicule**, c'est-à-dire les propriétés des RM qui ne sont pas sémantiques, par exemple leurs propriétés syntaxiques si l'on admet l'hypothèse d'un langage de la pensée.

La *quatrième* fait appel au rôle fonctionnel des concepts complexes et des croyances :

(CF) — Un concept Γ a pour constituant le concept Ψ si et seulement si l'on peut inférer, en vertu de la structure de Γ , du jugement « ceci est Γ » un jugement contenant Ψ , par exemple « ceci est Ψ ». Une croyance Φ contient le concept Ψ si et seulement si Φ peut prendre part à un raisonnement ayant Ψ pour moyen terme.

Le concept CHAT NOIR est, d'après (CF), composé du concept CHAT : si je juge qu'un item appartient à l'extension de CHAT NOIR, je peux inférer, en vertu de la structure de CHAT NOIR, qu'il appartient à l'extension du concept CHAT.

La formulation de (CF) distingue les jugements contenant Ψ du jugement « ceci est Ψ ». Parmi les adjectifs, appelés souvent « modificateurs », on trouve en effet des adjectifs subsectifs et des adjectifs dépendants du contexte. Un adjectif A est dépendant du contexte si une entité ne satisfait A que par référence à

une classe d'objets désignée par un nom N : « grand » est ainsi un adjectif dépendant du contexte, puisqu'une chose n'est grande que relativement à la classe à laquelle elle appartient. Une grande souris n'est grande que par rapport à l'espèce des souris. Un adjectif est subsectif (par opposition à intersectif) si et seulement si il n'a pas de référence par lui-même (comme les adjectifs dépendants du contexte) et si sa contribution sémantique à un AN est de sélectionner un sous ensemble de la classe désignée par N. Par exemple, « habile » est subsectif parce qu'il sélectionne un sous-ensemble des classes auxquelles réfèrent les noms qu'il modifie, par exemple la classe des musiciens ou des chirurgiens. Je suppose que les concepts exprimés par des adjectifs subsectifs et dépendants du contexte ont les mêmes propriétés : un concept est dépendant du contexte si et seulement s'il ne s'applique à une entité que par référence à une classe d'objets désignée par un autre concept. GRAND est ainsi un concept dépendant du contexte. Un concept est subsectif (par opposition à intersectif) si et seulement s'il n'a pas de référence par lui-même (comme les concepts dépendants du contexte) et si sa contribution sémantique à un concept complexe est de sélectionner un sous ensemble de la classe désignée par l'autre concept. Par exemple, HABILE est un concept subsectif (Kamp et Partee, 1995). De ce fait, si Ψ est un concept subsectif ou dépendant du contexte et Γ un concept complexe contenant Ψ , on ne peut déduire de « ceci est Γ » le jugement « ceci est Ψ » : un tel jugement, nous venons de le voir, n'a pas de sens. On peut en revanche déduire un jugement de la forme « pour un Δ , c'est Ψ » : ce jugement contient bien Ψ , conformément à (CF). De même, si une entité satisfait un concept complexe obtenu par disjonction, par exemple CHAT OU CHIEN, on ne peut pour autant déduire qu'elle satisfait l'un des concepts constituants. En revanche, il est possible de déduire que si cette entité n'est pas un chat, alors elle est un chien.

Pour autant, cette analyse ne peut être retenue. En effet, elle repose sur la notion de structure d'un concept complexe. De ce fait, l'analyse de la notion de constituant est circulaire, puisqu'il est difficile d'expliquer la notion ici pertinente de structure d'une RM conceptuelle complexe sans faire appel à la notion de constituant. En outre, la notion de structure ne peut être enlevée de (CF). En effet, certains philosophes caractérisent le contenu des concepts au moyen de leur rôle inférentiel : certaines inférences sont constitutives du contenu même du concept en question. Par exemple, si le rôle inférentiel de CELIBATAIRE contient l'inférence CELIBATAIRE (X) \rightarrow NON MARIE (X), l'inférence « Paul est un célibataire ; donc Paul n'est pas mari » repose sur le contenu même de CELIBATAIRE. Autrement dit, je peux inférer du jugement « Pierre est un célibataire » le jugement « Pierre n'est pas mari ». Pourtant, d'après la sémantique des rôles inférentiels, le concept CELIBATAIRE n'a pas pour constituant le concept NON MARIE et l'inférence précédente ne dépend pas de la structure de CELIBATAIRE. Si l'on veut expliquer de manière fonctionnelle (CONST), il faut donc pouvoir distinguer parmi les inférences qu'autorise un jugement de la forme « ceci est Γ », celles qui dépendent d'un éventuel rôle fonctionnel et celles qui n'en dépendent

pas. La notion de structure est donc nécessaire, mais circulaire²³.

Proposons donc *finalement* l'analyse suivante de la notion de constituant et donc de (CONST). Elle fait appel à la relation de coactivation (Fodor) :

(CCo) — Un concept complexe AB a pour constituant le concept A si et seulement si A est nécessairement activé chaque fois que AB l'est (*mutatis mutandis* pour les autres RM conceptuelles complexes).

Être le constituant d'un complexe, c'est donc être nécessairement²⁴ coactivé quand le concept complexe l'est : CHAT NOIR ne peut être activé sans que le concept CHAT le soit également (ainsi que NOIR)²⁵. La croyance LE CHAT EST SUR LE TAPIS ne peut être entretenue sans que le concept TAPIS soit activé.

À la différence des deux premières analyses, (CCo) porte bien sur les relations entre les concepts complexes et les concepts composants et non sur les relations entre leurs réalisations physiques. Ce n'est qu'incidemment que cette solution concerne les relations entre les réalisations physiques de A et de AB : A doit être réalisé physiquement quand AB l'est. Néanmoins, elle ne statue pas sur la relation entre la réalisation des complexes et celles des constituants : la réalisation physique de A n'a ni à être une partie de celle de AB (vs. (CM)) ni même à déterminer les propriétés physiques de la réalisation de AB (vs. (CPh)). En outre, à la différence de (CF), (CCo) n'implique aucun cercle vicieux : aucune notion mentionnée dans (CCo), en particulier la notion de coactivation, n'implique la notion de constituant. Enfin, (CCo) ne concerne pas le contenu des concepts, mais bien leur véhicule, puisque la coactivation est une relation entre RM et non pas une propriété de leur contenu.

2.1.2. Commentaires

1. (CCo) permet d'expliquer **la notion de structure** d'une RM conceptuelle complexe. De manière générale, une RM complexe Φ , produite à partir de Ψ et de Γ selon le mode de composition α , possède une structure si et seulement si elle possède un ensemble Z de propriétés intrinsèques $\{p_1, \dots, p_n\}$ tel que parmi les RM complexes, toute RM produite à partir de Ψ et Γ par α , et seules elles, possèdent Z²⁶. Distinguons deux types de structures, **les structures explicites** et **les structures implicites** . Φ est explicitement structurée si et seulement si parmi les propriétés de Z, on trouve les deux propriétés suivantes, *avoir pour constituant Ψ* et *avoir pour constituant Γ* et si une ou plusieurs autres

²³ (CF) pose un autre problème : la nature de la possibilité (« peut inférer ») doit être expliquée.

²⁴ La notion de nécessité permet de distinguer les associations entre concepts des véritables constituants. Supposons que le concept ABOIE soit activé chaque fois que Jean pense aux chiens. Cette association constante n'implique pas que ABOIE est un constituant de CHIEN, parce que ABOIE n'est pas *nécessairement* activé quand CHIEN l'est : CHIEN pourrait être activé sans que ABOIE le soit.

²⁵ À l'avenir, j'emploierai le terme « constituant » pour désigner des RM ayant la propriété décrite par (CCo).

²⁶ Cette définition laisse ouverte la possibilité de RM non complexes possédant Z.

propriétés correspondent à α . De ce fait, une RM conceptuelle complexe qui satisfait (CONST) et (CCo) possède une structure explicite si les constituants de cette RM possèdent des propriétés qui correspondent à la manière dont ils ont été composés. Une RM est implicitement structurée si et seulement si aucune des propriétés de Z ne correspond directement à ses composants.

Considérons les RM explicitement structurées. Si les propriétés qui correspondent aux modes de composition valent uniformément pour toutes les RM conceptuelles complexes (si une même propriété correspond à α pour toutes les RM), alors la possession d'une structure explicite peut *indiquer* aux parties de l'esprit dont le comportement dépend des RM la manière dont celles-ci ont été composées (ou bien permettre l'application aux RM de processus mentaux sensibles à la structure des RM²⁷).

Des remarques analogues valent pour la structure des énoncés linguistiques. Soit les deux énoncés suivants :

(8) *Pierre regarde le chien.*

(9) *Le chien regarde Pierre.*

Par convention, le français indique l'ordre des arguments des prédicats à deux places exprimés par des verbes transitifs au moyen de relations spatiales entre les parties de la phrase : en général, le premier argument (le sujet) précède le verbe tandis que le second (le complément) suit le verbe. Par l'ordre linéaire de droite à gauche (donc par une propriété relationnelle entre les constituants des énoncés), (8) indique donc que le prédicat « regarder » est appliqué à la paire (Pierre, le chien). De même, les différentes propriétés spatiales des énoncés (8) et (9) indiquent que deux paires d'arguments différentes satisfont le prédicat « regarder ». Les langages habituels de la logique des prédicats utilisent aussi des propriétés spatiales, mais d'une autre manière, pour distinguer les différentes paires d'arguments de prédicats à deux places :

(11) *Rab*

(12) *Rba*

Dans certaines langues, des modifications des racines des mots qui sont les constituants des phrases (des déclinaisons) correspondent aussi à la manière dont les mots sont composés : c'est le cas du latin.

En général, des propriétés diverses (relations spatiales, déclinaisons...) sont employées conjointement pour correspondre à la manière dont les atomes (les mots ou les morphèmes) sont composés. Elles ont toutes pour fonction d'indiquer au locuteur la manière dont les mots ont été combinés. Néanmoins, dans les langues naturelles, les propriétés des énoncés qui indiquent les modes de combinaison sont souvent ambiguës : ce sont les ambiguïtés de structure qui disparaissent dans les langages symboliques.

2. (CCo) explique en partie pourquoi les RM conceptuelles complexes sont susceptibles d'être *décomposées* dans leurs concepts constituants. Il suffit que les processus mentaux soient sensibles aux

²⁷ C'est le cas dans les modèles *computo-représentationnels* de l'esprit.

propriétés des RM complexes qui indiquent comment les constituants ont été composés : ces propriétés et les constituants forment donc un *input* possible pour un processus décomposant les concepts complexes.

2.2. Composition et compositionnalité

Un point important mérite d'abord d'être souligné : **la composition (CONST) n'est une condition nécessaire d'aucune des formes de compositionnalité.** (CONST) n'est que l'un des processus possibles de type R. D'autres processus de type R, par exemple la combinaison des prototypes (cf. plus loin), peuvent prétendre satisfaire ces différentes formes de compositionnalité (ce qui ne signifie pas qu'ils les satisfont effectivement).

2.2.1. Lien entre (CONST) et (CS)

De nombreux philosophes et psychologues²⁸ ne distinguent pas nettement la thèse selon laquelle les concepts se composent selon des règles de composition au sein de RM conceptuelles complexes dont ils sont les constituants de la thèse selon laquelle les propriétés sémantiques d'une RM conceptuelle complexe sont une fonction des propriétés sémantiques des concepts constituants et de la manière dont ces concepts sont composés. Cela est facteur de confusion.

Les résultats de Zadrozny (1994), Janssen (1997) et Hodges (1998) mentionnés dans le commentaire 6 de (CS) montrent qu'un système de symboles qui possède une syntaxe récursivement définie ne satisfait *ipso facto* la compositionnalité sémantique que si les propriétés sémantiques des symboles ne sont pas contraintes. En conséquence, puisque les propriétés sémantiques des concepts ne sont pas arbitraires, à supposer même que les règles de composition du système conceptuel soient analogues à une syntaxe récursivement définie, il est possible que le système conceptuel satisfasse (CONST) sans satisfaire (CS). La thèse de la composition des concepts, (CONST), n'implique donc pas à elle seule la thèse de la compositionnalité sémantique, (CS). *A fortiori*, les deux thèses ne sont pas équivalentes et ne doivent donc pas être confondues. On a donc :

(13) $\alpha((CONST) \wedge \neg(CS))$

2.2.2. Lien entre (CONST) et (CCP)

Le fait que les concepts se composent n'implique pas qu'il soit suffisant pour avoir un concept complexe d'avoir les concepts constituants et de connaître la règle de composition :

(14) $\alpha((CONST) \wedge \neg(CCP))$

2.2.3. Lien entre (CONST) et (CI)

Si les concepts se composent selon (CONST), alors ils satisfont la compositionnalité inverse.

²⁸ Par exemple, Shanks, (1997). Fodor lui-même emploie le terme « compositionality » pour désigner aussi bien (CONST) que l'un des 3 types de compositionnalité proposés ici ((CS), (CCP), (CI)). D'où l'ambiguïté profonde des appels à la compositionnalité dans bien des arguments pour ou contre les prototypes et plus généralement dans les arguments concernant la nature des concepts.

Selon (CI) les concepts composants apportent leurs conditions de possession aux concepts complexes. En outre, si p est une propriété d'un concept Γ , on peut toujours formuler, à partir de p , une condition de possession de Γ - qui doit être héritée par les concepts complexes dont Γ est un constituant (cf. le premier point des commentaires de A/ 2/). Par exemple, si la référence à la classe des chats est une propriété du concept CHAT, alors avoir une RM qui fait référence à la classe des chats est une condition de possession du concept CHAT. Et cette condition de possession doit être héritée par les concepts dont CHAT est un constituant, par exemple CHAT BOTTE ou CHAT NOIR, si les concepts satisfont (CI). En conséquence, avoir une RM qui fait référence à la classe des chats est une condition de possession de CHAT BOTTE. Comment cela est-il possible si CHAT BOTTE ne fait pas référence à la classe des chats, mais à la classe des chats bottés ? (CONST) et (CCo) expliquent cela : avoir une RM qui fait référence à la classe des chats est une condition de possession de CHAT BOTTE parce que CHAT est un constituant de CHAT BOTTE. Autrement dit, CHAT est activé quand CHAT BOTTE l'est. En conséquence, quiconque satisfait les conditions de possession de CHAT BOTTE doit aussi satisfaire les conditions de possession de CHAT. Donc quiconque satisfait les conditions de possession de CHAT BOTTE doit avoir une RM qui fait référence à la classe des chats. Bref, on a

(15) (CONST) \rightarrow (CI).

En résumé, j'ai analysé dans cette deuxième partie les notions de composition, de constituant et de structure explicite : les concepts complexes sont composés si et seulement si leurs composants les constituent. Les constituants sont nécessairement coactivés quand le complexe l'est. Une RM est explicitement structurée si elle a des constituants et possède des propriétés qui indiquent leur mode de composition. En outre, j'ai examiné les relations entre la composition et les formes de compositionnalité distinguées antérieurement : la composition implique (CI), mais n'implique ni (CS) ni (CCP).

3. Les concepts se composent-ils ?

Il s'agit d'examiner les données (plus ou moins) empiriques pouvant soutenir (CONST)²⁹.

3.1. Analogie avec les langues naturelles

Les énoncés déclaratifs servent entre autres à exprimer et à communiquer les croyances des sujets. Or, par hypothèse, les mots expriment des concepts et, d'autre part, les mots sont les constituants des énoncés déclaratifs (en un sens analogue à (CCo)). Le modèle

²⁹ On pourrait s'étonner qu'un philosophe avance des arguments empiriques. Néanmoins, l'analyse conceptuelle ne peut se passer d'une interaction avec des données empiriques : il s'agit d'évaluer si les notions proposées ont une pertinence eu égard à l'état actuel des connaissances empiriques. Notons ici que les arguments proposés sont tous de nature abductive — voir plus loin sur ce type d'argument

suivant de la compréhension est donc tentant : un locuteur comprend un énoncé déclaratif en attachant à chaque mot le concept conventionnellement associé et en déterminant la croyance exprimée par cet énoncé par composition des concepts conformément à la syntaxe de l'énoncé. Autrement dit, les croyances résultent de la composition des concepts associés aux mots. Puisque les mots sont des constituants des énoncés, il semble donc possible que les concepts soient les constituants des RM conceptuelles complexes :

Concepts	Mots
Croyances	Énoncés déclaratifs

Ainsi formulé, cette analogie est en fait peu convaincante. En effet, il *suffit* de supposer un processus de type R qui s'applique aux concepts et produit des croyances — que les concepts soient ou non des constituants des croyances — pour rendre compte de la compréhension des croyances exprimées par les énoncés déclaratifs. Le fait que les mots sont les constituants des énoncés ne nous contraint nullement à croire que les concepts sont aussi les constituants des croyances : la présence constitutive des mots au sein des énoncés a peut-être pour fonction d'indiquer l'*input* du processus supposé de production de croyances — et ce afin de permettre aux auditeurs d'acquérir la RM exprimée par l'énoncé. En conséquence, cette analogie est insuffisante pour défendre l'idée selon laquelle les concepts sont les constituants des croyances, *a fortiori* des RM conceptuelles complexes de manière générale.

3.2. Systématicité de la pensée

Lors de leur critique des architectures connexionnistes, Fodor et Pylyshyn (Fodor et Pylyshyn, 1988) ont insisté sur une propriété de la pensée conceptuelle humaine, la systématicité³⁰. La pensée est systématique si la capacité d'avoir certains états mentaux est corrélée de manière non accidentelle à la capacité d'avoir d'autres états mentaux dont le contenu est sémantiquement relié. Par exemple, la capacité d'un individu d'avoir une croyance dont le contenu est la proposition [Pierre aime Marie] est corrélée de manière non accidentelle à la capacité d'avoir une croyance dont le contenu est la proposition [Marie aime Pierre]. L'argument de Fodor et Pylyshyn est abductif : seul le principe (CONST) explique la systématicité, étant donné nos connaissances actuelles.

Une telle affirmation est-elle défendable ? Supposons de nouveau que les RM conceptuelles complexes soient produites selon un processus L de type R tels que les concepts composants ne soient pas des constituants des RM complexes. Autrement dit, L n'est *pas* un processus de composition. En outre, le contenu des RM conceptuelles complexes produites est systématique-

³⁰ Cette notion est proche de la contrainte de généralité proposée par Evans (1982). La RM F qui est un composant de la RM complexe prédicative Fa satisfait la contrainte de généralité ssi elle peut se composer avec toutes les autres RM susceptibles de former une RM prédicative, par exemple b, c... Néanmoins, alors que Evans fait de la contrainte de généralité une propriété nécessaire définitionnelle des concepts (par stipulation, rien n'est un concept s'il ne satisfait la contrainte de généralité), la systématicité est présentée comme une propriété empirique des concepts.

ment déterminé en fonction du contenu des *inputs*. Autrement dit, L est sémantiquement compositionnel. Un tel processus prend donc comme *inputs* des concepts comme PIERRE, MARIE et AIMER pour produire des RM conceptuelles complexes comme la croyance dont le contenu est la proposition [Pierre aime Marie] (qui n'a pour constituants ni PIERRE, ni MARIE ni AIMER). Un tel processus suffit-il à expliquer la systématisme ? Autrement dit, la thèse β de (CONST) est-elle véritablement nécessaire pour expliquer la systématisme de la pensée conceptuelle ? Si les croyances résultent d'un processus prenant comme *inputs* d'autres RM (les concepts) et si le contenu des croyances est compositionnellement déterminé par le contenu des *inputs*, alors la possibilité de remplacer les *inputs* les uns par les autres semble expliquer la production de RM dont le contenu est sémantiquement relié, donc la systématisme. Le remplacement de l'*input* PIERRE par l'*input* MARIE, et de l'*input* MARIE par l'*input* PIERRE, dans le processus de type R envisagé semble ainsi suffisant pour obtenir la croyance dont le contenu est [Marie aime Pierre]. En conséquence, la coactivation de PIERRE et de la RM correspondant à la croyance ne semble jouer aucun rôle dans l'explication de la systématisme. La possession des concepts, l'existence d'un processus de type R et la compositionnalité sémantique semblent suffire à expliquer la systématisme, même si les concepts ne satisfont ni (CONST) ni (CCo). Puisque l'argument en faveur de (CONST) est abductif et qu'une hypothèse plus simple que (CONST) explique la systématisme, il ne semble pas légitime d'admettre (CONST) sur cette base.

3.3. La capacité de raisonner est-elle reliée à la composition des concepts au sein des pensées conceptuelles ?

De manière générale, raisonner, c'est passer de certains états mentaux (les prémisses) à un autre état mental (la conclusion). Une explication simple et immédiate (qui n'est pas pour autant juste) d'une telle capacité est la suivante : l'esprit passe des prémisses à la conclusion parce que α les prémisses possèdent des constituants et, de ce fait, des éléments communs et parce que β les processus mentaux impliqués dans le raisonnement sont sensibles à la structure de ces états mentaux. Supposons en revanche un processus de type R tel que les états mentaux complexes produits par ce processus n'aient pas pour constituants les états mentaux composants³¹. Ces états mentaux pourraient-ils prendre part à un raisonnement ? Puisque, par hypothèse, les prémisses n'ont pas de constituants, la transition vers la conclusion ne peut être expliquée par des processus sensibles à la structure des états mentaux. Différentes positions peuvent alors être soutenues.

- Le contenu des prémisses explique la transition : les processus impliqués sont sensibles aux propriétés sémantiques communes des états mentaux servant de prémisses. Mais une telle position est délicate si l'on admet que les propriétés sémantiques sont en tant que telles causalement inertes. Une analyse externaliste du contenu mental semble en outre incompatible avec des processus mentaux n'ayant

accès qu'aux propriétés internes, c'est-à-dire n'impliquant pas une relation entre l'esprit et le monde, des RM.

- Certains sous-systèmes de l'esprit ont pour fonction d'interpréter les RM et d'identifier comme commun le contenu commun aux prémisses (le moyen terme)³². Cette solution est différente de la précédente : les processus n'ont pas à être sensibles aux propriétés sémantiques, mais seulement à l'identification opérée par ces sous-systèmes. Néanmoins, les problèmes précédents se reposent : comment une partie de l'esprit peut-elle interpréter d'autres parties de l'esprit puisque a/ au moins certaines propriétés des RM sont relationnelles (en particulier, leur référence) et b/ les propriétés sémantiques en tant que telles sont causalement inefficaces. Admettons pourtant qu'ils peuvent être résolus. Le problème principal est en effet ailleurs.

L'interprétation du contenu des états mentaux ne peut bien sûr être déterminée ni par la structure explicite des états mentaux (puisque par hypothèse, ceux-ci n'en ont pas) ni par les composants auxquels le processus de type R est appliqué (puisque ceux-ci, n'étant pas des constituants, ne sont pas nécessairement activés). En conséquence, chaque état mental est un symbole atomique. Les parties de l'esprit supposées interpréter les RM doivent donc déchiffrer un nombre colossal (ou infini³³) de RM. Autrement dit, le problème que pose l'interprétation d'un système productif de symboles se manifeste alors à l'intérieur même de l'esprit³⁴. Comment un sous-système de l'esprit pourrait-il interpréter un nombre gigantesque de symboles différents ? Bref, si l'on soutient que la capacité générale de raisonner ne suppose pas la composition des états mentaux selon un principe analogue à (CONST), il faut expliquer comment des sous-systèmes ont la possibilité d'interpréter un nombre gigantesque d'états mentaux possibles. En l'absence de réponse, nous sommes donc conduits à admettre l'idée selon laquelle les états mentaux complexes qui interviennent dans un raisonnement doivent avoir pour constituants les états mentaux composés³⁵.

³² Telle semble être la position de Millikan. Cette position permet en outre d'expliquer, sans recourir à (CCo), aussi bien (CF) que la possibilité de décomposer des RM conceptuelles complexes. Néanmoins, une telle explication se heurte à la difficulté soulevée dans le texte : étant donné la productivité de l'esprit, comment comprendre qu'un système interne à l'esprit puisse interpréter tous les concepts (atomiques et complexes) ?

³³ Cf. plus loin la distinction des deux types de productivité.

³⁴ Malgré tout, l'hypothèse ii/ peut s'appliquer correctement à certains types de raisonnements. Supposons en effet une classe finie de RM dont le contenu est complexe. Ces RM, à l'exclusion des autres RM, participent à des raisonnements spécifiques — ceux-ci ont peut-être une fonction particulière. Supposons en outre qu'il existe des parties de l'esprit ayant pour fonction l'interprétation de ces états mentaux : puisque la classe est par hypothèse finie, ceux-ci peuvent les reconnaître et déterminer leur contenu. Donc ces raisonnements ne supposent pas la composition des prémisses.

³⁵ Je ne suis pas absolument certain du caractère convaincant de cet argument : il suppose que les processus mentaux, c'est-à-dire les transitions entre RM, sont au même niveau que l'interprétation des RM, ce que rejettent certaines formes de

³¹ La structure argumentative est la même que précédemment.

Le raisonnement a été décrit précédemment de manière générale et rien de ce qui a été défendu ne porte spécifiquement sur le raisonnement conceptuel. On peut parfaitement admettre des types de raisonnements employant des états mentaux de nature non conceptuelle, auxquels l'exigence de constituants s'appliquerait tout autant. Citons ainsi Millikan :

« *Imagine a creature that carries about in its head intentional icons that are three-dimensional maps of various places it has been — not maps in three spatial dimensions, of course, but dimensions in some neurally realizable medium. It has a map of the locale in which it last found water, and another of the locale in which it last saw lions. On each of these maps its den is marked. Now imagine that it overlaps these maps, using its den as pivots, and arrives at a third map showing the proximity of lions to the source of water. Reacting to this new map, it seeks a new source of water (...). (...) this would be a paradigm case of mediating inference. As is characteristic of all mediate inference, two vehicles of information have been combined, using a middle term, so as to produce a third vehicle containing new information.* » (Millikan, 1993a, 104)

Bien sûr, j'applique ce résultat spécifiquement au raisonnement conceptuel, qui implique par exemple des croyances. La nature du raisonnement et la productivité nous font donc admettre (CONST)³⁶, parce que (CONST) suffit à expliquer la possibilité du raisonnement conceptuel à la différence de processus qui ne satisfont pas β de (CONST).

En résumé, j'ai montré dans cette troisième partie que, à la différence de la compréhension des énoncés linguistiques et de la systématité, le raisonnement semble requérir la composition des RM conceptuelles.

4. Les concepts possèdent ces trois types de compositionnalité

Les données empiriques sur la pensée conceptuelle permettent-elles d'attribuer aux concepts l'une de ces formes de compositionnalité ?

connexionnisme (Smolensky, 1992). Néanmoins, le connexionnisme doit relever le défi suivant : si les propriétés qui guident le cours du raisonnement ne sont pas reliées aux propriétés sémantiques, comment le raisonnement peut-il être sémantiquement fiable ? Autrement dit, comment se fait-il que la conclusion d'un raisonnement soit sémantiquement impliquée par les prémisses si les propriétés qui guident le cours du raisonnement n'ont rien à voir avec le contenu des RM ? Certains travaux suggèrent que le défi n'est pas insurmontable, même s'il n'a pas encore été relevé (Smolensky, 1988, 1991, 1992 ; Van Gelder, 1990).

³⁶ Ce raisonnement est une abduction : cf. plus loin sur ce type de raisonnement.

4.1. (CI)

4.1.1. Objection de Shanks

On a vu que (CONST)→(CI) ; par *modus ponens*, on a donc (CI)³⁷.

Cette thèse est néanmoins contestée par Shanks (1997) au nom de la dépendance contextuelle des propriétés des concepts³⁸.

« *We do not really want COFFEE to have any privileged connection with CUP WITH COFFEE over and above micro-feature overlap³⁹, because it is in fact mistaken to believe that the concept COFFEE is the same when it appears in isolation as when it appears as an element of a more complex concept. The reason for this is that mental representations are inevitably context-dependent in a way that is not naturally dealt with in the symbolic theory.* » (p. 204)

Son argument est le suivant (p. 202-205). Les activités et les tâches cognitives mobilisant des concepts sont extrêmement sensibles au contexte dans lequel elles se déroulent. C'est le cas entre autres des jugements de typicalité : le degré de typicalité d'un objet varie en fonction du contexte d'évaluation. Ainsi une vache est-elle typique de la classe des animaux dans le contexte de production de lait et fort peu typique dans le contexte de l'équitation (Roth et Shoben, 1983, cité par Shanks, 1997). C'est aussi le cas de la catégorisation : selon les contextes, un liquide comme le thé est ou non jugé être de l'eau (Malt, 1994). L'accès lexical est pareillement sensible au contexte : les temps de réaction dans diverses tâches lexicales supposant l'activation d'un même concept dépendent du contexte. Enfin, les propriétés qui sont jugées caractériser une catégorie

³⁷ Inversement, (CI) fournit un argument par abduction en faveur de (CONST). Notons un problème avec cette manière d'établir (CI). Si (CI) est ainsi établi, ce principe ne peut être un argument dans les contextes où la composition de la pensée conceptuelle n'est pas présupposée. (CI) peut être établi autrement, par exemple empiriquement. En particulier, supposons que les enfants n'acquiescent jamais un mot composé sans acquiescer auparavant ou en même temps les mots qui le composent. Un tel fait est prévisible si l'on admet (CI) : comprendre un mot, c'est avoir associé un son et un concept. Un mot composé correspond à un concept complexe. Si les concepts satisfont (CI), il est impossible de maîtriser un concept complexe sans posséder ses composants. Donc quiconque acquiert une compétence linguistique avec un mot composé doit posséder les concepts exprimés par les mots le composant. Il n'a qu'à associer les concepts composants avec ces mots-ci. *Notons néanmoins qu'aucun argument précis et développé de cette sorte n'a été fourni.*

³⁸ Soulignons que Shanks est tout à fait cohérent : il rejette (CI) pour rejeter (CONST) (*modus tollens*). A ses yeux, les concepts ne satisfont pas (CONST) mais (COMB) (cf. plus loin). Il devrait donc répondre aux arguments empiriques avancés en faveur de la composition des concepts — ce qu'il ne fait pas.

³⁹ Dans un réseau de neurones, les activations d'unités correspondant aux représentations distribuées COFFEE et CUP WITH COFFEE se recourent en partie : c'est ce que Shanks appelle à la suite de Smolensky « micro-featural overlap ». L'exemple du café vient de Smolensky. Cependant la position de Shanks est ici différente de celle de Smolensky (1990) : Shanks ne retient que la première partie de la réponse à Fodor et Pylyshyn (1988), puisqu'il n'envisage pas la possibilité de structure implicite (cf. plus loin).

(par exemple les oiseaux) varient en fonction du contexte. Or Shanks admet avec la plupart des psychologues, mais non avec tous (Armstrong *et al.*, 1983 ; Rips, 1995 ; Leslie, 2000), le principe suivant :

(Ψ) — Les propriétés des activités et tâches cognitives mobilisant des concepts dépendent des propriétés des concepts (par exemple, des informations qu'ils contiennent).

(Ψ) a une conséquence épistémologique importante : les propriétés des activités cognitives révèlent plus ou moins directement les propriétés des concepts. Dans le cas présent, la dépendance contextuelle des tâches cognitives mobilisant des concepts est supposée révéler la dépendance contextuelle des propriétés constitutives des concepts : par exemple, puisque la typicalité des vaches n'est pas la même dans le contexte de production de lait et d'équitation et que les propriétés des concepts sont supposées, dans de nombreux modèles des concepts, rendre compte des jugements de typicalité, c'est que les propriétés des concepts varient d'un contexte à l'autre (Barsalou, 1987)⁴⁰. Cette dépendance contextuelle est étendue des contextes dans lesquels une tâche cognitive a lieu aux concepts avec lesquels un concept donné se compose : les propriétés d'un concept A sont supposées varier en fonction des concepts avec lesquels il se combine. Donc les propriétés de A dans AB ne sont pas celles de A dans AC.

4.1.2. Réponse à l'objection de Shanks

La dépendance contextuelle des tâches cognitives mobilisant les concepts est une découverte empirique robuste et importante. Pourtant, l'argument de Shanks ne tient pas.

Tout d'abord, en général, les données empiriques sur la dépendance contextuelle ne portent pas sur les propriétés d'un concept donné dans divers contextes. Elles concernent plutôt la manière dont un concept est appliqué étant données certaines informations ou certaines situations. Néanmoins, il n'est pas certain que l'extension de ce type de contexte au premier soit légitime. Admettons cependant que des données empiriques analogues aux données précédentes puissent être avancées en faveur de la variation des propriétés d'un concept donné dans différents contextes. La tâche suivante pourrait en effet fournir des données en faveur de cette variation. Elle consiste simplement à demander au sujet d'énoncer les propriétés qui caractérisent une table de cuisine et celles qui caractérisent une table de bureau. Il est probable que des propriétés différentes seraient énoncées.

D'autre part, l'argument de Shanks tel qu'il est formulé souffre d'un certain nombre de confusions. Si les propriétés constitutives du concept qui compose AB et celles du concept qui compose AC ne sont pas les mêmes, AB et AC n'ont pas pour constituant un unique et même concept A. C'est donc qu'il n'y a pas un seul et même concept CAFE, mais plusieurs : CAFE₁ pris isolément, CAFE₂ qui est un constituant de TASSE DE CAFE, CAFE₃ etc.. Certes, ces concepts partagent un

air de famille. Mais ce n'est pas le même concept pour autant. Cela revient donc à dire qu'aucun concept n'est le constituant d'un concept complexe. On peut donc, de manière polémique, poser la question suivante : en quel sens Shanks peut-il donc soutenir que le concept COFFEE est un élément de CUP WITH COFFEE (plus loin, il parle de la même manière de « partie ») ?

Une lecture plus charitable de l'objection de Shanks est possible. Les concepts changent en un sens de nature parce que la production d'un concept complexe à partir de concepts composants ne respecte pas le principe (CONST). Autrement dit, les concepts composants ne sont pas des constituants des concepts complexes. Au contraire, la production d'un concept complexe AB à partir des concepts A et B consisterait alors à produire un concept qui ne posséderait ni A ni B pour constituants (à la manière dont les prototypes se combinent — cf. (COMB) plus loin). Mais alors pourquoi chercher **au sein du concept complexe** TASSE DE CAFE un concept CAFE (comme le fait aussi Smolensky) ?

Cette objection *ad hominem* montre que Shanks ne peut avoir le beurre et l'argent du beurre : il ne peut soutenir à la fois que les concepts sont des constituants des concepts dont ils sont les composants et rejeter en même temps (CI).

Enfin, l'objection principale est la suivante : le principe (Ψ) consiste à admettre que les propriétés constitutives des concepts rendent compte, entre autres, des jugements de catégorisation et/ou de typicalité. Or c'est là l'hypothèse fondatrice de la théorie des prototypes. L'argument commet donc une pétition de principe : il attribue implicitement aux concepts une certaine nature afin d'établir des propriétés incompatibles avec (CI) et, par *modus tollens*, avec (CONST). Or attribuer aux concepts une telle nature, ce n'est rien d'autre que rejeter (CI).

4.2. (CS), (CCP) et productivité

Nous avons admis les deux thèses suivantes :

(13) $\diamond((CONST) \wedge \neg(CS))$

(14) $\diamond((CONST) \wedge \neg(CCP))$.

Donc ni (CS) ni (CCP) ne peuvent être dérivées de (CONST). Ces deux thèses supposent un détour par la productivité.

4.2.1. Les deux notions de productivité

Distinguons d'abord deux versions de la productivité :

La notion forte de productivité : l'esprit humain est capable d'avoir un nombre infini de pensées conceptuelles, par exemple de croyances.

La notion faible de productivité : l'esprit entretient de fait un très grand nombre de pensées.

La productivité forte peut être attribuée à l'esprit sur la base de la récursivité de certaines constructions syntaxiques des langues naturelles, comme c'est le cas, par exemple, de la relative (Chomsky) :

(16) *Un chien qui regarde un chien qui regarde un chien qui regarde un chien...*

⁴⁰ La conclusion inverse pourrait être tirée : l'ampleur de la dépendance contextuelle des tâches cognitives montre au contraire que le principe (Ψ) est très discutable.

En raison de ce type de structures, la syntaxe des langues naturelles permet la formation d'énoncés d'une complexité arbitraire, donc d'un nombre infini d'énoncés déclaratifs. Puisque les énoncés déclaratifs sont supposés exprimer des croyances des sujets, l'esprit doit lui-même être capable d'entretenir un nombre infini de croyances afin de pouvoir comprendre les croyances véhiculées par les énoncés d'autrui. En outre, si, de fait, les modes de composition de nos pensées sont récursifs, alors la pensée est *ipso facto* productive. De toute évidence, quelle que soit la manière dont la productivité forte de la pensée est établie, elle suppose une idéalisation : il faut distinguer la compétence conceptuelle du sujet de sa performance conceptuelle. Car en un sens sa capacité est finie. Les ressources nécessaires pour la formation des pensées conceptuelles, peut-être l'emploi de la mémoire de travail, sont finies : elles n'autorisent donc pas une complexité arbitraire de pensées. Autrement dit, les sujets ne peuvent former que des pensées conceptuelles d'une complexité limitée. Donc les facultés cognitives mises en jeu dans la composition des concepts n'autorisent pas la formation d'un nombre infini de pensées — mais pour une raison qui est extrinsèque à la capacité de composer des concepts. Si l'on fait abstraction de ces facultés cognitives ou si on prend en compte des facultés idéales, c'est-à-dire non limitées, les sujets sont en mesure d'entretenir un nombre infini de pensées⁴¹.

En revanche, la notion faible de productivité ne suppose pas cette idéalisation (peut-être contestable) : *de fait* nous entretenons un nombre gigantesque de pensées conceptuelles, par exemple de croyances.

4.2.2. L'argument à partir de la productivité

L'argument habituel en faveur de la compositionnalité des concepts fait appel à la notion forte de productivité (Fodor). De la capacité d'avoir un nombre infini de croyances, on infère en effet par abduction (CS). En effet, chaque pensée est caractérisée par un contenu différent, si bien que la productivité implique la capacité de penser à un nombre infini de choses. L'argument abductif est le suivant : la détermination fonctionnelle du contenu des concepts complexes et des croyances par le contenu des composants est la meilleure explication de la manière dont un système cognitif aux ressources finies peut entretenir un nombre infini de pensées dont le contenu est à chaque fois différent. *Le même argument permet de soutenir (CCP)*⁴².

Mais (CS) et (CCP) résultent aussi, par un raisonnement abductif identique, de la notion faible de productivité, qui reconnaît seulement un nombre très élevé de pensées possibles. Or, alors que l'infinité des pensées conceptuelles suppose une idéalisation, ce n'est pas le cas du nombre très élevé de pensées : notre

performance elle-même fournit donc un argument en faveur de (CS) et (CCP)⁴³.

4.2.3. Valeur de l'argument

Quel est le statut de cet argument ? Cet argument ne *prouve* pas que le système conceptuel est compositionnel aux premier et deuxième sens. C'est en effet un argument abductif : une thèse est affirmée dans la mesure où elle explique au mieux, étant donné toutes les explications actuellement envisageables, certaines données empiriques. La forme de l'abduction est la suivante :

$$(17) \quad q \\ p \rightarrow q \\ \therefore p$$

De ce fait, des explications alternatives, par exemple neurologiques, sont toujours possibles (Fodor, 1991, 1998, p. 96 ; Schiffer, 1987, 1991). Schiffer (1991) soutient ainsi la possibilité d'une explication de la productivité de la pensée (ainsi que sa systématité) qui ne suppose pas la compositionnalité sémantique des états mentaux⁴⁴. Nous n'aurions alors aucune raison de supposer que le système conceptuel est sémantiquement compositionnel.

Le principe de survenance des attitudes propositionnelles⁴⁵ affirme que chaque attitude, par exemple la croyance que *p*, dépend d'une propriété non sémantique Φ telle que l'occurrence de Φ est métaphysiquement (ou nomologiquement) suffisante pour l'occurrence de cette attitude. En outre, les croyances sont des RM conceptuelles complexes produites par la composition de concepts⁴⁶. Les concepts atomiques sont en nombre fini en raison du caractère fini de l'esprit. Il est possible que chaque concept atomique type possède de manière essentielle une propriété non sémantique Ψ (ou plus). Par ailleurs, il est possible que les bases de survenance des croyances soient déterminées à partir des propriétés non sémantiques des atomes et de la manière dont ces atomes sont composés (règles de composition). Chaque croyance possède donc une propriété non sémantique Φ nomologiquement ou métaphysiquement suffisante pour l'occurrence d'une croyance douée d'un contenu sémantique donné. Ces propriétés sont en nombre infini en raison de la productivité du système conceptuel. Néanmoins, puisque ces propriétés sont, par hypothèse,

⁴¹ De même pour la productivité de notre compétence linguistique.

⁴² Cet argument permet de soutenir également (CONST) : c'est d'ailleurs à cette fin qu'il est utilisé par Fodor et Pylyshyn (1988). L'argument de la productivité est aussi présent, sous une forme ou sous une autre, chez Katz et Fodor (1963) et Davidson (1965). Chez ces derniers, il est utilisé pour défendre la compositionnalité sémantique des langues naturelles.

⁴³ Un autre aspect de notre performance conceptuelle pourrait servir de prémisses à une inférence par abduction de la compositionnalité, notre capacité à avoir des pensées conceptuelles nouvelles. Aviez-vous déjà pensé aux vélos arctiques avant cette question ?

⁴⁴ L'argument de Schiffer est précisément dirigé contre la nécessité de reconnaître au mentalisme admis par Fodor une sémantique compositionnelle. J'adapte cet argument au présent contexte.

⁴⁵ Il y a plusieurs manières de définir la survenance. Par exemple, les états mentaux surviennent sur les états neuraux ssi tout individu qui posséderait les mêmes états neuraux possèderaient *ipso facto* les mêmes états mentaux.

⁴⁶ Puisque Schiffer tourne son attention vers le mentalisme, il admet implicitement (CONST) : les croyances ont pour constituants les concepts.

dérivées des propriétés, en nombre fini, des concepts atomiques, et des règles de composition — également en nombre fini — il n'y a pas incompatibilité entre la productivité de la pensée et le caractère fini de l'esprit. Par ailleurs, les propriétés des concepts atomiques ne sont pas des propriétés sémantiques — par hypothèse de nouveau. Par conséquent, les propriétés sémantiques des croyances ne sont pas dérivées des propriétés sémantiques des concepts : cette hypothèse explique la productivité des croyances sans attribuer au système conceptuel une sémantique compositionnelle.

La réponse de Fodor est pertinente (Fodor, 1991). Il concède qu'il existe des explications de la productivité de la pensée qui ne supposent aucune sémantique compositionnelle. Néanmoins, le problème est de savoir quelle est la meilleure explication de la productivité. Or l'hypothèse avancée est instable : on est tenté de voir dans les propriétés physiques attribuées aux concepts des bases de la survenance du contenu des concepts — ce qui revient à admettre la compositionnalité sémantique du système conceptuel. Qu'est-ce donc qui nous empêche de voir dans ces propriétés des propriétés sur lesquelles surviennent les propriétés sémantiques des concepts ? Schiffer a recours à des arguments indépendants contre la possibilité de sémantiques compositionnelles pour les systèmes linguistiques (et même plus généralement contre les systèmes des symboles) — arguments qu'il tire des problèmes rencontrés dans l'analyse sémantique des ascriptions d'attitudes propositionnelles (Schiffer, 1987)⁴⁷. La valeur de son argument dépend donc de la possibilité d'une telle sémantique compositionnelle. Notons que l'argument de Schiffer n'a de sens que si l'on contraint les propriétés sémantiques auxquelles il est permis de recourir dans l'interprétation des attitudes propositionnelles. Dans le cas contraire, le théorème de Zdrozny s'applique : puisque les constructions syntaxiques des attitudes propositionnelles sont récursives, l'on peut affirmer qu'il y a une sémantique compositionnelle.

Dans le présent contexte d'une analyse des concepts, le problème principal est néanmoins le suivant : les concepts sont par hypothèse des entités interprétées. Ces états mentaux possèdent des propriétés sémantiques — de manière paradigmatique une référence. Par ailleurs, les concepts composés, par exemple CHAT NOIR, sont également des états mentaux interprétés, alors que ce ne sont pas des croyances, au mieux des constituants de croyances. Or l'hypothèse avancée par Schiffer est incompatible avec l'interprétation des concepts : les croyances ont un contenu en raison de leur base de survenance, mais les constituants des croyances, les concepts, n'en ont pas — même s'ils déterminent la base de survenance des croyances dont ils sont les composants. Une position à la Schiffer sur la nature du système conceptuel est donc contradictoire⁴⁸.

⁴⁷ Ces arguments pourraient être avancés contre la possession d'une sémantique compositionnelle pour le système conceptuel : certaines de nos croyances sont des métareprésentations, donc, entre autres, des croyances sur d'autres croyances.

⁴⁸ En outre, toute donnée empirique relative à l'existence des concepts doués de propriétés sémantiques implique le rejet de l'explication suggérée par Schiffer.

En résumé, j'ai défendu dans cette quatrième partie (C1) contre les critiques de Shanks : il est d'une part incohérent de rejeter ce principe et de chercher au sein des concepts complexes les concepts composants ; d'autre part, c'est commettre une pétition de principe que de poser que les concepts varient d'un contexte à l'autre. Ensuite, j'ai montré pourquoi la productivité de la pensée nous autorisait par abduction à poser la compositionnalité sémantique et la compositionnalité des conditions de possession.

5. Notion de combinaison des prototypes et des théories

Soit le principe suivant :

(COMB) *Les prototypes et les théories se combinent en prototypes et en théories complexes.*

La notion de combinaison⁴⁹ est donc la suivante : à partir de deux prototypes (deux théories), un troisième prototype (une troisième théorie) peut être produit(e). À partir du prototype *POISSON* et du prototype *ANIMAL FAMILIER (PET)*, (COMB) affirme que l'on peut produire un prototype *POISSON FAMILIER* (Hampton, 1987, 1997a). (COMB) affirme la même chose des théories (Rips, 1995).

Tout comme (CONST), (COMB) décrit un processus de type R.

Cette thèse est empirique : la question « les prototypes se conforment-ils ou non à (COMB) ? » est une question factuelle.

5.1. Analyse de (COMB)

On peut distinguer quatre manières de comprendre la notion de combinaison par degré croissant d'exigence.

La première interprétation correspond au **NO Peeking Principle** de Rips (Rips, 1995) à propos de la combinaison des théories : le prototype complexe/la théorie complexe ne doit pas contenir d'informations acquises suite à notre expérience des objets auxquels ce prototype/cette théorie s'applique. En revanche, toutes les connaissances possédées antérieurement peuvent être mises à contribution pour constituer le prototype ou la théorie complexe.

(COMB 1) *La combinaison doit respecter le No Peeking Principle.*

Par exemple, le prototype *POISSON FAMILIER* viole le No Peeking Principle, parce qu'il contient des informations extraites de notre expérience des poissons familiers, à savoir que les poissons familiers sont en

⁴⁹ La combinaison de deux prototypes ou de deux théories pour former un troisième n'est que l'une des manières dont un nouveau prototype (théorie) peut être formé à partir d'un autre prototype (théorie) ou de plusieurs. En particulier, les prototypes peuvent être modifiés par les RM exprimées par des adjectifs modificateurs. Par exemple, le prototype *VOITURE* est modifié par la RM exprimée par l'adjectif « rouge », même si cette modification ne correspond pas à la combinaison de deux prototypes, puisque « rouge » n'exprime sans doute pas un prototype (Murphy, 1988, 1990 ; Smith et al., 1988).

général des poissons rouges. En conséquence, le prototype *POISSON FAMILIER* n'est pas un prototype complexe résultant de la combinaison de deux autres prototypes (cf. ci-dessous).

La deuxième interprétation de la notion de combinaison exige que les prototypes et les théories complexes résultent des **seuls** prototypes et théories composants. Elle renforce donc le No Peeking Principe de Rips : non seulement les informations ne doivent pas être dérivées des classes d'objets auxquels les prototypes (théories) complexes s'appliquent, mais en outre toutes les informations doivent être dérivées des prototypes et des théories qui sont composés. Si l'on obtient le prototype complexe *POISSON FAMILIER* par combinaison des prototypes *POISSON* et *ANIMAL FAMILIER*, toutes les informations contenues dans le premier prototype doivent être dérivées des informations contenues dans les deux autres prototypes. En conséquence, la combinaison ne peut être un processus holiste qui aurait recours à toutes les connaissances d'arrière-plan. En revanche (COMB 1) n'exige nullement cela : les informations qui composent le prototype (la théorie) complexe peuvent être dérivées de n'importe quelle connaissance antérieure.

Comme pour (COMB 1), le processus par lequel les prototypes et les théories complexes sont produits à partir des composants n'est pas spécifié : il n'a donc pas à obéir à un ensemble défini et fixe de règles. Autrement dit, à partir de deux prototypes ou de deux théories (ou davantage), le système cognitif doit être en mesure de projeter un troisième prototype ou une troisième théorie sans recourir aux connaissances d'arrière-plan. Mais cette projection peut s'apparenter davantage à une résolution de problèmes qu'à l'application d'un ensemble de règles prédéterminées. Par ailleurs, ces processus de combinaison peuvent être sensibles au contexte.

(COMB 2) *La combinaison est telle que le prototype (la théorie) complexe dépend **uniquement** des prototypes (des théories) composant(e)s.*

La troisième interprétation caractérise la combinaison par la notion de régularité. Cette notion consiste à renforcer (COMB 2) en spécifiant de manière abstraite la nature du processus de combinaison : il doit exister un ensemble défini de règles permettant de passer des composants au complexe (Hampton, 1997a). (COMB 3) caractérise donc la combinaison par deux conditions séparément nécessaires et conjointement suffisantes :

(COMB 3) *La combinaison doit être **régulière** :*

- elle obéit à un ensemble défini de règles ;
- et le prototype complexe dépend uniquement des prototypes composants. *Mutatis mutandis* pour les théories.

Bien sûr, différents types de règles sont possibles. Par ailleurs, on peut imaginer que le processus de combinaison se déroule en deux étapes, que la première soit régulière tandis que la seconde pourrait s'apparenter davantage à une résolution de problèmes.

La dernière interprétation de la combinaison exige l'existence d'une bijection entre les prototypes et les théories composants et les prototypes et les théories complexes. Plus précisément, pour tout triplet composé

de deux prototypes et d'un mode de combinaison α ($P1, P2, \alpha$) il existe un unique $P3$ obtenu par combinaison de $P1$ et $P2$ selon α et, inversement, pour tout prototype $P1$ résultant d'une combinaison, il existe un unique triplet de prototypes ($P2, P3, \beta$)⁵⁰ telle que la combinaison de $P1$ et $P2$ selon le mode β produise $P3$. Les autres versions de (COMB) n'exigent nullement que le prototype complexe corresponde ainsi à un unique triplet ainsi formé : un même prototype peut être produit par combinaison de différentes paires de prototypes selon un même mode de combinaison. *Mutatis mutandis* pour les théories.

(COMB 4) *La combinaison est régulière et bijective :*

- elle obéit à un ensemble défini de règles ;
- le prototype (la théorie) complexe dépend uniquement des prototypes (des théories) composant(e)s ;
- il existe une bijection entre les triplets composés de deux prototypes (théories) composant(e)s et d'un mode de combinaison et les prototypes (théories) composés(e)s.

Ces quatre interprétations de (COMB) semblent conceptuellement cohérentes. La combinaison des prototypes et des théories se conforme-t-elle à (COMB 1), à (COMB 2) à (COMB 3) ou à (COMB 4)⁵¹ ? La réponse est bien sûr empirique (cf. F/).

5.2. Combinaison des prototypes et des théories vs composition des concepts

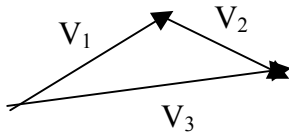
5.2.1. (CONST) et (COMB)

La principale différence entre (CONST) et (COMB) est la suivante : d'après les quatre analyses proposées de (COMB), à la différence des RM conceptuelles complexes, ni les prototypes complexes ni les théories complexes n'ont pour **constituants** les prototypes et les théories composés (au sens de (CCo)).

Une analogie avec la combinaison vectorielle éclaire la distinction entre (CONST) et (COMB). Soient deux vecteurs V_1 et V_2 qui se combinent en un vecteur V_3 :

⁵⁰ Éventuellement un quadruplet si le mode de combinaison suppose trois prototypes (théories).

⁵¹ L'approche PTC du connexionnisme, défendue par Smolensky (« *Proper Treatment of Connectionism* »), caractérise la composition des RM (représentées dans un réseau par un vecteur d'activation d'unités) de manière conforme à (COMB 4). Les RM atomiques sont des vecteurs d'activation. Les différentes places dans une structure de RM (ce qui correspond aux modes de combinaison), par exemple dans un arbre de construction syntaxique, sont aussi codées par des vecteurs d'activation. On code le fait qu'une RM atomique occupe une place donnée dans une structure par une opération (appelée « *tensor product* ») sur les vecteurs d'activation correspondant à la RM combinée et à la place qu'elle occupe dans la structure. On code la combinaison de plusieurs RM en additionnant leurs vecteurs. Dans certaines conditions, mais non dans toutes, un vecteur complexe correspond de manière bijective à un nombre donné de RM atomiques et à leur mode de combinaison (représenté par la structure). On a donc là une réalisation formelle de (COMB 4) (Smolensky, 1988, 1991, 1992 ; Van Gelder, 1990).



V_3 résulte bien de la combinaison de V_1 et V_2 . Mais V_3 n'a pas pour constituants les deux vecteurs dont il dérive par combinaison : quand V_3 est dessiné sur le papier, ni V_1 ni V_2 n'ont à l'être. En conséquence, comme les prototypes complexes, V_3 peut exister sans V_1 ni V_2 . En un sens analogue, la combinaison des prototypes n'implique pas (CI) : un individu peut posséder un prototype A identique *qua* type (autrement dit, contenant les mêmes informations) au prototype que l'on peut obtenir par combinaison de deux autres prototypes, B et C , alors même qu'il ne possède ni B ni C , *a fortiori* qu'il ne les a pas combinés. En outre, V_3 est produit conformément aux règles de la combinaison vectorielle (celle-ci satisfait l'exigence de régularité). La combinaison vectorielle est donc précisément l'analogue de (COMB 3).

5.2.2. Spécificité de (COMB 4)

Néanmoins, il y a une différence importante entre (COMB 4) et les trois premières analyses de (COMB). S'il est vrai que d'après (COMB 4), les prototypes combinés ne sont pas des constituants des prototypes complexes au sens de (CCo), néanmoins, les RM qui se combinent selon (COMB 4) possèdent une structure implicite. En effet, puisque le prototype complexe correspond à une unique paire de prototypes composés selon un mode donné de combinaison, l'identité du prototype complexe encode les prototypes à partir desquels il a été produit ainsi que la manière dont il a été produit (le mode de combinaison) : les composants sont encodés au sens où pour un prototype complexe, il y a un unique triplet (P_1, P_2, α). En revanche, aucune des trois autres combinaisons ne produit des RM complexes douées d'une structure implicite.

Comparons (COMB 4) et (CONST). Dans le cas des RM conceptuelles complexes, les constituants sont coactivés avec l'activation de la RM complexe. Les concepts complexes peuvent donc servir d'*inputs* à des processus mentaux sensibles à la structure des RM. En revanche, si certaines RM se combinent selon (COMB 4), lorsqu'une RM complexe est activée, les RM composantes ne sont pas activées par là-même. Donc les RM ne peuvent prendre part à des processus mentaux sensibles à la structure explicite des RM. En conséquence, les processus mentaux ne peuvent être sensibles à la structure explicite des RM⁵². En outre, à la différence de (CONST) et (CCo), (COMB 4) n'implique pas (CI) : puisque les RM combinées selon ce principe ne sont pas coactivées, la possession d'une RM conceptuelle qui *pourrait* être obtenue par combinaison de deux RM n'implique pas la possession de ces deux RM. (COMB 4) affirme seulement que si une RM Φ a été combinée, alors elle a été produite à partir d'un unique triplet ou éventuellement quadruplet. Mais ce principe n'exclut nullement qu'un individu possède Φ sans l'avoir

obtenu par combinaison. Cela n'est pas exclu par (COMB 4) puisque les RM composantes ne sont pas des constituants de Φ . Bien sûr, cela est impossible pour les RM qui satisfont (CONST) et (CCo). Bref, en un sens, à la différence de (COMB 1), (COMB 2) et (COMB 3), une RM combinée conformément à (COMB 4) encode les RM composées et la manière dont elles ont été combinées ; mais, à la différence de (CONST) et (CCo), les RM composantes ne sont pas explicitement encodées par la RM complexe.

Reprenons l'analogie avec la combinaison vectorielle et supposons (faussement) qu'elle satisfasse un principe analogue à (COMB 4). En conséquence, il n'y aurait qu'une seule manière de produire V_3 à partir d'autres vecteurs, et donc qu'une seule manière de décomposer ce vecteur. Cela n'est plus le cas si la combinaison vectorielle satisfait des principes analogues aux autres interprétations de (COMB) : un même vecteur peut alors résulter de la combinaison de différents couples de vecteurs. Si cela est le cas, la décomposition d'un vecteur donné n'a pas de sens : il n'y a pas un unique couple de vecteurs qui lui corresponde. Pareillement, si la combinaison des prototypes et des théories satisfait (COMB 4), les prototypes et les théories complexes ne peuvent pas être décomposés. En revanche, si la combinaison des prototypes et des théories satisfait (COMB 4), les prototypes et les théories complexes peuvent être décomposés⁵³.

En résumé, j'ai distingué dans cette cinquième partie quatre types de combinaison. Toutes les combinaisons respectent le No Peeking Principle, alors que les combinaisons peuvent ou non être des processus holistes et réguliers. J'ai aussi montré la spécificité du quatrième type de combinaison : les RM complexes qui en résultent encodent les RM composantes sans que celles-ci les constituent ; elles possèdent donc une structure implicite.

6. Les prototypes et les théories se combinent-ils ? Et comment ? Données empiriques

6.1. Rejet de (COMB 4)

Dans la littérature psychologique sur les prototypes et les théories, on ne trouve guère de défense de la combinaison selon la ligne présentée par (COMB 4). On a là une différence entre les partisans des prototypes et des théories mentales et certains connexionnistes, comme Smolensky : alors que ces derniers ont essayé de montrer comment les représentations des réseaux de neurones étaient susceptibles d'avoir une structure implicite, les psychologues de la catégorisation ont cherché à montrer comment les prototypes et les théories pouvaient être l'objet d'un processus de type R.

⁵² C'est là l'une des caractéristiques des réseaux connexionnistes selon l'interprétation PTC : il faut distinguer le niveau de l'interprétation sémantique du niveau causal.

⁵³ Néanmoins, les processus décomposant les complexes ne pourraient être guidés par une structure explicite. (PTC) fournit un exemple de processus mentaux permettant la décomposition de pensées alors même que les pensées complexes n'ont pas leurs composants pour constituants.

L'idée que les prototypes et les théories complexes puissent avoir une structure implicite n'est donc pas défendue.

Les données empiriques ne semblent en outre pas soutenir l'existence d'une structure implicite. Autrement dit, elles ne soutiennent pas l'idée que les prototypes et les théories se composent conformément à (COMB 4). En effet, la combinaison des prototypes semble dépendre, en partie au moins, des connaissances d'arrière-plan des sujets. Elles interviennent en particulier lorsque les prototypes composés, par exemple *NATIONALISTE* et *PACIFISTE*, représentent des propriétés incompatibles ou difficilement conciliables. Les connaissances d'arrière-plan permettent de résoudre l'incompatibilité (Medin et Shoben, 1988 ; Hampton, 1997a). Or (COMB 4) est incompatible avec un processus holiste.

6.2. Les prototypes et les théories n'ont pas de constituants : la combinaison est distincte de la composition

Cette partie de la réponse à la question posée est négative : si les prototypes et les théories se conforment à (COMB), alors ils ne doivent pas avoir de constituants. Est-ce le cas ?

Examinons le cas des prototypes. De fait, ils ne satisfont pas (CI). Les prototypes sont des ensembles d'informations relatives aux propriétés fréquentes d'une classe d'objets. Ces informations sont en général extraites des interactions avec les objets de cette classe. Par conséquent, un individu qui aurait des expériences *sui generis* pourrait acquérir directement (vs. par combinaison) un prototype relatif à une classe d'objets AB que l'on pourrait définir comme l'intersection d'une classe d'objets A et d'une classe d'objets B — prototype identique à celui qui pourrait être construit par combinaison de deux autres prototypes, l'un relatif aux A et l'autre aux B. Il est par exemple concevable qu'un individu possède un prototype *POISSONS FAMILIERS* identique au prototype qui *pourrait* être construit, conformément à l'une des interprétations de (COMB), à partir de *POISSON* et *ANIMAL FAMILIER*, sans posséder pour autant un prototype de poissons ou un prototype des animaux familiers. De même pour les théories.

Puisque (CONST)→(CI), on rejette par *modus tollens* (CONST) pour les prototypes et les théories complexes : les concepts n'ont pas de constituants.

6.3. Les prototypes et/ou les théories se combinent.

La deuxième partie de l'argument est positive : il est possible de produire des prototypes et des théories complexes conformément à l'une des versions de (COMB).

1. La plupart des prototypes et des théories **apparemment** complexes ne le sont pas en réalité. Je fais ici référence aux prototypes et aux théories qui **semblent** résulter de la combinaison d'autres prototypes, par exemple le prototype des poissons familiers. Mais de tels prototypes contiennent des propriétés qui ne peuvent être dérivées ni des

prototypes apparemment composés (Hampton appelle ce type de propriétés « des propriétés émergentes » (Hampton, 1997a)) ni des connaissances d'arrière-plan. En fait, ces prototypes et ces théories sont pour la plupart dérivés de notre expérience des objets eux-mêmes. Par exemple, le prototype de poisson familier est obtenu à partir des poissons rouges, et non par combinaison du prototype de poissons et du prototype des animaux familiers : c'est ce que Hampton appelle le *feedback extensionnel* (Hampton, 1988ab, 1997a).

2. Nous sommes capables de catégoriser ou de décrire des entités inconnues comme des étudiants d'Oxford qui sont des ouvriers (cet exemple vient de Hampton et Dillane, 1993). Ces classes d'entités ont la propriété d'être l'intersection de classes pour lesquelles il existe des prototypes et/ou des théories indépendantes (c'est-à-dire un prototype/théorie relatif aux étudiants d'Oxford et un prototype/théorie relatif aux ouvriers). En outre, antérieurement à la tâche expérimentale (décrire ou catégoriser), l'objet est inconnu (qui a déjà rencontré un étudiant d'Oxford qui est un ouvrier ?) : nous ne disposons donc ni d'un prototype ni d'une théorie relatifs à cette classe d'objets. Tout d'abord, puisqu'il est communément admis que la catégorisation repose sur des prototypes ou sur des théories, notre capacité de catégoriser des entités au sein de la classe des étudiants d'Oxford qui sont des ouvriers suggère l'existence de certaines procédures permettant de déterminer un prototype ou une théorie complexe à partir des connaissances antérieurement possédées (Rips, 1995). De même, notre capacité de décrire des entités inconnues signifie que nous sommes capables de produire des informations relatives à ces classes d'entités à partir des connaissances que nous possédons. Il existe donc un prototype ou une théorie s'appliquant aux entités inconnues alors même qu'aucune information n'a pu être dérivée des objets eux-mêmes (parce qu'ils sont inconnus). La production de ce principe ne viole donc pas le *No Peeking Principle* : en résumé, les prototypes et les théories semblent au moins satisfaire (COMB 1).

Dans les deux cas, **deux questions (empiriques) restent ouvertes** :

- Les connaissances à partir desquels les prototypes et les théories complexes sont produites sont-elles dérivées exclusivement des prototypes et des théories correspondant aux classes intersectées (par exemple la classe des étudiants d'Oxford et la classe des ouvriers) ?
- Ces procédures sont-elles ou non des règles ?

Une réponse positive à la première question est une condition suffisante pour que les prototypes et les théories puissent satisfaire (COMB 2), et une réponse positive à la troisième question est une condition suffisante pour que les prototypes et les théories puissent satisfaire (COMB 3).

En résumé, j'ai montré que les prototypes se combinent et que leur combinaison ne correspond pas à (COMB 4).

7. Le meilleur argument contre les assimilations (A) et (B)

Si mes analyses conceptuelles sont correctes (1., 2., 5.) et si, de manière factuelle, les concepts, les théories et les prototypes les satisfont (3., 4., 6.), alors le rejet des assimilations (A) et (B) repose sur la distinction entre la composition des concepts et la combinaison des prototypes et des théories : le détour par la compositionnalité semble inutile.

Conclusion

Quels sont les points cruciaux de cette analyse ?

Tout d'abord, il convient de distinguer (au moins) deux processus par lesquels une RM complexe peut être produite à partir de RM composantes, à savoir la combinaison et la composition. Dans le second cas, mais non dans le premier, les composants constituent le complexe, au sens où ils sont nécessairement coactivés quand le complexe l'est.

Ensuite, ces processus peuvent posséder trois types de compositionnalité, la compositionnalité sémantique, la compositionnalité des conditions de possession et la compositionnalité inverse. Selon la compositionnalité sémantique, les propriétés sémantiques des représentations mentales complexes sont une fonction des propriétés sémantiques des représentations composantes et de la règle de composition ; selon la compositionnalité des conditions de possession, les conditions de possession d'une représentation complexe sont intégralement déterminées par les conditions de possession des représentations composantes ; selon la compositionnalité inverse, les conditions de possession des représentations composantes sont aussi des conditions de possession des représentations complexes. La compositionnalité sémantique et la compositionnalité des conditions de possession ne sont ni impliquées par ni inconsistantes avec la composition et la combinaison. En revanche, la compositionnalité inverse est impliquée par la composition et inconsistante avec la combinaison.

En outre, le raisonnement conceptuel requiert la composition des concepts au sein de RM conceptuelles complexes alors que la possibilité de catégoriser et de décrire des objets inconnus et la violation de la compositionnalité inverse suggèrent que les prototypes se combinent.

Enfin, le rejet des assimilations (A) et (B) ne repose pas sur la compositionnalité, mais sur la distinction de la composition et de la combinaison.

Références bibliographiques

- [Armstrong et al., 1983] Armstrong S. L., Gleitman L. R., Gleitman H. (1983). What some concepts might not be?. *Cognition*. 13, 263-308.
- [Barsalou, 1987] Barsalou L. W. (1987). The instability of graded structures: implications for the nature of concepts. In Vosniadou S. et Ortony, A. (eds.) *Similarity and Analogical Reasoning*, CUP : Cambridge.
- [Carey, 1985] Carey S. (1985). *Conceptual Change in Childhood*. MIT Press : Cambridge, Mass.

[Carey et Johnson, 2000] Carey S., Johnson S. C. (2000). Metarepresentations and Conceptual Change: Evidence from Williams Syndrome. In Sperber D. (ed.) *Metarepresentations*. OUP : Londres.

[Davidson, 1965] Davidson D. (1965). Theories of Meaning and Learnable Languages. Repris dans Davidson D. (1984). *Inquiries into Truth and Interpretation*. OUP : Oxford.

[Evans, 1982] Evans G. (1982). *The Varieties of Reference*. Ed. J. McDowell. Clarendon Press, OUP : Oxford.

[Fodor, 1991] Fodor J. A. (1991). Replies. In Loewer B. et Rey G. (eds.) (1991). *Meaning in Mind, Fodor and his Critics*. Blackwell : Cambridge, Mass..

[Fodor, 1998] Fodor J. A. (1998). *Concepts*. OUP : Oxford.

[Fodor, 1999a] Fodor J. A. (1999). There are no Recognitional Concepts : not even Red. In Fodor J. A. (1999d).

[Fodor, 1999b] Fodor J. A. (1999). There are no Recognitional Concepts: not even Red; Part 2: The Plot Thickens. In Fodor J. A. (1999d).

[Fodor et Lepore, 2001] Fodor J. A., Lepore E. (brouillon). Why Compositionality Won't Go Away: Reflections on Horwich's "Deflationary" Theory.

[Fodor et Pylyshyn, 1988] Fodor J. A., Pylyshyn Z. W. (1988). Connectionism and cognitive architecture: a critical analysis. *Cognition*. 28, 3-71.

[Gopnik et Meltzoff, 1998] Gopnik A., Meltzoff A. (1998). *Words, Thoughts and Theories*, MIT Press : Cambridge, Mass.

[Hampton, 1987] Hampton J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory and Cognition*. 15, 55-71.

[Hampton, 1988a] Hampton J. A. (1988). Disjunction of natural concepts. *Memory and Cognition*. 16, 579-91.

[Hampton, 1988b] Hampton J. A. (1988). Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology : Learning, Memory, and Cognition*. 14, 12-32.

[Hampton, 1997a] Hampton J. A. (1997a). Conceptual combinations. In Lamberts K. et Shanks D. (eds.) (1997). *Knowledge, Concepts and Categories*. MIT Press : Cambridge, Mass.

[Hampton, 1997b] Hampton J. A. (1997b). Psychological representations of concepts. In Conway M. A. (ed.) (1997). *Cognitive Models of Memory*. MIT Press : Cambridge, Mass.

[Hampton, 2000] Hampton J. A. (2000). Concepts and Prototypes. *Mind and Language*. 15, 2-3, 299-307.

[Hampton et Dillane, 1993] Hampton J. A., Dillane M. (1993). Taking a point of view on attributes inheritance. Papier présenté à la 34th Annual Convention of the Psychonomic Society, Washington, D.C.

[Hodges, 1998] Hodges W. (1998). Compositionality is not the problem. *Logic and Philosophy*. 6, 7-33.

[Horgan, 1998] Horgan T. (1998). Recognitional Concepts and the Compositionality of Concept Possession. In Villanueva E. (ed.).

[Horwich, 1998] Horwich P. (1998). Concept Constitution. In Villanueva E. (ed.) (1998).

[Horwich, 1999] Horwich P. (1999). *Meaning*. OUP : Oxford.

[Janssen, 1997] Janssen M. V. Th. (1997). Compositionality. In van Benthem J. et ter Meulen A. (eds.). *Handbook of Logic and Language*. Elsevier Science B.V.

[Kamp et Partee, 1995] Kamp H., Partee B. (1995). Prototype theory and compositionality. *Cognition*. 57, 129-191.

- [Katz et Fodor, 1963] Katz J. J., Fodor J. A. (1963). The Structure of a semantic Theory. *Language*. 39, 170-210. Repris dans Fodor J. A. et Katz J. J. (eds.) (1964). *The Structure of Language, Readings in the Philosophy of Language*. Prentice-Hall, Inc. : Englewood Cliffs, New Jersey.
- [Leslie, 2000] Leslie A. M. (2000). How to acquire a representational theory of mind. In Sperber D. (ed.) (2000). *Metarepresentations*. OUP : Londres.
- [Malt, 1994] Malt B. C. (1994). Water is not H₂O. *Cognitive Psychology*. 27, 42-70.
- [Medin et Shoben, 1988] Medin D. L., Shoben E. J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*. 20, 158-90.
- [Millikan, 1990] Millikan R. G. (1990). Compare and Contrast Dretske, Fodor, and Millikan on Teleosemantics. Repris in Millikan R. G. (1993a).
- [Millikan, 1993a] Millikan R. G. (1993a). *White Queen Psychology and Other Essays for Alice*. MIT Press : Cambridge, Mass.
- [Millikan, 1993b] Millikan R. G. (1993b). On Mentalese Orthography. In Dahlbom D. (ed.) (1993). *Dennett and his Critics*. Blackwell Publishers : Oxford. Repris en partie in Millikan R. G. (1993a).
- [Millikan, 1998] Millikan R. G. (1998a), A More Plausible Kind of "Recognitional Concept". In Villanueva E. (ed.) (1998).
- [Murphy, 1988] Murphy G. L. (1988). Comprehending complex concepts. *Cognitive Science*. 12, 529-62.
- [Murphy, 1990] Murphy G. L. (1988). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*. 29, 259-88.
- [Osherson et Smith, 1981] Osherson D. N., Smith E. E., (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition*. 11, 35-58.
- [Osherson et al., 1990] Osherson D. N., Smith E. E., Wilkie O., Lopez A., Shafir, E. (1990). Category-based induction. *Psychological Review*. 97, 185-200.
- [Partee, 1995] Partee B. H. (1995). Lexical Semantics and Compositionality. In Gleitman L. et Liberman M. (eds.) (1995). *An Invitation to Cognitive Science, Language, vol. 1*. MIT Press : Cambridge, Mass.
- [Peacocke, 1992] Peacocke Ch. (1992). *A Study of Concepts*. MIT Press : Cambridge, Mass.
- [Peacocke, 2000] Peacocke Ch. (2000). Fodor on Concepts: Philosophical Aspects. *Mind and Language*. 15, 2-3, 327-340.
- [Rey, 1983] Rey G., (1983). Concepts and Stereotypes. *Cognition*. 15, 237-262.
- [Rips, 1995] Rips L. J. (1995) The current status of research on concept combination. *Mind and Language*. 10, 72-104.
- [Schiffer, 1987] Schiffer St. (1987). *Remnants of Meaning*. MIT Press : Cambridge, Mass.
- [Schiffer, 1991] Schiffer St. (1991). Does Mentalese Have a Compositional Semantics ?. In Loewer B. et Rey G. (eds.) (1991). *Meaning in Mind, Fodor and his Critics*. Blackwell : Cambridge, Mass.
- [Shanks, 1997] Shanks D. R. (1997). Distributed representations and implicit knowledge : A brief introduction. In Lamberts K. et Shanks D. (eds.) (1997).
- [Smith et al., 1988] Smith E. E., Osherson D. N., Rips L. J., Keane M. (1988). Combining prototypes: a selective modification model. *Cognitive Science*. 12, 485-527.
- [Smolensky, 1988] Smolensky P. (1988). On the proper treatment of connectionism. *Behavioural and Brain Sciences*. 11, 1-74.
- [Smolensky, 1991] Smolensky P. (1991). Connectionism, Constituency and the Language of Thought. In Loewer B. et Rey G. (eds.) (1991).
- [Smolensky, 1992] Smolensky P. (1992). IA Connexionniste, IA symbolique et cerveau. In Andler D. (ed.). *Introduction aux sciences cognitives*. Gallimard : Paris. Publié en 1987. *Artificial Intelligence Review*. 1, 95-109.
- [Van Gelder, 1990] Van Gelder (1990). Compositionality: a connectionist variation on a classical theme. *Cognitive Science*. 14, 355-384.
- [Zadrozny, 1994] Zadrozny W. (1994). From Compositional Semantics to Systematic Semantic. *Linguistics and Philosophy*. 17, 329-342.

L'auteur

Édouard Machery est Allocataire de recherches Moniteur dans le département de philosophie de l'Université de Paris-Sorbonne (Paris IV) où il enseigne la logique. Ancien élève de l'École Normale Supérieure de la rue d'Ulm et agrégé de philosophie, il rédige actuellement une thèse de philosophie intitulée « Qu'est-ce qu'un concept ? La notion de concept en philosophie de l'esprit et en psychologie cognitive » sous la direction de Daniel Andler. Il a été récemment *Visiting Graduate Student* dans le département de philosophie de l'Université de Rutgers (EU) à l'invitation de Stephen Stich. Il est particulièrement intéressé par les théories de la catégorisation et du raisonnement, par la psychologie développementale, et aussi, récemment, par la psychologie évolutionniste.

