

CHAPITRE 2

La reconnaissance de formes

Vous êtes un génie créatif.

Votre génie est à ce point parfait qu'il vous apparaît,
ainsi qu'aux autres, comme naturel.

Pourtant, il devance largement les efforts les plus vaillants des ordinateurs actuels
les plus puissants. Pour l'invoquer il vous suffit d'ouvrir les yeux.

Donald HOFFMAN (1998)

SOMMAIRE

1. La description des formes 34
2. Les étapes du traitement de l'information 44
3. La reconnaissance de mots 49

La citation qui ouvre ce chapitre est tirée du premier paragraphe de l'ouvrage de Donald Hoffman, Visual Intelligence. Elle fait référence à notre remarquable aptitude à reconnaître des formes, une aptitude qui est encore loin d'être reproduite par les ordinateurs les plus puissants du monde. Il nous est difficile d'apprécier cette capacité parce qu'elle semble tout à fait naturelle. Avec un peu de chance, vous l'apprécierez davantage après avoir pris connaissance des progrès accomplis par les scientifiques de la cognition dans la recherche sur notre capacité à réaliser une tâche aussi difficile.

L'étude de la reconnaissance de formes[♦] est surtout une étude qui porte sur la façon dont nous identifions les objets dans leur environnement. Notre aptitude à reconnaître des formes devrait impressionner, compte tenu du nombre de variantes pour une même forme. Par exemple, chaque lettre de l'alphabet est une forme en soi. La figure 2.1 montre diverses écritures manuscrites. Manifestement, tout le monde n'écrit pas de la même façon et certaines écritures sont beaucoup moins lisibles que d'autres. Cependant, sauf à être tout à fait illisibles, nous parvenons en général à les lire, c'est-à-dire que nous reconnaissons les mots.

Notre supériorité sur les ordinateurs, en tant que « reconnaiseurs » de formes, présente l'avantage pratique que la reconnaissance de formes peut servir de test pour déterminer qui tente d'accéder à Internet: une personne ou un logiciel? Si vous avez

◆ **reconnaissance de formes**

L'étape perceptive au cours de laquelle un stimulus est identifié.

Figure 2.1

Differentes écritures manuscrites.
Source: *Man machine engineering*,
A. Chapanis. Copyright © 1965
de Brooks/Cole, une section
de Thomson Learning.

We all read different styles of handwriting so easily and so commonly that it is easy for us to overlook what an extraordinary ability this is. Note the extreme differences in the way different people write certain letters of the alphabet. Now consider what kind of a machine would be necessary to "recognize" all these letters. In part, we are able to read these samples of handwriting because of the context and redundancy in the passage. But to a page dyer, our ability to read this passage is also due to the remarkable capacity the human organism has for "perceptual generalization".

l'habitude de passer du temps sur Internet, il vous est sans doute arrivé de devoir identifier un mot déformé pour être autorisé à accéder à un site. S'il nous est facile d'identifier ce mot, pour les ordinateurs, il n'en est rien (voir « Dans la Presse » 2.1).

La facilité et la précision avec lesquelles nous reconnaissions les formes rendent difficile l'étude de cette aptitude. Il n'est ni intéressant ni révélateur que quelqu'un identifie aisément une grande variété de formes. Pour compliquer la tâche, les psychologues procèdent souvent à une présentation très rapide des formes. Lorsque les formes ne sont présentées que quelques millisecondes, les sujets se trompent et les psychologues étudient les types d'erreurs commises.

Une grande partie de la littérature sur la reconnaissance de formes est consacrée aux différentes façons de les décrire. La première section de ce chapitre porte sur trois types de descriptions qui représentent des théories différentes de la reconnaissance de formes. La deuxième section aborde les modèles de traitement de l'information dans le cadre de la reconnaissance de formes visuelles. Nous examinerons plus en détail la recherche de Sperling et la façon dont ses résultats ont influencé les théories ultérieures. La troisième section traite de la reconnaissance de mots et sera pour nous l'occasion de considérer quelques facteurs influençant la lecture.

1. La description des formes

Considérons les explications suivantes au sujet de la façon dont nous reconnaissions les formes. Notre mémoire à long terme (MLT) contient les descriptions de nombreux types de formes. Lorsque nous voyons ou entendons une forme, nous en construisons une description que nous comparons à celles qui sont stockées dans notre MLT. Nous sommes capables de reconnaître la forme si sa description correspond en tout point à une description stockée dans la MLT. Bien que cette explication soit plausible, elle est plutôt vague. Par exemple, sous quelles formes se présentent ces descriptions ? Examinons trois explications qui ont été suggérées : (1) la description des gabarits, (2) la description des traits et (3) la description structurelle.

1.1 *Les théories des gabarits*

Selon les théories des gabarits, les formes ne sont absolument pas « décrites ». Bien plutôt, les *gabarits*[♦] sont des entités holistiques ou non analysées que nous comparons avec d'autres formes, en évaluant dans quelle mesure les formes se recouvrent. Imaginez que vous découpez une série de lettres dans du carton. Si vous avez réalisé un découpage pour chaque lettre de l'alphabet et que je vous donne le découpage d'une lettre que j'ai réalisé moi-même, vous pourrez mesurer jusqu'à quel point ma lettre recouvre chacune des vôtres – les gabarits. L'identité de ma lettre sera déterminée par celui des gabarits possédant le degré de recouvrement le plus élevé. Le même principe serait d'application si vous remplacez vos lettres en carton par une image visuelle de chacune d'elles et que vous utilisiez les images pour effectuer des comparaisons mentales.

Utiliser le degré de recouvrement comme mesure de la reconnaissance de formes présente de nombreux problèmes. Pour commencer, la comparaison impose que le gabarit soit dans la même position, dans la même orientation et de la même taille que la forme que vous cherchez à identifier. Donc, la position, l'orientation et la taille des gabarits devraient être continuellement ajustées, de façon à correspondre à la position, l'orientation et la taille de chaque forme que vous voulez reconnaître. Un deuxième problème est la grande variabilité des formes, comme nous l'avons constaté dans la figure 2.1. Il serait difficile de construire, pour chaque lettre, un gabarit qui correspondrait à toutes les variantes de cette lettre. Troisièmement, une théorie des gabarits ne révèle pas de quelle manière deux formes diffèrent. Nous pourrions apprendre d'une telle théorie que les lettres capitales *P* et *R* sont similaires parce que l'une recouvre l'autre en grande partie. Mais pour savoir comment les deux lettres diffèrent, nous devons être capables d'analyser ou de décrire les lettres. Par contre, la théorie des traits, abordée dans la section suivante, nous permet d'analyser les formes dans leurs parties. Un quatrième problème est qu'une théorie des gabarits n'autorise aucune description alternative d'une forme. Il se peut que vous ayez vu des figures ambiguës se prêtant à plus d'une interprétation comme, par exemple, une jeune ou une vieille dame. Les deux interprétations sont basées sur des descriptions différentes : par exemple, la bouche de la vieille dame serait le collier de la plus jeune.

♦ gabarit (ou patron)

Une forme non analysée qui est comparée à des formes alternatives, en utilisant le degré de recouvrement comme mesure de la ressemblance.

Ces faiblesses de la théorie des gabarits la rendent peu prometteuse en tant que théorie générale de la reconnaissance de formes et elle est généralement vite abandonnée. Toutefois, il est des situations où elle peut fournir un modèle utile. Souvenez-vous qu'au chapitre 1 nous avons vu que le *registre sensoriel** conserve un bref moment les informations sensorielles, de sorte que l'observateur dispose de plus de temps pour reconnaître les formes. Mais comment les formes sont-elles conservées dans le registre sensoriel si elles ne sont pas reconnues ? Une possibilité serait que les formes soient représentées comme gabarits non analysés dont les traits sont étudiés par la suite au niveau de la reconnaissance de formes.

Cette interprétation du registre sensoriel est présentée très clairement par Phillips (1974). Dans l'expérience de Phillips, les sujets voyaient des formes réalisées par un remplissage aléatoire des cellules d'une matrice carrée. La première forme était présentée durant 1 seconde et était suivie, après un intervalle de temps variable, par une forme, soit identique, soit similaire. La tâche du sujet consistait à décider aussi vite que possible si les deux formes étaient les mêmes ou si elles étaient différentes. Dans une moitié des épreuves, la seconde forme apparaissait exactement au même endroit que la première. Puisque la seconde forme était parfaitement superposée à l'image sensorielle de la première, le registre sensoriel pouvait être utilisé pour établir la correspondance entre les gabarits. Dans l'autre moitié des épreuves, les participants se prononçaient également sur l'identité ou la différence des deux formes, mais la seconde forme était déplacée horizontalement d'une cellule. Cette légère modification devait empêcher la correspondance entre les gabarits puisque les deux formes n'étaient pas correctement alignées.

La figure 2.2 présente les résultats de l'expérience de Phillips. L'*intervalle interstimulus** – le temps écoulé entre les deux formes – était de 20, 60, 100, 300 ou 600 msec (millisecondes). Lorsque les deux formes étaient présentées aux mêmes endroits (condition « identique » dans la figure 2.2), la précision diminuait à mesure que l'intervalle interstimulus augmentait. Ce résultat suggère que les sujets utilisaient un registre sensoriel se dégradant rapidement. Lorsque la seconde forme était

◆ **registre sensoriel**

Partie de la mémoire qui, le temps d'une fraction de seconde, conserve l'information sensorielle non analysée, permettant ainsi un éventuel examen après la fin de l'excitation physique par un stimulus.

◆ **intervalle interstimulus**

Le temps écoulé entre la fin d'un stimulus et le début du suivant.

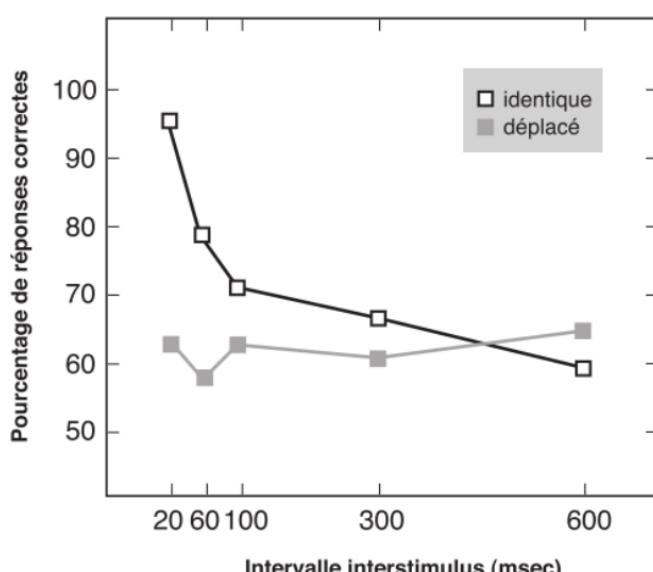


Figure 2.2

Pourcentage de réponses correctes pour une matrice 8x8 en fonction de l'intervalle interstimulus et du déplacement.

Source : «On the distinction between sensory stage and short-term visual memory», W.A. Phillips, 1974, *Perception and Psychophysics*, 16, 283-289. Copyright © 1974 de la Psychometric Society. Reproduction autorisée.

déplacée, les sujets ne pouvaient plus utiliser le registre sensoriel pour apprimer les gabarits et, en conséquence, la précision n'était pas influencée par l'intervalle séparant les formes. Notez que l'utilisation du registre sensoriel permettait plus de précision lorsque l'intervalle interstimulus était inférieur à 300 msec. Ce qui suggère que le registre sensoriel n'est opérant que sur une durée approximative d'un quart de seconde. Lorsque l'intervalle n'était que de 20 msec (sur les deux courbes du graphique, les points situés à l'extrême gauche) et que les formes étaient présentées au même endroit, la performance était presque parfaite, y compris pour les formes les plus complexes.

Lorsque la seconde forme était déplacée, les sujets devaient compter sur une description de la première forme et non sur une image sensorielle. Cette description pourrait être une image visuelle mais, à la différence d'une image sensorielle dans laquelle la forme perdure physiquement, une image visuelle doit être recherchée dans la mémoire. La description était donc moins précise que l'image sensorielle. Puisque la comparaison des descriptions de deux formes nécessite plus de temps que l'appariement de deux gabarits, la vitesse de réponse était hautement corrélée avec la précision. Dans la figure 2.2, les résultats concernant le temps de réaction seraient beaucoup plus semblables aux résultats concernant la précision si nous remplaçions « Pourcentage de réponses correctes » par « Vitesse de réponse ». Les temps de réaction étaient très courts dans la condition « identique » mais s'allongeaient à mesure que l'intervalle interstimulus augmentait. Les temps de réaction étaient plus longs dans la condition « déplacé » et ils n'étaient pas influencés par l'intervalle interstimulus. Les résultats concernant la précision et le temps de réaction suggèrent que le registre sensoriel peut être utilisé pour un appariement rapide des gabarits si les deux formes sont séparées par moins de 300 msec et qu'elles sont présentées au même endroit.

1.2 *Les théories des traits*

Pour que la théorie des gabarits explique les résultats obtenus par Phillips, il faut que la seconde forme soit appariée avec une image sensorielle de la première forme. L'image sensorielle peut être conçue comme une sorte d'image rémanente, dans laquelle la forme continue à être physiquement présente jusqu'à ce que le registre sensoriel se dégrade. Mais on peut se demander si le registre sensoriel joue un rôle important en dehors des laboratoires ; et serait-ce le cas, l'information est rapidement perdue. Comme nous ne pouvons généralement pas nous fier au registre sensoriel pour apprimer des formes, nous devons comparer des descriptions de formes. Les *théories des traits*[♦] nous permettent de décrire une forme au moyen d'une liste de ses parties. Par exemple, nous pourrions décrire un ami comme quelqu'un qui a de longs cheveux blonds, un nez court et des sourcils broussailleux.

Les théories des traits sont commodes pour décrire l'apprentissage fondé sur la perception et l'une des meilleures discussions au sujet de ces théories se trouve dans *Principles of Perceptual Learning and Development* de Gibson (1969). La théorie de Gibson soutient qu'un apprentissage qui est basé sur la perception se fait par la découverte de traits qui *distinguent* les formes entre elles.

Une partie des preuves en faveur des théories des traits provient de l'enregistrement des potentiels d'action au niveau de cellules individuelles dans le cortex visuel. En

♦ théorie des traits

Une théorie de la reconnaissance de formes qui décrit les formes selon leurs traits distinctifs ou caractéristiques.

plaçant des micro électrodes dans le cortex visuel d'animaux, Hubel et Wiesel (1962, 1963) ont découvert que les cellules ne répondent qu'à des types particuliers de stimuli, tels qu'une ligne d'une certaine largeur, orientée selon un certain angle et située à un certain endroit dans le champ visuel. D'autres cellules sont même concernées par la longueur de la ligne. En 1981, Hubel et Wiesel ont reçu le prix Nobel pour ce travail.

Bien que la plupart des théoriciens de la reconnaissance de formes recourent au concept de trait, trouver une bonne liste de traits représente souvent un vrai défi. Gibson (1969) a proposé les critères suivants, comme base pour sélectionner une série de traits dans le cas des lettres majuscules :

1. Les traits doivent être réellement discriminatoires, présents dans certains éléments de la série et pas dans d'autres, de façon à offrir un contraste.
2. L'identité des traits doit rester inchangée malgré les modifications de luminosité, de taille et de perspective.
3. Les traits doivent produire une forme unique pour chaque lettre.
4. Le nombre de traits proposés doit être raisonnablement petit.

Gibson a utilisé ces critères, des données empiriques et son intuition pour établir la série de traits des lettres majuscules. Les traits consistent essentiellement en différentes lignes et courbes qui sont les composantes des lettres. Les lignes, qui peuvent être horizontales, verticales ou diagonales, penchent de l'un et/ou de l'autre côté, telles que nous les trouvons dans la majuscule A. Les courbes comprennent un cercle fermé (la lettre O), un cercle ouvert vers le haut (la lettre U) ou un cercle ouvert sur le côté (la lettre C). La plupart des lettres comprennent plus d'un trait, comme la lettre Q qui présente un cercle fermé et une ligne diagonale.

Une série de traits est habituellement évaluée en fonction de son pouvoir de prédition des *confusions perceptives*[♦] puisque les items susceptibles d'être confondus devraient avoir de nombreux traits en commun. Par exemple, l'unique différence de trait entre les lettres P et R est la présence d'une ligne diagonale dans la lettre R; pour cette raison, les deux lettres devraient être facilement confondues. Les lettres R et O diffèrent par cinq traits et ne devraient donc être que rarement confondues.

Une méthode permettant de générer des confusions perceptives consiste à demander à un observateur d'identifier des lettres qui lui sont présentées très rapidement par un tachitoscope (Townsend, 1971). Dans de telles conditions, il est souvent difficile de discriminer des lettres physiquement proches et les erreurs fournissent une mesure de la perception de similarités. Holbrook (1975) a comparé deux modèles de traits pour déterminer avec quel succès chacun des deux pouvait prédire le type d'erreurs trouvées par Townsend. L'un était le modèle de Gibson et l'autre était une modification de ce modèle, proposée par Geyer et De Wald (1973). La modification la plus importante était la spécification du nombre de traits pour une lettre (comme deux lignes verticales pour la lettre H) plutôt qu'une simple liste de présence des traits.

Une comparaison des deux modèles a révélé que la série de traits suggérée par Geyer et De Wald était un meilleur prédicteur des erreurs de confusion commises par les adultes (Townsend, 1971) et par les enfants de 4 ans (Gibson *et al.*, 1963). Les prédictions des deux modèles ont été améliorées lorsque les traits furent pon-

♦ **confusion perceptive**

Une mesure de la fréquence à laquelle deux formes sont erronément perçues comme identiques.

dérés de façon optimale afin de tenir compte du fait que certains traits sont plus importants que d'autres pour expliquer les confusions. Puisque la distinction droite/courbe est particulièrement importante, elle devrait être plus soulignée que les autres.

1.2.1 Les traits distinctifs

Les enfants apprennent à identifier un objet parce qu'ils sont capables de repérer les différences entre cet objet et les autres. Par exemple, lorsque l'enfant est confronté pour la première fois aux lettres *E* et *F*, il peut ne pas savoir en quoi elles diffèrent. Apprendre à les discriminer nécessite qu'il découvre qu'une ligne horizontale est présente dans le bas de la lettre *E* mais pas dans la lettre *F*. La ligne horizontale du bas est un *trait distinctif** entre *E* et *F*, c'est-à-dire qu'elle nous permet de distinguer une forme d'une autre.

L'apprentissage fondé sur la perception peut être facilité par une procédure d'apprentissage qui met l'accent sur les traits distinctifs. Une méthode efficace pour insister sur un trait distinctif consiste à commencer par le colorer différemment du reste de la forme et à le ramener ensuite graduellement à sa couleur originale. Egeland (1975) a utilisé cette procédure pour enseigner à des enfants de prégarde comment distinguer entre des paires de lettres facilement confondues *R-P*, *Y-V*, *G-C*, *Q-O*, *M-N* et *K-X*. Une lettre de chaque paire était présentée au sommet d'une carte et au-dessus de six autres lettres dont trois s'apparaissaient avec elle, tandis que les trois autres constituaient la seconde lettre de la paire. On demandait aux enfants de choisir les lettres qui correspondaient exactement à la lettre du sommet.

Un groupe d'enfants a bénéficié d'une procédure d'entraînement dans laquelle le trait distinctif de la lettre était mis en évidence par sa couleur rouge – par exemple, la ligne diagonale du *R* dans la paire à discriminer *R-P*. Pendant la session d'entraînement, le trait distinctif était graduellement ramené à la couleur noire du reste de la lettre. Un autre groupe d'enfants ne voyait que des lettres noires. On leur faisait savoir si leurs choix étaient corrects, mais on ne leur disait rien à propos des traits distinctifs. Les deux groupes ont passé deux tests – un test immédiatement après la session d'entraînement et un test une semaine plus tard. Le groupe « traits distinctifs » a commis significativement moins d'erreurs aux deux tests, même lorsque les traits n'étaient pas mis en évidence lors du test. Les enfants de ce groupe se sont également moins trompés durant les sessions d'entraînement.

Accentuer les traits distinctifs a été doublement avantageux. D'abord, cela a permis aux enfants d'apprendre ces traits, de sorte qu'ils ont pu continuer à différencier les lettres lorsque ces traits n'étaient plus soulignés. Ensuite, cela leur a permis de connaître les traits sans commettre aucune erreur durant la session d'entraînement. Les échecs et les frustrations, dont beaucoup d'enfants font l'expérience au début de l'apprentissage de la lecture (la discrimination des lettres), peuvent diminuer leur intérêt pour les apprentissages scolaires ultérieurs.

1.2.2 Les visages

* trait distinctif

Un trait présent dans une forme mais absent dans une autre et qui concourt à la distinction entre les deux.

Les contributions d'Eleanor Gibson et de ses collègues sont centrées sur la discrimination de lettres et de formes similaires. Les théories des traits sont cependant très générales et peuvent être utilisées pour de nombreuses formes, dont les

visages. Pour décrire le visage d'une femme, vous vous référeriez probablement à des traits tels que son nez, ses yeux et sa bouche. Les deux premières rangées de la figure 2.3 présentent des variations pour deux types de traits faciaux.

Les photos de bouches et d'oreilles illustrent le genre de traits qui sont identifiés par les programmes d'intelligence artificielle développés par Ullman (2007). Ullman a découvert que des traits plus complexes, tels que les bouches et les oreilles (et les phares arrière de voiture), sont particulièrement utiles pour identifier des objets. Cependant, ses programmes sont également capables de reconnaître des traits de bas niveau, comme les lignes et les courbes. Ces lignes et ces courbes concernent toutes les formes, alors que les traits plus complexes, tels que les bouches et les feux arrière, sont spécifiques de certaines catégories, comme les visages et les voitures. Les traits de bas niveau permettent de distinguer entre les différents types d'un même trait complexe, comme lorsque l'on décrit le sourire d'Angelina Jolie en tant que combinaison particulière de courbes.

Cette distinction entre différents niveaux de description soulève un problème pratique à propos de la reconnaissance de formes : à quel niveau a lieu cette reconnaissance ? Un niveau peut être très général (un visage), intermédiaire (un visage asiatique) ou très spécifique (le visage d'Angie Chen). Reconnaître des visages à l'intérieur d'une même catégorie peut s'avérer difficile du fait que nous accentuons plutôt les traits nous permettant de distinguer *entre* les catégories que les traits nous permettant de distinguer *à l'intérieur* des catégories (Goldstone, Lippa & Shiffrin, 2001 ; Goldstone & Steyvers, 2001). Comme l'indique l'article « Dans la Presse » 2.1, il est particulièrement malaisé de distinguer parmi les visages d'une catégorie ethnique différente de la nôtre.

Les recherches corroborent cette explication. Les étudiants reconnaissent les visages d'autres races avec plus de précision si on les informe préalablement que



Figure 2.3

Exemples de traits plus complexes identifiés par les programmes de reconnaissance de formes d'Ullman.

Source : « Object recognition and segmentation by a fragment-based hierarchy », S. Ullman, 2007, *Trends in the Cognitive Sciences*, 11, 58-64.

Dans la presse 2.1

Pourquoi ont-ils tous l'air pareil ?

Siri CARPENTER

Il est notoire que nous reconnaissions très mal les visages des personnes appartenant à d'autres races. C'est une petite habitude chez les êtres humains qui est souvent expliquée par l'opinion selon laquelle nous regardons plus les membres de notre propre race et acquérons, de ce fait, une «expertise en perception» pour les traits de cette race.

Une version influente de cette hypothèse soutient que le déficit de reconnaissance inter-raciale peut être modélisé en supposant que, dans les autres races, les visages sont psychologiquement plus semblables les uns aux autres que dans la sienne propre. Mais cet argument laisse insatisfait Daniel Levin, psychologue de la cognition à la Kent State University.

«La position de l'expertise en perception est assez intuitive et sensée», dit-il. «Mais je soutiens qu'elle n'est pas vraiment exacte. Le problème n'est pas que nous ne pouvons pas coder les détails des autres races – le problème est que nous ne le faisons pas.»

Au contraire, selon lui, nous insistons démesurément sur les catégories raciales – quelqu'un est blanc, noir ou asiatique – et nous ignorons les

informations qui nous aideraient à reconnaître les gens en tant qu'individus. Dans une recherche récente, Levin a montré que les sujets peuvent, en fait, percevoir des différences subtiles sur les visages d'autres races – pour autant qu'ils utilisent ces différences pour classer des races.

Par exemple, Levin explique que «lorsqu'une personne de race blanche regarde le nez d'une autre personne de race blanche, elle se dit probablement "C'est le nez de John". Lorsqu'elle regarde le nez d'une personne de race noire, il est probable qu'elle pense "C'est un nez noir."»

Levin maintient que les résultats sont importants parce qu'ils aident à expliquer cette question de longue date: pourquoi sommes-nous médiocres dans la reconnaissance des visages de personnes d'autres groupes raciaux? Pareille compréhension pourrait être utile dans nombre de situations, dont l'entraînement des policiers et d'autres professionnels de la justice à identifier plus précisément les visages.

Source: traduit de «Pourquoi ont-ils tous l'air pareil?», Siri Carpenter, *Monitor on Psychology*, Décembre 2000, p. 44. Copyright © de la American Psychological Association. Reproduit avec autorisation.

cette tâche est généralement difficile et que, en conséquence, ils doivent faire tout particulièrement attention aux différences entre visages. Des étudiants blancs, qui avaient reçu ces instructions, n'ont par la suite montré aucune différence dans leur capacité à identifier des visages de Blancs et de Noirs (Hugenberg, Miller & Claypool, 2007). Des étudiants qui n'avaient pas reçu ces instructions ont obtenu l'habituel taux de reconnaissance faible dans le cas de visages d'une autre race. Des instructions peuvent donc rapidement éliminer les effets de la différence de race sur la reconnaissance faciale.

Se concentrer sur les traits distinctifs aide à distinguer les lettres et cela pourrait également aider à distinguer les visages. Pour tester cette hypothèse, Brennan (1985) a utilisé des *caricatures* générées par ordinateur, qui accentuent le caractère distinctif des traits. Par exemple, si quelqu'un avait de grandes oreilles et un petit nez, la caricature avait des oreilles encore plus grandes et un nez encore plus petit que le portrait fidèle. Lorsque l'on montrait aux étudiants des dessins au trait de connaissances, ils identifiaient plus vite les personnes qui étaient caricaturées que celles qui étaient représentées par un portrait fidèle (Rhodes, Brennan & Carey, 1987). Autrement dit, rendre des traits distinctifs encore plus distinctifs par le moyen de l'exagération facilitait la reconnaissance.

1.3 Les théories structurelles

◆ **caricature**

Une exagération des traits caractéristiques afin de rendre la forme plus distinctive.

Les théories des traits se heurtent au fait que les descriptions de formes nécessitent souvent que nous précisions de quelle façon les traits sont regroupés.

Prendre en compte la manière dont les traits sont assemblés est un des principes directeurs de la psychologie Gestalt. Les psychologues de la Gestalt considèrent qu'une forme est plus que la somme de ses parties. Les relations entre les traits d'une forme ont été formalisées par des chercheurs travaillant dans le domaine de l'intelligence artificielle qui ont découvert que l'interprétation des formes dépend en général de l'idée que nous nous faisons de la façon dont leurs lignes se rejoignent (Clowes, 1969).

Les *théories structurelles*[♦] reposent sur les théories des traits. Avant de pouvoir préciser les relations entre les traits, nous devons définir ces traits. Une théorie structurelle permet d'indiquer la façon dont les traits sont assemblés. Par exemple, la lettre *H* est constituée de deux lignes verticales et d'une ligne horizontale. Mais nous pourrions réaliser un grand nombre de formes différentes à partir de deux lignes verticales et une ligne horizontale. Ce qui est indispensable, c'est une spécification précise de la manière dont les lignes doivent être jointes – la lettre *H* est formée de deux lignes verticales connectées en leur milieu par une ligne horizontale.

La figure 2.4 montre des formes-squelettes (*skeleton shapes*) représentant des animaux différents, basées sur des descriptions structurelles d'abord proposées par Blum (1973) afin de choisir parmi les formes biologiques. Wilder, Feldman et Singh (2011) adoptèrent ses méthodes afin de faire des prédictions sur la catégorisation de ces formes en *feuille* ou *animal*. La réussite de leurs prédictions soutient la supposition que les gens utilisent cette sorte de description dans la classification. Les formes-squelettes animales ont des membres dessinés avec des lignes plus ondulées que les feuilles, qui sont représentées par un plus petit nombre de lignes plus droites.

Passer d'un monde bidimensionnel de symboles, tels que des lettres et des nombres, à un monde tridimensionnel d'objets, génère de nouveaux défis pour identifier et décrire la relation entre les traits. La figure 2.5 illustre un problème d'identification de traits dû à la difficulté relative de percevoir les trois formes comme des cubes (Kopfermann, 1930). La forme de gauche est la plus difficile à percevoir comme un cube et celle du milieu la plus facile. Avant de lire ce qui suit, essayez d'en deviner la raison (indice: pensez à la difficulté à identifier les traits pour chacun des trois exemples).

Le livre d'Hoffman (1998) à propos de l'intelligence visuelle avance que nous obéissons à des règles lorsque nous créons des descriptions de formes. La première des nombreuses règles décrites dans ce livre est de toujours interpréter une ligne droite dans une image comme une ligne droite dans un espace tridimensionnel.

- ♦ **théorie structurelle**
Une théorie qui spécifie de quelle manière les traits d'une forme sont reliés les uns aux autres.

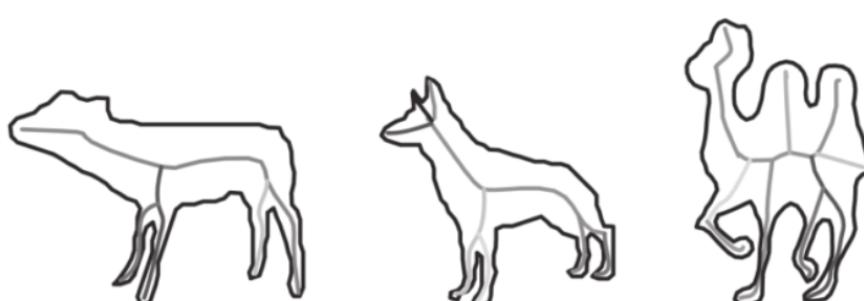


Figure 2.4

Exemples de formes de squelettes d'animaux.

Source: «Superordinate shape classification using natural shape statistics», J. Wilder, J. Feldman & M. Singh, 2011, *Cognition*, 119, 167-174. Copyright 2011 by Elsevier.
Reproduit avec autorisation.

C'est pourquoi, dans la figure 2.5, nous percevons la longue ligne verticale au centre de la forme de droite comme une seule ligne. Cependant, il est nécessaire de diviser cette ligne en deux lignes distinctes pour obtenir un cube parce que les lignes appartiennent à des surfaces différentes. Il est particulièrement difficile de voir la figure de gauche comme un cube parce que nous devons en outre diviser les deux longues lignes diagonales en deux lignes plus courtes, de façon à ne pas percevoir une forme plate.

La forme du milieu est facilement identifiable en tant que cube et vous avez peut-être reconnu le fameux cube de Necker. Le cube de Necker est bien connu parce que la perception des surfaces frontale et arrière du cube change en fonction de la façon dont on les regarde (Long & Toppino, 2004). Encore un exemple d'une description structurelle qui peut changer, alors que les traits ne changent pas !

1.3.1 Le modèle des composants de Biederman

Les descriptions des objets tridimensionnels seraient assez compliquées si nous devions en relever chacune des lignes et des courbes. Par exemple, les cubes de la figure 2.5 consistent chacun en 12 lignes (que vous trouverez peut-être plus faciles à compter dans les cubes de droite et de gauche que dans le cube de Necker, après avoir divisé les lignes).

Il serait plus aisés de décrire les objets tridimensionnels en termes de volumes simples tels que cubes, cylindres, côtés et cônes, que de décrire tous les traits de ces volumes. Comme les traits des lettres, ces composants peuvent être combinés de nombreuses façons différentes pour produire une diversité d'objets. Par exemple, la tasse et le seau dans la figure 2.6 contiennent les deux mêmes composants dans un arrangement différent. Les études sur la recherche de formes ont montré qu'à la fois les composants et leurs relations déterminent la similarité perçue des formes (Arguin & Saumier, 2004). Par exemple, l'attaché-case et le tiroir de la figure 2.6 sont similaires parce qu'ils partagent les mêmes composants. Cependant, l'attaché-case ressemble plus au seau qu'à la tasse à cause de la *relation* entre les composants – la poignée se trouve au sommet pour l'attaché-case et pour le seau.

Pouvoir former un grand nombre d'arrangements à partir d'un petit nombre de composants est avantageux car nous n'avons alors besoin que d'un nombre relati-

Figure 2.5

Percevoir des cubes.

Source: *Visual intelligence*, D.D. Hoffman.

Copyright 1998 de W.W. Norton.

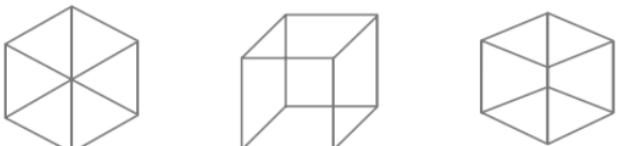


Figure 2.6

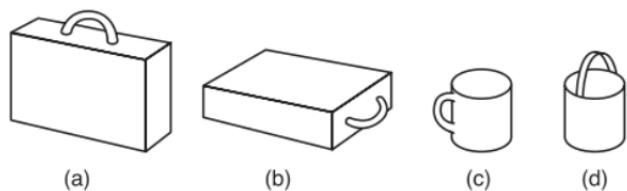
Les mêmes composants arrangés différemment peuvent produire des objets différents.

Source: «Human image understanding: Recent research and a theory»,

I. Biederman, 1985, *Computer Vision, Graphics, and Image Processing*, 32,

29-73. Copyright 1985 de Academic Press. Reproduction autorisée

par Elsevier Science.



vement réduit de composants pour décrire les objets. Biederman (1985) a proposé que seuls 35 formes élémentaires (qu'il a baptisés géons[♦]), approximativement, nous sont nécessaires pour décrire les objets du monde. L'avantage de pouvoir créer différents arrangements à partir de quelques composantes est de n'avoir besoin que d'un nombre limité d'entre elles pour décrire des objets. Biederman (1985) propose l'idée que nous n'avons besoin que d'environ 35 volumes simples (qu'il appelle géons) afin de décrire tous les objets dans le monde. Les recherches de Biederman, Yue et Davidoff (2009) établissent qu'il est plus facile de faire la différence entre un géon et un deuxième géon qu'entre deux variations d'un même géon. Par exemple, des étudiants américains différencient plus facilement les objets du milieu de la figure 2.7 de l'objet à leur gauche (un géon différent avec des côtés droits) que de l'objet à leur droite (une variation du même géon avec une plus grande courbe).

Ces résultats soulèvent la question des différences culturelles relatives à la capacité de distinguer un géon d'un autre. Distinguer une ligne courbe d'une ligne droite est un aspect fondamental de la culture occidentale, ainsi que nous l'avons déjà vu pour la différence entre les lettres de l'alphabet. Par contre, il y a un moindre besoin de pouvoir distinguer entre les lignes et les courbes de la part des Himba, un peuple semi-nomade vivant dans les régions reculées de la Namibie. Néanmoins, les Himba voient plus facilement la différence entre les géons (les deux objets à gauche) qu'entre des variations du même géon (les deux objets à droite).

♦ géons

Différentes figures tridimensionnelles qui se combinent pour réaliser des formes tridimensionnelles.

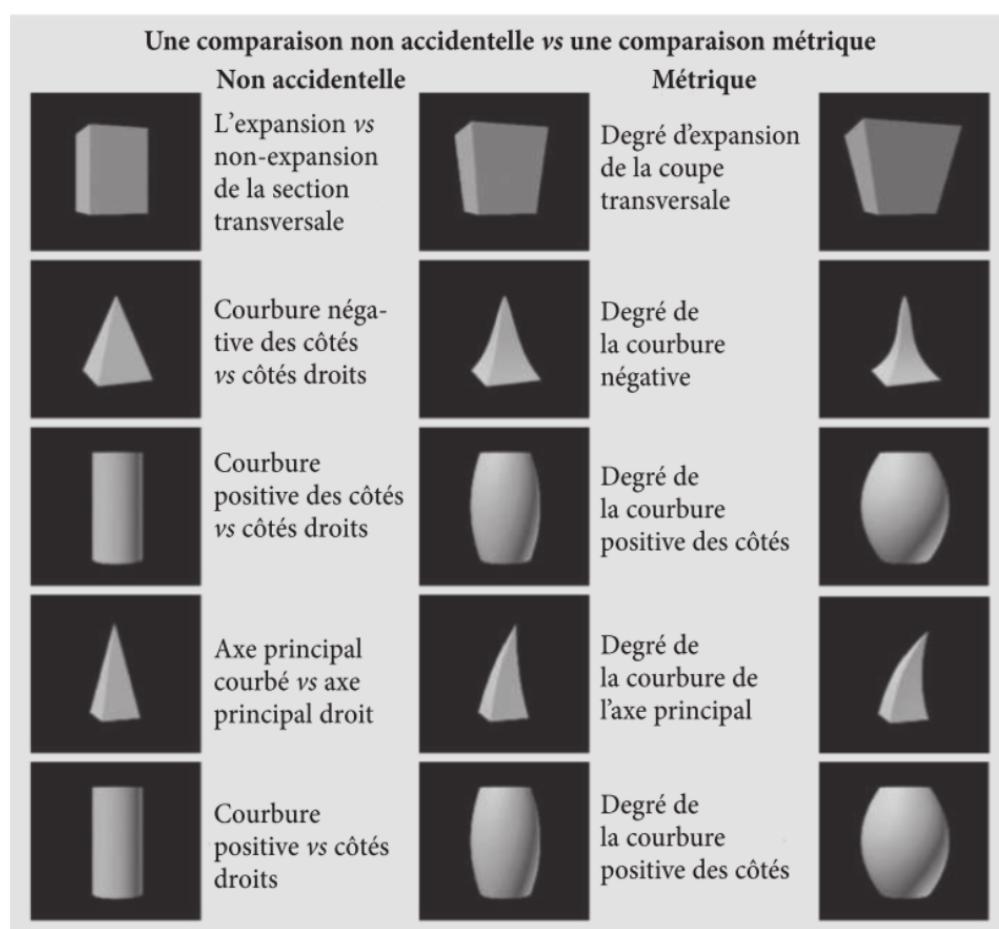


Figure 2.7

Différencier entre des géons différents (au milieu et à gauche) est plus facile que de différencier entre des variations d'un même géon (au milieu et à droite).
Source: « Representation of shape in individuals from a culture with minimal exposure to regular, simple artifacts », I. Biederman, X. Yue, & J. Davidoff, 2009, *Psychological Science*, 20, 1437-1442. Copyright: 2009. Reproduit avec autorisation.

Si la reconnaissance des formes consiste principalement en la description des relations entre un ensemble limité de composantes, il s'ensuit que la suppression de ces informations devrait réduire la capacité à reconnaître des formes. Afin de tester cette hypothèse, Biederman a retiré 65 % du contour d'objets comme les deux tasses de la figure 2.8. Pour la tasse de gauche, les contours au milieu des segments ont été supprimés, permettant aux observateurs de voir la relation entre les segments. Pour la tasse de droite, les contours aux angles ont été supprimés afin de rendre difficile la perception de la relation entre les segments. Lorsque le dessin des objets était présenté pendant 100 msec, les participants ont reconnu 70 % des objets si les parties supprimées étaient au milieu des segments. Si les parties supprimées étaient aux angles, les participants reconnaissaient moins de 50 % des objets (Biederman, 1985). Ainsi qu'il l'avait prévu, la destruction des informations relationnelles réduit la capacité à reconnaître un objet.

En conclusion, les théories structurelles poussent plus loin les théories des traits, en spécifiant comment les traits sont reliés. Sutherland (1968) fut l'un des premiers à soutenir que si nous voulons rendre compte de notre capacité véritablement impressionnante à reconnaître des objets, nous aurons besoin du langage descriptif plus puissant de la théorie structurelle. Les expériences présentées dans cette section démontrent que Sutherland avait raison. Voyons maintenant comment la reconnaissance de formes se déroule dans le temps.

2. Les étapes du traitement de l'information

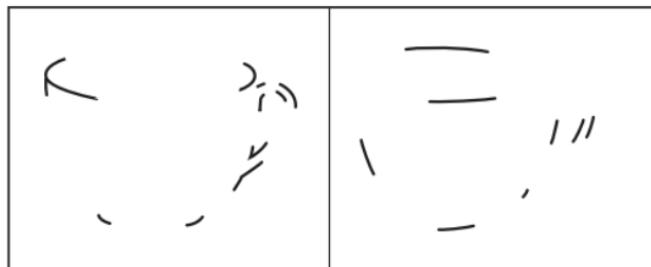
2.1 *La technique du rapport partiel*

Pour comprendre la façon dont nous exécutons une tâche de reconnaissance de formes, il nous faut identifier ce qui a lieu lors de chacune des étapes du traitement de l'information abordées dans le chapitre 1. On doit à Sperling (1960) la première construction d'un modèle de traitement de l'information pour une tâche de reconnaissance visuelle. Les sujets de l'expérience de Sperling voyaient des groupes de lettres présentées brièvement (habituellement, durant 50 msec) par un tachitoscope et on leur demandait de rapporter toutes les lettres qu'ils pouvaient se rappeler. Les réponses étaient très précises lorsque le nombre de lettres était inférieur à 5. Mais lorsque ce nombre augmentait, les sujets ne rapportaient correctement jamais plus qu'une moyenne de 4,5 lettres, quel que soit le nombre de lettres présentées.

Figure 2.8

Illustration du retrait de 65 % du contour au niveau de la moitié des segments (objet de gauche) ou des sommets (objet de droite).

Source : « Human image understanding: Recent research and a theory », I. Biederman, 1985, *Computer Vision, Graphics, and Image Processing*, 32, 29-73. Copyright 1985 de Academic Press. Reproduction autorisée par Elsevier Science.



Un problème général dans l'élaboration d'un modèle de traitement de l'information est l'identification de la cause de la limitation d'une performance. Sperling cherchait à mesurer le nombre de lettres qui peuvent être reconnues à l'occasion d'une brève exposition, mais il était conscient que la limite supérieure de 4,5 pouvait être due à une incapacité à se remémorer davantage. En d'autres mots, il se pouvait que les sujets aient reconnu la plupart des lettres mais qu'ils en aient oublié certaines avant de pouvoir rapporter ce qu'ils avaient vu. Pour cette raison, Sperling a changé sa *procédure de rapport intégral** (rapporter toutes les lettres) en une *procédure de rapport partiel** (rapporter seulement quelques lettres).

La présentation la plus typique comprenait trois rangées de quatre lettres chacune. Les sujets ne pouvaient être capables de se rappeler les 12 lettres, mais ils devaient être à même de s'en rappeler 4. La procédure du rappel partiel requiert que les sujets ne rapportent qu'une seule rangée. La hauteur d'un signal sonore indiquait laquelle des trois rangées rapporter : la tonalité aiguë pour la rangée supérieure, la tonalité intermédiaire pour la rangée du milieu et la tonalité grave pour la rangée inférieure. Le signal avait lieu juste après la disparition des lettres afin que les sujets regardent les trois rangées et non simplement l'une d'entre elles (figure 2.9). L'utilisation de la technique du rapport partiel est basée sur l'hypothèse selon laquelle le nombre de lettres rapportées dans une rangée indiquée est égal au nombre moyen de lettres perçues dans chacune des rangées puisque les sujets ne pouvaient pas anticiper quelle rangée regarder. Les résultats de cette procédure ont montré que les sujets peuvent rapporter correctement trois des quatre lettres d'une rangée, ce qui implique qu'ils en ont reconnu 9 dans la présentation.

◆ **procédure de rapport intégral**
Une tâche qui requiert que des observateurs rapportent tout ce qu'ils voient lors d'une présentation d'items.

◆ **procédure de rapport partiel**
Une tâche dans laquelle il est demandé aux observateurs de ne rapporter, parmi un ensemble d'items, que certains d'entre eux.

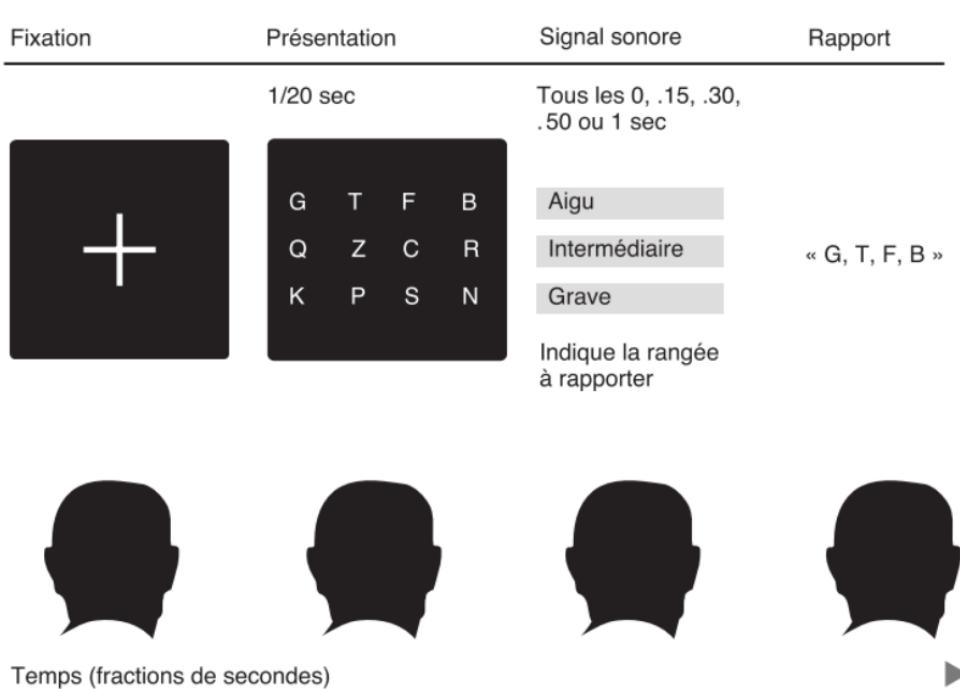


Figure 2.9

L'étude de la mémoire sensorielle par Sperling (1960). Après que les sujets aient fixé la croix, les lettres apparaissent sur l'écran le temps nécessaire à la création d'une image rémanente. Des signaux sonores aigus, intermédiaires et graves signalaient la rangée de lettres à rapporter.
Source : « The information available in brief visual presentations », G. Sperling, 1960, *Psychological Monographs*, 74 (Whole No. 498).

2.2 Le modèle de Sperling

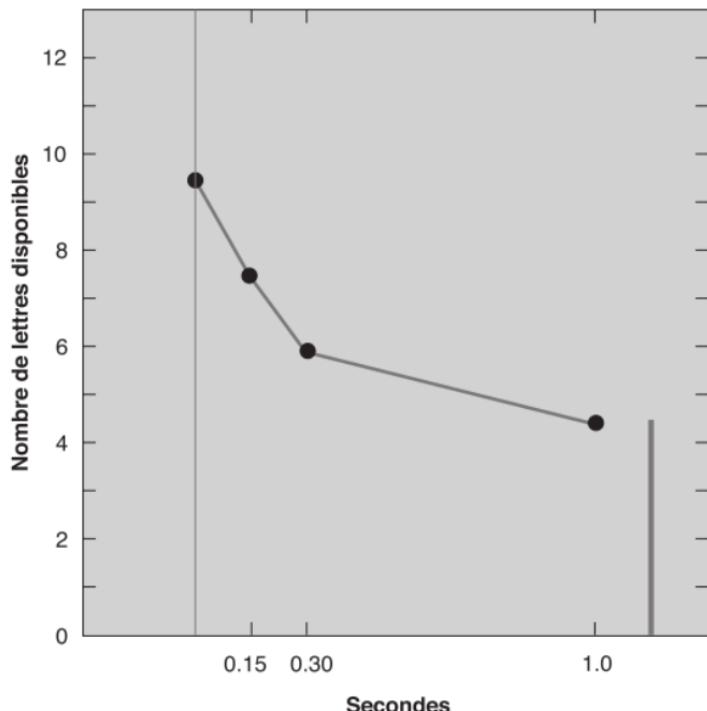
Il arrive bien souvent que ce dont on se rappelle le mieux du travail d'un scientifique ne soit pas ce qu'il avait l'intention d'étudier. Bien que Sperling ait conçu la technique du rapport partiel pour réduire les exigences à l'égard de la mémoire dans son expérience et pour obtenir une mesure « pure » de la perception, son travail est surtout connu pour la découverte de l'importance du registre sensoriel visuel. Comment cela s'est-il produit ? L'estimation que les sujets avaient perçu 9 lettres a été obtenue dans le cas où le signal sonore suivait immédiatement la fin de la présentation (50 msec). Les sujets pouvaient alors rapporter environ trois quarts des lettres et les trois quarts de douze font 9. Mais lorsque le signal sonore survenait jusqu'à 1 sec après la présentation, la performance chutait à 4,5 lettres seulement. C'est-à-dire qu'il y avait une perte progressive de 9 lettres à 4,5 lettres à mesure que le délai du signal sonore augmentait de 0 à 1 seconde (figure 2.10).

Ce qu'il y a de plus intéressant à propos du nombre 4,5, c'est qu'il est exactement égal à la limite supérieure de la performance pour le rapport intégral, représentée par la barre bleue dans la figure 2.10. La procédure de rapport partiel n'offre aucun avantage sur la procédure de rapport intégral si le signal sonore se produit après 1 seconde ou plus. Pour expliquer ce déclin progressif de la performance, Sperling a proposé que les sujets utilisent un registre sensoriel visuel pour reconnaître les lettres de la rangée indicée. Lorsqu'ils entendaient le signal sonore, ils faisaient sélectivement attention, dans le registre, à la rangée indicée et essayaient d'identifier les lettres de cette rangée. Le succès dû au signal sonore dépendait de la clarté des informations dans leur registre sensoriel. Lorsque le signal sonore retentissait immédiatement après la fin du stimulus, les informations étaient suffisamment

Figure 2.10

Rappel en fonction du délai d'apparition d'un signal sonore.

Source : « The information available in brief visual presentations », de G. Sperling, 1960, *Psychological Monographs*, 74 (Whole No. 498).



claires pour que des lettres supplémentaires soient reconnues dans la rangée indiquée. Mais à mesure que la clarté de l'image sensorielle diminuait, il devenait de plus en plus difficile de reconnaître des lettres supplémentaires. Lorsque le signal survanait après 1 seconde, les sujets ne pouvaient pas du tout utiliser le registre sensoriel pour se concentrer sur la rangée indiquée, de sorte que la performance était déterminée par le nombre de lettres qu'ils avaient reconnues dans l'ensemble de la présentation pour cette rangée. Pour cette raison, leur performance était équivalente à celle de la procédure de rapport intégral, dans laquelle ils faisaient attention à l'ensemble de la présentation.

En 1963, Sperling a proposé un modèle de traitement de l'information pour la performance lors de sa tâche de rapport visuel. Le modèle consistait en un registre d'information visuel, un balayage visuel, la répétition et un registre d'information auditif.

Le *registre d'information visuel** est un registre sensoriel qui conserve l'information pendant un bref moment, variant d'une fraction de seconde à une seconde. La vitesse de dégradation dépend de facteurs comme l'intensité, le contraste et la durée du stimulus, et également du fait que l'exposition au stimulus est suivie ou non par une seconde exposition. Le masquage visuel a lieu lorsqu'une seconde exposition, qui consiste en un champ vivement éclairé ou en une série de formes différentes, réduit l'efficacité du registre d'information visuel.

Pour que la reconnaissance de formes ait lieu, l'information contenue dans le registre sensoriel doit être parcourue. Sperling a d'abord considéré que ce balayage avait lieu item par item, comme si nous regardions à travers un trou pratiqué dans un morceau de carton, tout juste assez large pour nous laisser voir une lettre.

Les deux autres composants du modèle étaient la *répétition** (se dire la lettre en soi-même) et un *registre d'information auditif** (se rappeler le nom des lettres). Les sujets expliquaient en général que, pour se souvenir des items jusqu'au moment du rappel, ils se les répétaient. Une preuve supplémentaire de la répétition verbale a été trouvée dans des erreurs de rappel dues à des confusions auditives – autrement dit, une production de lettre dont la prononciation est proche de celle de la lettre exacte. L'avantage du registre auditif est que cette vocalisation interne des lettres les maintient actives dans la mémoire. Le registre auditif de Sperling fait partie de la mémoire à court terme (MCT), un sujet que nous aborderons plus loin dans ce livre.

Sperling a révisé son modèle initial en 1967. À cette époque, des preuves ont commencé à s'accumuler, suggérant que les formes n'étaient pas parcourues une à une mais étaient analysées simultanément. Cette distinction entre la réalisation d'une seule opération cognitive à la fois (*traitement sériel**) et la réalisation de plusieurs opérations cognitives à la fois (*traitement parallèle**) est fondamentale en *psychologie cognitive**. Pour cette raison, Sperling a modifié son idée du *composant de balayage visuel** de façon à ce que la reconnaissance de formes se produise simultanément dans toute la présentation, bien que la vitesse de reconnaissance à un endroit précis dépende de l'endroit où le sujet focalise son attention.

Comme je l'ai indiqué dans le chapitre 1, il s'agissait de l'un des premiers modèles qui tentaient de montrer comment plusieurs étapes (registre sensoriel, reconnaissance de formes et MCT) se combinent pour influencer la performance dans une

◆ **registre d'information visuel (RIV)**

Un registre sensoriel qui maintient l'information visuelle durant environ un quart de seconde.

◆ **répétition**

Répéter l'information verbale pour la maintenir active dans la MCT ou pour la transférer dans la MLT.

◆ **registre d'information auditif**

Dans le modèle de Sperling, ce registre maintient l'information verbale dans la MCT au moyen de la répétition.

◆ **traitement sériel**

Exécuter une opération à la fois, comme prononcer un mot à la fois.

◆ **traitement parallèle**

Exécuter plus d'une opération à la fois, comme regarder une œuvre d'art et tenir une conversation.

◆ **composant de balayage visuel**

Dans le modèle de Sperling, le composant de l'attention qui détermine ce qui est reconnu dans le registre visuel.

tâche de traitement visuel. Cela a contribué à la construction du modèle général illustré par la figure 1.1 et a mené au développement ultérieur de modèles plus détaillés de la façon dont nous reconnaissons les lettres dans une présentation visuelle (Rumelhart, 1970).

2.3 *Le modèle de Rumelhart*

En 1970, Rumelhart a proposé un modèle mathématique détaillé de la performance pour un vaste champ de tâches de traitement de l'information, dont les procédures de rapports intégral et partiel étudiées par Sperling. Ce modèle repose sur les hypothèses-clés de Sperling, telles que l'importance du registre d'information visuel et l'utilisation d'un balayage en parallèle pour reconnaître les formes. Mais Rumelhart a été plus précis dans sa description de la façon dont la reconnaissance a lieu. Il suppose que la reconnaissance se produit par identification des traits d'une forme.

La reconnaissance de traits a lieu simultanément dans l'ensemble de la présentation, mais elle prend du temps; plus l'observateur dispose de temps, plus grand est le nombre de traits qu'il peut reconnaître. Imaginez que vous regardez l'écran d'un tachitoscope et que l'expérimentateur présente brièvement les lettres *F*, *R* et *Z*. Si l'exposition est très courte, il se pourrait que vous ne voyiez que la ligne verticale du *F*, la courbe de la lettre *R* et la diagonale de la lettre *Z*. Si vous êtes tenu de deviner à ce moment, il est plus que probable que vous utiliserez ces informations. Il se pourrait que vous deviniez que le *R* était un *R*, un *P* ou un *B* parce que ces lettres possèdent un segment courbe à leur sommet. Si l'exposition est un peu plus longue, vous serez capable de voir plus de traits et il vous sera plus aisés de deviner ou même de reconnaître la lettre entière. Votre réussite dans l'identification des lettres sera alors déterminée, non seulement par la durée de l'exposition, mais aussi par la rapidité avec laquelle vous reconnaissiez les traits.

La vitesse de reconnaissance des traits dans le modèle de Rumelhart est influencée par la clarté de l'information et par le nombre d'items présentés. À la fin de la présentation, la clarté diminue à mesure que le registre d'information visuel se dégrade. Le nombre d'items présentés affecte la vitesse de reconnaissance des traits parce que le modèle postule que nous possédons une quantité limitée d'attention répartie sur tous les items. Lorsque le nombre d'items augmente, la quantité d'attention qui peut être focalisée sur chaque item diminue, ce qui ralentit la vitesse de reconnaissance de chacun d'eux.

L'hypothèse selon laquelle la vitesse de reconnaissance des traits dépend du nombre d'items présentés et de la clarté de l'information est utilisée par Rumelhart pour expliquer la performance dans la tâche de Sperling. Le modèle de Rumelhart proposait que les sujets peuvent rapporter une moyenne de 4,5 lettres dans le cas de la procédure de rapport intégral, en raison d'une limitation perceptive plutôt que du fait d'une limitation de la mémoire. Pendant que le nombre de lettres se rapproche de 12, les sujets continuent à essayer de reconnaître toutes les lettres simultanément. Mais, pour chaque lettre, la vitesse de reconnaissance diminue à mesure que d'autres lettres sont ajoutées dans la présentation. Bien que plus de lettres pourraient être reconnues, leur accroissement est compensé par une plus faible probabilité de reconnaître chacune d'elles.

Le modèle de Rumelhart postulait que, dans le cas de la procédure de rapport partiel, l'observateur tente de reconnaître les lettres dans toute la présentation avant d'entendre le signal sonore. Puis, percevant le signal, il se concentre uniquement sur la rangée indicée dans le registre d'information visuel et tente de reconnaître les lettres supplémentaires de cette rangée particulière. La vitesse de reconnaissance est plus élevée parce que l'observateur ne doit plus faire attention qu'à 4 lettres au lieu de 12. Mais, à mesure que le registre d'information se dégrade, non seulement le temps disponible pour l'utiliser diminue mais il devient aussi plus difficile à utiliser car il est de moins en moins précis. Réussir à fixer son regard sur la ligne indicée dépend donc de façon très critique du délai jusqu'au signal sonore, comme l'illustre la figure 2.10. Les hypothèses de Rumelhart fournissent des prédictions quantitatives précises sur la performance, non seulement dans les tâches de Sperling, mais aussi dans un certain nombre d'autres tâches.

Des études plus récentes ont confirmé un grand nombre de ces hypothèses, dont celle qui avance que les sujets passent d'un regard sur l'ensemble de la présentation à un regard sur la rangée indicée seulement lorsqu'ils savent laquelle ils doivent rapporter (Gegenfurtner & Sperling, 1993). Cependant, avant de recevoir l'indice, les observateurs font surtout attention à la rangée du milieu et sont en conséquence plus précis lorsqu'on leur demande de rapporter les lettres de cette rangée.

Vous avez sans doute deviné maintenant qu'il est difficile d'utiliser le paradigme du rapport partiel pour répondre à la question initiale de Sperling au sujet du nombre de lettres que nous percevons lors d'une brève exposition. Les observateurs commencent par essayer de percevoir les lettres de l'ensemble de la présentation, avec une attirance prononcée pour la rangée du milieu; ensuite, ils entendent le signal sonore et décident où porter leur attention; enfin, ils tentent de reconnaître les lettres uniquement dans la rangée indicée.

Une procédure plus à même de répondre à cette question, appelée le *paradigme de détection*[♦], a été conçue par Estes et Taylor (1966). Cette procédure nécessite que l'observateur rapporte laquelle de deux lettres cibles fait partie de la présentation d'un ensemble de lettres. Par exemple, on pourrait dire à un sujet que la présentation contiendra un *B* ou un *F* et la tâche consistera à rapporter quelle lettre est présentée. L'effort de mémoire est minime parce que le sujet ne doit rapporter qu'une seule lettre. En se basant sur le pourcentage d'essais corrects et en tenant compte des réponses où les sujets devinent, Estes et Taylor ont été capables de calculer le nombre moyen de lettres perçues à chaque essai. La procédure de détection a également été analysée par Rumelhart (1970), en tant que partie de son modèle général – un modèle qui fournissait un compte rendu impressionnant des performances lors des tâches visuelles de traitement de l'information étudiées au cours des années 1960.

3. La reconnaissance de mots

3.1 L'effet de supériorité du mot

Dans les années 1970, la recherche sur la reconnaissance de formes a, en grande partie, délaissé la question de savoir comment nous reconnaissions des lettres isolées pour s'intéresser à la façon dont nous reconnaissions les lettres dans les

♦ **paradigme de détection**
Une procédure dans laquelle les observateurs doivent spécifier laquelle de deux formes-cibles est présentée.

mots. Cette recherche a été stimulée par une découverte baptisée *effet de supériorité du mot*⁶. Dans sa thèse soutenue à la University of Michigan, Reicher (1969) a étudié la possibilité d'inclure le composant de balayage visuel dans le modèle que Sperling a élaboré en 1967. Si l'observateur essaie de reconnaître simultanément toutes les lettres présentées, peut-il reconnaître une unité de quatre lettres dans le même laps de temps que celui qu'il lui faut pour en reconnaître une seule ? Pour répondre à cette question, Reicher a mis au point une expérience dans laquelle les observateurs voyaient une seule lettre, un mot de quatre lettres ou un non-mot de quatre lettres. La tâche consistait toujours à identifier une seule lettre en choisissant parmi deux alternatives. La présentation du stimulus était immédiatement suivie par un cache visuel comprenant les deux réponses alternatives, situées immédiatement au-dessus de la lettre en question. Par exemple, une série de stimuli était le mot WORK, la lettre K et le non-mot OWRK. Ici, les deux alternatives étaient les lettres D et K, présentées au-dessus de la lettre critique K (figure 2.11). Les observateurs indiquaient la lettre qui, selon eux, occupait cette place, un D ou un K.

Cet exemple illustre plusieurs caractéristiques du dispositif de Reicher. Premièrement, le mot de quatre lettres est composé des mêmes lettres que le non-mot de quatre lettres. Deuxièmement, la place de la lettre critique est la même dans le mot et dans le non-mot. Troisièmement, les deux alternatives de réponse forment un mot (WORD ou WORK) dans la condition « mot » et un non-mot dans la condition « non-mot ». Quatrièmement, l'effort de mémoire est minime parce que les sujets n'identifient qu'une seule lettre, même lorsque quatre lettres sont présentées.

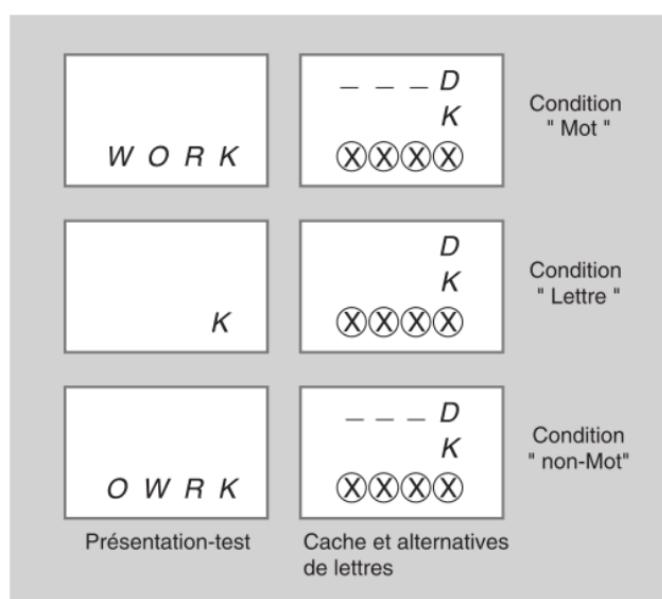
Les résultats ont montré que, lorsque la lettre critique faisait partie d'un mot, les sujets étaient significativement plus précis dans leur identification que lorsque la lettre critique faisait partie d'un non-mot ou que lorsqu'elle était présentée seule (l'*effet de supériorité du mot*⁶). Huit des neuf sujets ont obtenu des résultats meilleurs avec des mots isolés qu'avec des lettres isolées. L'unique sujet qui soit allé à l'encontre de cette tendance était le seul qui avait dit avoir vu les mots comme quatre lettres

◆ effet de supériorité du mot

La découverte que la précision avec laquelle une lettre est reconnue est meilleure lorsque la lettre fait partie d'un mot que lorsqu'elle apparaît seule ou dans un non-mot.

Figure 2.11

Exemple des trois conditions expérimentales de l'expérience de Reicher (1969). Le cache et les alternatives de réponse suivaient la position-test. La tâche consistait à décider laquelle des deux alternatives était apparue à cette position.



séparées qu'il assemblait ensuite en mot; les autres sujets disaient avoir vu les mots en tant que tels et non comme quatre lettres formant ensuite un mot.

L'effet de supériorité du mot est un autre exemple du traitement top-down. Nous avons vu précédemment comment le fait que nous sachions quels mots conviennent dans un contexte verbal particulier nous aide à reconnaître un mot. L'effet de supériorité du mot révèle la façon dont notre connaissance des mots nous aide à reconnaître plus rapidement les lettres d'un mot. En conséquence, le traitement top-down, partant de la connaissance stockée dans la MLT, peut concourir de diverses manières à la reconnaissance de formes.

3.2 Un modèle de l'effet de supériorité du mot

L'un des défis majeurs pour les psychologues étudiant la reconnaissance de mots a été d'expliquer les raisons de l'effet de supériorité du mot (Pollatsek & Rayner, 1989). Un modèle particulièrement influent, le *modèle d'activation interactive*♦ proposé par McClelland et Rumelhart (1981), contient plusieurs hypothèses de base reposant sur celles du modèle préalable de reconnaissance de lettres de Rumelhart. La première hypothèse est que la perception visuelle implique un traitement parallèle. Ce traitement se fait selon deux sens différents. Le traitement visuel est spatialement parallèle, les quatre lettres d'un mot étant traitées simultanément. Cette hypothèse est en accord avec le balayage visuel en parallèle de Sperling et avec le modèle de Rumelhart à propos de la façon dont les sujets tentent de reconnaître des lettres dans une présentation.

Le traitement visuel est également parallèle au sens où la reconnaissance a lieu en même temps dans trois niveaux d'abstraction différents. Les trois niveaux – le niveau « Trait », le niveau « Lettre » et le niveau « Mot » – sont présentés à la figure 2.12. Une hypothèse-clé du modèle d'activation interactive est que les trois niveaux interagissent pour déterminer ce que nous percevons. La connaissance des mots d'une langue interagit avec les informations entrantes relatives aux traits pour fournir des indices au sujet des lettres qui composent le mot. Ceci est illustré, à la figure 2.12, par les flèches qui montrent que le niveau « Lettre » reçoit des informations du niveau « Trait » et du niveau « Mot ».

Il existe deux types de connexions entre les niveaux : les connexions excitatrices et les connexions inhibitrices. Les *connexions excitatrices*♦ fournissent des indices positifs et les *connexions inhibitrices*♦ des indices négatifs sur l'identité d'une lettre ou d'un mot. Par exemple, une ligne diagonale apporte un indice positif pour la lettre K (et toutes les autres lettres qui

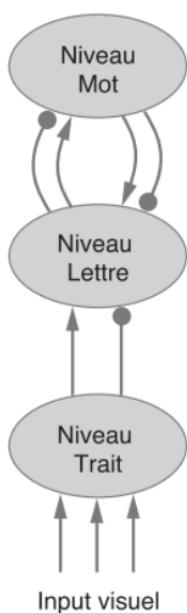


Figure 2.12

Les trois niveaux du modèle d'activation interactive. Les flèches indiquent les connexions excitatrices et les cercles indiquent les connexions inhibitrices.

Source: «An interactive-activation model of context effects in letter perception: Part 1. An account of basic findings», J.L. McClelland et D.E. Rumelhart, 1981, *Psychological Review*, 88, 375-407.
Copyright © 1981 de la American Psychological Association. Reproduction autorisée.

◆ **modèle d'activation interactive**
Une théorie qui propose que la connaissance des traits et celle des mots se combinent pour fournir des informations sur l'identité des lettres d'un mot.

◆ **connexion excitatrice**
Une réaction à un indice positif pour un concept, comme lorsqu'une ligne verticale renforce la possibilité qu'une lettre soit un K.

◆ **connexion inhibitrice**
Une réaction à un indice négatif pour un concept, comme lorsque la présence d'une ligne verticale fournit un témoignage négatif pour la lettre C.

contiennent cette ligne) et un indice négatif pour la lettre *D* (et toutes les autres lettres qui ne contiennent pas de diagonale). Les connexions excitatrices et inhibitrices sont également présentes entre le niveau «Lettre» et le niveau «Mot», selon que la lettre est correctement placée dans le mot dont elle fait partie. Reconnaître que la première lettre d'un mot est un *W* accroît le niveau d'activation de tous les mots commençant par un *W* et diminue le niveau d'activation de tous les autres.

Le modèle d'activation interactive repose sur les hypothèses de la théorie de Rumelhart relatives à la reconnaissance de lettres abordée précédemment. Dans la présentation, chaque trait a une certaine probabilité d'être détecté, variable selon la qualité visuelle de la présentation. Les traits détectés augmentent le niveau d'activation des lettres les possédant et diminuent le niveau d'activation des lettres ne les possédant pas. Les influences excitatrices et inhibitrices se combinent pour déterminer l'activation totale de chaque lettre. Par exemple, détecter une ligne verticale et une ligne diagonale activerait fortement les lettres qui contiennent ces deux traits (comme *K* et *R*).

Tous les psychologues ne considèrent pas que le modèle d'activation interactive soit exact. Massaro et Cohen (1991), en particulier, ont soigneusement comparé les prédictions de ce modèle avec celles d'un modèle dans lequel les informations relatives aux lettres et aux mots se combinent indépendamment plutôt qu'elles n'interagissent. Contrairement au modèle d'activation interactive, illustré par la figure 2.12, leur modèle ne possède pas de connexions descendant du niveau «Lettre» vers le niveau «Mot». Les informations au sujet des mots n'influencent donc pas directement l'activation des lettres. Les données de Massaro et Cohen suggèrent que faire l'hypothèse d'une intégration indépendante des informations issues des niveaux «Lettre» et «Mot» permet plus de prédictions précises que supposer une interaction entre les informations des deux niveaux.

Une critique plus récente du modèle d'activation interactive a vu le jour à cause d'un message électronique que vous avez peut-être vu. Le message contient des exemples de syntagmes formés de mots brouillés qui restent remarquablement lisibles. Par exemple: L'epsirt huiamn lit-il les mtos cmome un tuot? Les exemples démontrent qu'un texte composé de mots dont les lettres sont réarrangées pouvait être assez facilement lu. Cette découverte est un problème pour tous les modèles qui, comme le modèle d'activation interactive, font l'hypothèse que toutes les lettres d'un mot sont placées dans l'ordre exact. Ce qui n'est pas le cas pour les autres modèles qui sont plus souples au niveau de l'agencement des lettres.

J'ai consacré une place considérable au modèle d'activation interactive et vous pourriez vous demander si cela en valait la peine, compte tenu du fait que certains psychologues le rejettent. Je vous répondrai deux choses. D'un point de vue général, il y a très peu de théories (sinon aucune) en psychologie qui n'ont pas été contestées. Les psychologues tentent sans arrêt de formuler de meilleures théories et il y a souvent des débats très nourris pour déterminer laquelle des théories est la meilleure. Plus spécifiquement, le modèle d'activation interactive a eu un impact considérable sur la formulation des théories psychologiques car il a contribué à raviver l'intérêt pour les modèles de réseau neuronal de la cognition. Bien que ces modèles aient été développés pour de nombreuses tâches cognitives, telles que le stockage de l'information dans la MCT (Burgess & Hitch, 1992), la sélection de problèmes analogues dans la résolution de problèmes (Holyoak & Thagard, 1989) et la compréhension de

texte (Kintsch, 1988), ils ont surtout été utilisés pour modéliser la reconnaissance de formes. Considérons maintenant les hypothèses générales de cette approche.

3.3 Les modèles de réseau neuronal

Le modèle d'activation interactive fut le premier pas de McClelland et Rumelhart dans leur mise au point de modèles de réseau neuronal de la cognition. Ils qualifient ceux-ci de modèles de *traitement parallèle distribué* (TPD)♦ parce que l'information y est évaluée en parallèle et distribuée dans tout le réseau. Un *modèle de réseau neuronal*♦ est constitué de plusieurs composants (Rumelhart, Hinton & McClelland, 1986), dont quelques-uns ont déjà été examinés dans le modèle d'activation interactive. Ils comprennent :

1. Une série d'unités de traitement appelées *nœuds*♦. Dans le modèle d'activation interactive, ces nœuds sont des traits, des lettres et des mots. Ils peuvent connaître différents niveaux d'activation.
2. Un patron de *connexions* entre les nœuds. Les nœuds sont reliés les uns aux autres par des connexions excitatrices ou inhibitrices dont la force varie.
3. Des règles d'activation pour les nœuds. Les *règles d'activation*♦ spécifient la façon dont un nœud combine ses inputs excitateurs et inhibiteurs avec son état d'activation du moment.
4. Un état d'*activation*. Les nœuds peuvent être activés à des degrés divers. Nous prenons connaissance des nœuds lorsqu'ils sont activés au-dessus du seuil de conscience. Par exemple, nous sommes avertis de la présence de la lettre K dans le mot WORK lorsque cette lettre reçoit suffisamment d'excitation des niveaux « Trait » et « Lettre ».
5. Les *fonctions d'output* des nœuds. Ces fonctions établissent un lien entre les niveaux d'activation et les outputs – par exemple, elles déterminent le degré d'activation qui permettra d'atteindre le seuil de conscience.
6. Une *règle d'apprentissage*. Généralement, l'apprentissage repose sur un changement de poids des connexions excitatrices et inhibitrices entre les nœuds; le rôle des règles d'apprentissage est de spécifier comment doit s'opérer ce changement.

Le dernier composant, l'apprentissage, est l'une des caractéristiques les plus importantes d'un modèle de réseau neuronal parce qu'il permet au réseau d'améliorer ses performances. Un exemple serait un modèle de réseau neuronal apprenant à mieux discriminer les lettres, en augmentant le poids des traits distinctifs, ceux-là même qui aident le plus la discrimination.

En 1992, l'approche du réseau neuronal a mené à des milliers de recherches et à une industrie qui dépense annuellement des centaines de millions de dollars (Schneider & Graham, 1992). Plusieurs raisons rendent compte de cet engouement. Tout d'abord, un grand nombre de psychologues estiment que cette modélisation dépeint plus fidèlement le fonctionnement du cerveau que ne le font d'autres modèles du comportement, plus séquentiels. Ensuite, un réseau qui régule le poids des inhibitions et des excitations entre les nœuds est capable d'apprentissage et révèle sans doute de la sorte la manière dont nous apprenons. Enfin, ces modèles permettent un autre type de calcul dans lequel beaucoup de contraintes faibles (telles que les

♦ **traitement parallèle distribué (TPD)**

Lorsque l'information est recueillie simultanément à partir de différentes sources et combinée en vue d'une décision.

♦ **modèle de réseau neuronal**

Une théorie qui se sert du réseau neuronal comme métaphore et dans laquelle les concepts (les nœuds) sont reliés les uns aux autres par des connexions excitatrices et inhibitrices.

♦ **nœuds**

Le format des concepts dans un réseau.

♦ **règle d'activation**

Une règle qui détermine comment les connexions excitatrices et inhibitrices se combinent pour déterminer l'activation totale d'un concept.

informations provenant à la fois du niveau « Trait » et du niveau « Mot ») peuvent être prises en compte simultanément.

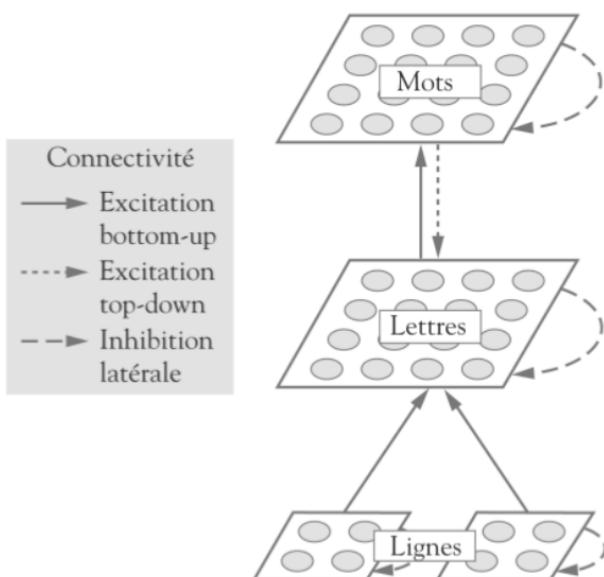
Les modèles de réseau neuronal seraient particulièrement intéressants si leurs hypothèses correspondaient au fonctionnement réel du cerveau. Comme exemple de ce type de modèles, citons celui qui a été développé par Huber et O'Reilly (2003) et qui se fonde sur les hypothèses du modèle d'activation interactive (McClelland & Rumelhart, 1981). Les deux modèles possèdent des niveaux « Trait », « Lettre » et « Mot », mais deux modifications font du modèle de Huber et O'Reilly une meilleure approximation de l'activité neuronale. Premièrement, les connexions de ce modèle sont temporairement sensibles à l'habitude, ce qui entraîne des taux de décharge inférieurs lorsque la durée d'exposition à un stimulus augmente. Deuxièmement, les connexions inhibitrices opèrent cette fois à *l'intérieur* des niveaux (l'inhibition latérale de la figure 2.13) et non plus *entre* les niveaux, ce qui produit des effets de masquage et de compétition entre interprétations alternatives d'un même stimulus.

La figure 2.14 illustre une tâche de reconnaissance de mots dans laquelle un mot-cible (SHADE) apparaît durant 50 msec, avant un masque et une présentation-test. Ce test demande d'identifier lequel des deux mots est apparu – GUEST (distracteur) ou SHADE (cible). Mais la partie la plus importante de l'expérience est la présentation du mot-cible (amorce-cible) ou du mot distracteur (amorce-distracteur), au début de chaque essai. Intuitivement, il semble évident que le fait de présenter l'amorce-cible devrait accroître le nombre de reconnaissances correctes, tandis que la présentation de l'amorce-distracteur devrait le diminuer. Cependant, comme le montre la partie gauche de la figure 2.15, les résultats sont plus complexes et plus intéressants que nos intuitions.

Relevons que l'efficacité relative de l'amorce-cible et de l'amorce-distracteur dépend de la durée de l'amorçage. Pour des durées inférieures à 150 msec, l'amorce-cible est plus efficace que l'amorce-distracteur, mais cet effet s'inverse à partir de 400 msec. L'amorce-distracteur est particulièrement perturbatrice lorsque la durée est de 50 msec parce que le sujet voit deux mots présentés rapidement durant 50 msec et éprouve des difficultés à identifier leur ordre d'apparition. Était-ce GUEST ou SHADE le second mot ? Augmenter la durée de présentation du distracteur facilite la reconnaissance de la cible.

L'aspect curieux des résultats est que l'augmentation de la durée de présentation de l'amorce-cible conduit à l'effet opposé. Pourquoi le fait de disposer de plus de temps pour voir l'amorce-cible rend-il plus difficile la reconnaissance de la cible ? Le concept d'habituation peut nous aider à répondre à cette question. Voir un mot plus longtemps entraîne une inhibition neuronale pour ce mot, ce qui rend sa reconnaissance plus laborieuse lorsqu'il apparaît une seconde fois en tant que cible. Dans le même temps, cette habituation rend la reconnaissance d'autres mots plus aisée parce qu'ils ne sont plus inhibés. Cette interaction entre une habituation pour les connexions excitatrices entre niveaux différents et les connexions inhibitrices au sein des niveaux eux-mêmes, permet au modèle de Huber-O'Reilly de prédire avec succès les résultats obtenus, comme le montre la partie droite de la figure 2.15.

En conclusion, les modèles de réseau neuronal représentent une approche de la modélisation de la cognition. Relier ces modèles aux découvertes effectuées en neurosciences cognitives (Huber *et al.*, 2008) offre des opportunités stimulantes pour améliorer notre compréhension de la cognition et du cerveau.

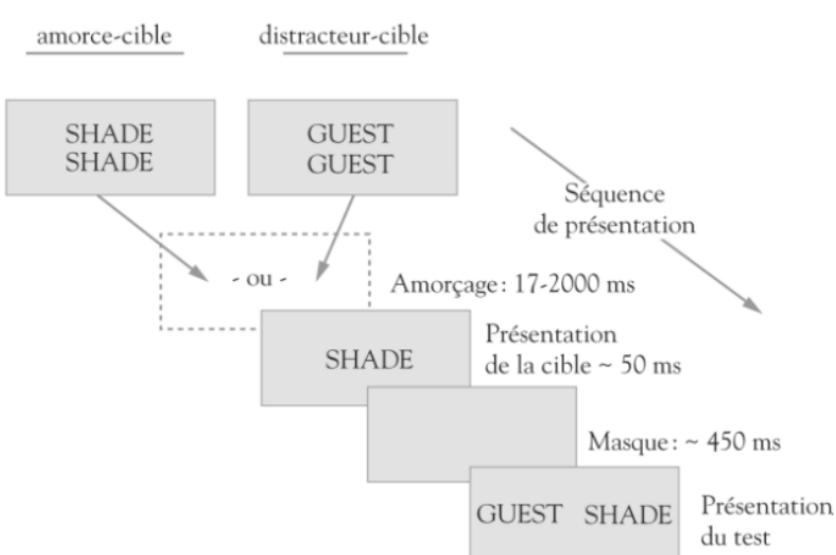
**Figure 2.13**

Le modèle de réseau neuronal développé par Huber et O'Reilly (2003).

Source: figure 3 de « Immediate priming and cognitive aftereffects », D.E. Huber, 2008, *Journal of Experimental Psychology: General*, 137, 324-347.

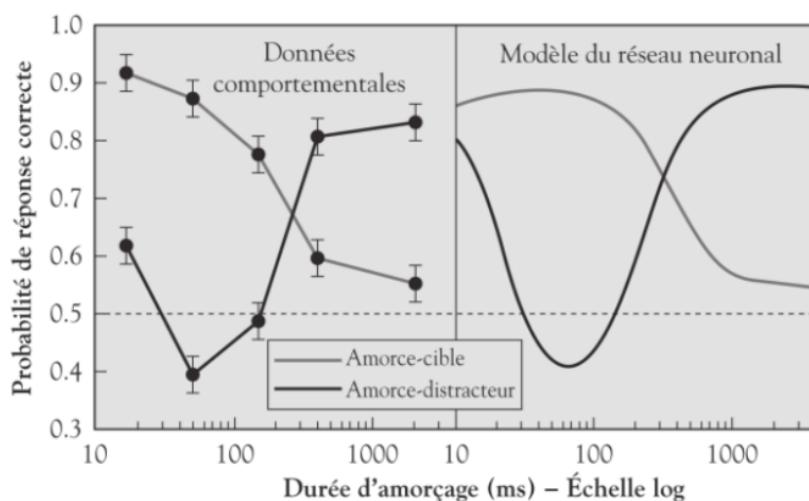
Copyright 2008 de American Psychological Association.

Reproduction autorisée.

**Figure 2.14**

Une méthodologie pour l'amorçage de la reconnaissance de mots.

Source: figure 1 de « Immediate priming and cognitive aftereffects », D.E. Huber, 2008, *Journal of Experimental Psychology: General*, 137, 324-347.

**Figure 2.15**

Résultats réels et simulés de l'expérience d'amorçage de mots de la figure 2.4.

Source: figure 2 de « Immediate priming and cognitive aftereffects », D.E. Huber, X. Tian, T. Curran, R.C. O'Reilly & B. Woroch. (2008).

Journal of Experimental Psychology: Human Perception and Performance, 34, 1389-1416.

Résumé du chapitre

La reconnaissance de formes est une compétence que nous exerçons très bien. Trois théories cherchent à l'expliquer : la théorie des gabarits, la théorie des traits et la théorie structurelle. Une théorie des gabarits propose que nous comparons deux formes en évaluant leur degré de recouvrement. Une théorie des traits éprouve des difficultés à rendre compte de nombreux aspects de la reconnaissance de formes, mais elle est pratique pour représenter l'information dans le registre sensoriel avant que celle-ci ne soit analysée durant l'étape de la reconnaissance de formes. Les théories les plus communes de la reconnaissance de formes supposent que les formes sont analysées à travers leurs traits. La discrimination perceptive nécessite de découvrir quels sont les traits critiques qui permettent de distinguer les formes. Les théories structurelles établissent explicitement la façon dont les traits d'une forme sont reliés. Elles fournissent une description plus complète des formes et sont particulièrement utiles pour la description de formes constituées de lignes entrecroisées.

L'intérêt de Sperling pour le nombre de lettres perceptibles durant une brève exposition par un tachitoscope a conduit à la construction des modèles de traitement de l'information pour les tâches visuelles. Sperling a proposé que l'information est conservée très brièvement par un registre d'information visuel dans lequel toutes les lettres peuvent être simultanément analysées. Lorsqu'une lettre est reconnue, son nom peut être répété verbalement et conservé dans un registre d'information auditif (STM).

Le modèle de Rumelhart propose que nous reconnaissions les formes en identifiant leurs traits. La vitesse d'identification des traits dépend à la fois de la clarté des items dans le registre d'information sensoriel et du nombre de lettres présentées. Le modèle explique les performances au cours de la tâche de rapport partiel de Sperling en supposant que l'observateur focalise son attention sur la rangée indiquée dès qu'il entend le signal sonore. La probabilité de reconnaître plus de lettres dans la rangée dépend de la précision du registre d'information visuel.

La reconnaissance de lettres dans un mot est influencée par les informations perceptives et par le contexte de la lettre. La découverte qu'une lettre est reconnue plus facilement lorsqu'elle fait partie d'un mot que lorsqu'elle fait partie d'un non-mot ou que lorsqu'elle est présentée seule a été appelée l'effet de supériorité du mot. Une modélisation influente de cet effet de supériorité du mot est le modèle d'activation interactive proposé par McClelland et Rumelhart. Son hypothèse de base est que la connaissance des mots d'une langue interagit avec les informations entrantes à propos des traits pour fournir un indice au sujet des lettres qui composent un mot. Cette approche s'est maintenue sous le label général de traitement parallèle distribué et elle s'inspire des modèles du réseau cérébral neuronal. Les développements récents des modèles de réseaux neuronaux se servent des découvertes faites en neuroscience cognitive, afin d'évaluer leurs hypothèses.



Questions

1. Certains pourraient déclarer qu'il n'existe pas de consensus définitif au sujet de ce qu'est une forme. Ne vous découragez pas – écrivez votre propre définition d'une *forme*. Ensuite, cherchez *forme* dans un bon dictionnaire. Cela vous aide-t-il ?
2. Dans quel sens peut-on considérer que la reconnaissance de formes est une compétence ? (Pour répondre à cette question, vous devez disposer d'une définition utilisable de compétence. Quelle est la vôtre ?)
3. Distinguez gabarit, trait et structure. Quelle est la nature de chacun et en quoi diffèrent-ils ?
4. Que nous apprend l'expérience que Phillips a menée en 1974 au sujet des caractéristiques présumées du registre sensoriel ? Quelle est la différence entre une image sensorielle et une image visuelle ?

5. Cet ouvrage présente une application de la théorie des traits qui aide les enfants à apprendre les différences entre lettres similaires. De quoi a-t-on tenu compte pour constituer une « bonne » liste de traits – c'est-à-dire : selon quels critères diriez-vous qu'une telle liste est bonne ?
6. Pourquoi discriminons-nous mieux les visages de notre propre groupe ethnique ? En quoi le concept de « trait distinctif » nous aide-t-il à différencier les visages d'autres groupes ethniques ?
7. En quoi la théorie structurelle va-t-elle au-delà des théories des traits ? Quel résultat prendriez-vous en compte pour soutenir sa prétention à être la meilleure description de la reconnaissance de formes visuelles ?
8. Pourquoi la recherche que Sperling a menée en 1960 est-elle toujours discutée dans la plupart des textes introductifs de psychologie ?
9. Comment fonctionne un cache visuel ? Que fait-il ? Pourquoi cette procédure est-elle essentielle dans l'étude de l'effet de supériorité du mot ?
10. Les modèles de réseau neuronal représentent une approche théorique majeure en psychologie, mais ils se laissent difficilement résumer en termes simples. Expliquez les hypothèses centrales de ces modèles avec vos propres termes.

Mots clés

- | | |
|-----------------------------------|--|
| ✓ caricatures | ✓ procédure de rapport partiel |
| ✓ composant de balayage visuel | ✓ reconnaissance de formes |
| ✓ confusions perceptives | ✓ registre d'information auditif |
| ✓ connexions excitatrices | ✓ registre d'information visuel |
| ✓ connexions inhibitrices | ✓ registre sensoriel |
| ✓ effet de supériorité du mot | ✓ règles d'activation |
| ✓ gabarits | ✓ répétition |
| ✓ géons | ✓ théories des traits |
| ✓ intervalle interstimulus | ✓ théories structurelles |
| ✓ modèle d'activation interactive | ✓ trait distinctif |
| ✓ modèle de réseau neuronal | ✓ traitement parallèle |
| ✓ nœuds | ✓ traitement parallèle distribué (TPD) |
| ✓ paradigme de détection | ✓ traitement sériel |
| ✓ procédure de rapport intégral | |

Lectures pour aller plus loin

Le livre d'Hoffman (1998), *Visual Intelligence*, fournit une analyse lisible et érudite de la façon dont nous construisons les descriptions d'objets. Loftus, Shimamura et Johnson (1985) décrivent les caractéristiques du registre d'information visuel. Fallshore et Schooler (1995) soutiennent que la description verbale de visages peut diminuer la reconnaissance ultérieure parce que les descriptions verbales ignorent les informations relatives à la configuration. D'autres domaines de la recherche sur la reconnaissance de formes incluent la reconnaissance du discours (Poeppel & Monahan, 2008), des visages (McKone, Kanwisher & Duchaine, 2007 ; Palermo & Rhodes, 2007 ; Tarr & Cheng, 2003), des scènes (Braun, 2003 ; Green & Hum-

mel, 2004), et des plans (Tversky, 2005), ainsi que l'enseignement de la lecture (Rayner, Foorman, Perfetti, Psetsky & Seidenberg, 2001). Un recueil de chapitres explore les processus analytiques et holistiques dans la reconnaissance de formes (Petersons & Rhodes, 2003) et Bar (2007) discute la façon dont les composants neurologiques soutiennent les prédictions. L'ouvrage de Martindale (1991), *Cognitive Psychology: A Neural-Network Approach* et celui de Pinker (1997), *How the Mind Works*, offrent une introduction très lisible à l'influence des théories du réseau neuronal sur les théories cognitives.