

Reconnaissance des formes

- Théories de la reconnaissance des formes
 - Théories de l'appariement des gabarits
 - Théories de l'extraction des caractéristiques (Théories des traits)
 - Théories structurales
- Traitement descendant et reconnaissance de formes

© Serge Larochelle

Reconnaissance de formes

- Définition : Étape de la perception durant laquelle un stimulus est identifié.
- Notre capacité à reconnaître les formes est étonnante.
 - Variations de la même forme



Par contre...

- Les ordinateurs éprouvent encore de grandes difficultés à reconnaître des formes.

Méthode

- La reconnaissance des formes se fait généralement de façon automatique.
- Pour étudier les processus de reconnaissance des formes,
 - les chercheurs ne peuvent pas utiliser des rapports introspectifs verbaux;
 - les chercheurs utilisent :
 - Une présentation rapide des stimuli
 - Une présentation floue, dégradée des stimuli
 - Pour étudier la nature des erreurs commises
 - Pour analyser les TR (temps de réponse) quand très peu d'erreurs sont commises.

- La reconnaissance d'une forme implique de mettre en correspondance une forme perçue (i.e., projetée sur la rétine) avec celles contenues en mémoire à long terme (MLT).
- Cette mise en correspondance peut se faire sur la base plusieurs types d'information.
- 3 types de théories :
 - Théories de l'appariement des gabarits
 - Théories de l'extraction des caractéristiques
 - Théories structurales

Théories à base de gabarits

- Description
- Problèmes

L'appariement de gabarits

- Gabarit : Forme non analysée, représentation globale d'une forme. (Analogie: un moule).
- Appariement de gabarits: processus par lequel une forme donnée est comparée (superposée) aux gabarits en MLT pour trouver à quel gabarit la forme ressemble le plus.
- Variable importante: Le taux de recouvrement entre une forme et un gabarit.
- La forme est reconnue comme étant une instance du gabarit le plus semblable
- Traitement global, holistique des formes.

Problèmes

- Problème 1 :
 - La forme projetée sur la rétine et le gabarit doivent être :
 - de la même taille
 - dans la même orientation
 - dans la même position
 - Il existe des transformations analogiques qui permettent de résoudre ces problèmes
- Problème 2 :
 - Peu informative sur la distinction entre 2 formes.
 - Les lettres P et R sont différentes, mais comment le système cognitif fait-il la différence ?
- Problème 3
 - La grande variabilité des formes (pensez à l'écriture manuelle)
 - ex: A A A A A A A A A A A A A A A A ...
 - Un gabarit pour chaque variante de chaque forme (pas très parcimonieux)

Théories à base de caractéristiques

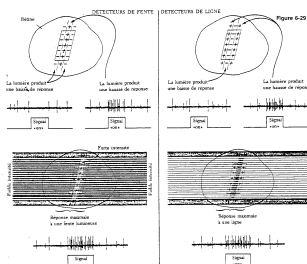
- Description
- Neurophysiologie (Hubel & Wiesel)
 - Détecteurs spécialisés dans le cerveau
 - Hyper-colonnes
- Psychologie
 - Les caractéristiques des lettres (E. Gibson)
- Conclusion

L'extraction des caractéristiques

- Cette théorie définit les formes en décrivant les différentes parties (les caractéristiques) qui les composent.
- Ex : $H = | + | + -$
- Les caractéristiques de la forme perçue sont ensuite comparées avec les caractéristiques des formes contenues en MLT.
- Traitement analytique, local des formes

Neurophysiologie

- Hubel & Wiesel (1962, 1977) ont découvert
 - des cellules qui répondent seulement à une arête, fente, ligne :
 - de largeur x
 - d'orientation y
 - de position z dans le champs visuel
 - organisation en colonne des détecteurs spécialisés



E. Gibson (1969)

- L'apprentissage perceptuel des formes se ferait donc par la découverte des *traits distinctifs*.
- Les *différences* d'un objet par rapport à un autre seraient ainsi particulièrement importantes.
- Ex: Discriminer le E et le F... la barre horizontale du bas permet de discriminer les 2 lettres.

Caractéristiques des lettres (E. Gibson, 1969)

- 4 critères de sélection des caractéristiques importantes des lettres :
 - ❶ Les caractéristiques doivent être discriminatoires.
 - ❷ Les caractéristiques ne changent pas selon les conditions visuelles (taille, perspective etc.).
 - ❸ L'ensemble de caractéristiques produit une forme unique pour chaque lettre.
 - ❹ Le nombre de caractéristiques ne doit pas être trop élevé.

Caractéristiques des majuscules

Alphabet	A	E	F	H	I	L	T	K	M	N	V	W	X	Y	Z	B	C	D	G	J	O	P	R	Q	S	U
Traits																										
Droite																										
horizontale	+	+	+	+	+	+	+									+		+					+	+		
verticale	+	+	+	+	+	+	+	+	+	+	+					+										
diagonale /	+							+	+	+	+	+	+	+												
diagonale \	+								+	+	+	+	+	+									+	+		
Courbe																										
fermée																+	+			+	+	+	+	+		
ouverte V																				+						+
ouverte H																				+	+	+				
Intersection	+	+	+	+			+	+								+						+	+	+		
Redondance																										
changement cyclique	+								+				+			+									+	
symétrie	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+			+					+
Discontinuité																										
verticale	+	+	+	+	+	+	+	+	+	+	+				+							+	+			
horizontale	+	+	+	+	+	+	+								+											

Le niveau de confusion des lettres

- Les caractéristiques sont habituellement établies sur la façon dont elles rendent compte des confusions perceptuelles.
- Les items susceptibles d'être confondus doivent avoir de nombreuses caractéristiques communes.

Gibson et al. (1963)

- Expérience A :
 - Lecture des 26 lettres chez des enfants de 4 ans.
 - Variable dépendante : Erreurs, on regarde quelles lettres sont confondues.
- Problèmes :
 - Ils font peu d'erreurs
 - Beaucoup de lettres n'ont jamais été confondues

Gibson et al. (1968)

- Discrimination de paires de lettres :
 - 2 lettres apparaissent à l'écran pendant une brève période de temps.
 - La tâche est de dire si les lettres sont identiques ou différentes.
- Enfants âgés de 7 ans et adultes.
- Variables dépendantes : TR et erreurs (assez peu)
- Ex chez adultes : G W = 458 ms ; P R = 571 ms

Les théories à base de caractéristiques: Conclusion

- Avantages:
 - Économiques: nombre d'unités de base à stocker en mémoire est très réduit.
 - Flexibles : moins sensible aux différences de tailles, de position et de perspective.
- Problèmes
 - Représentations sont incomplètes: Relations entre caractéristiques
 - Représentations sont-elles suffisantes pour représenter objets du monde réel?

Théories structurales

- Description
- Les principes de la Gestalt
- La reconnaissance par composantes
 - Les géons
 - Les relations entre géons
- Tests de la théorie des géons (Biederman)

Les théories structurales

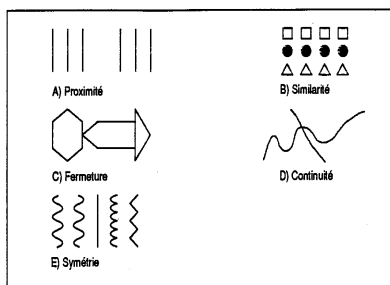
- On peut faire beaucoup de formes avec :
 - Deux lignes verticales
 - Une ligne horizontale
- H : $| - |$, $|| -$, $- ||$, $_ ||$, $\overline{||}$...
- Il est donc important de spécifier comment les lignes sont reliées !

Les théories structurales

- Le tout est plus important que la somme de ses parties.
- La reconnaissance des formes dépend en général de la façon dont nous interprétons les relations entre les arêtes, les surfaces et les composantes d'un objet.
- Les théories structurales tentent de prendre en compte ces relations.

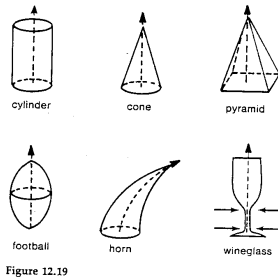
L'extraction des arêtes et les surfaces (Les principes de la gestalt)

- Proximité
- Similarité
- Fermeture
- Continuité
- Symétrie



Les géons

- Les géons sont des représentations bidimensionnelles de formes géométriques tridimensionnelles simples :
 - Cylindres
 - Blocs
 - Cônes



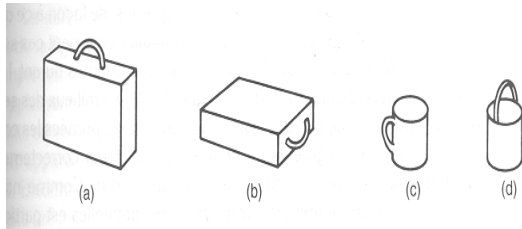
La nature des géons

- Les géons résultent de la combinatoire de propriétés élémentaires:
 - Arête de la base (2)
 - Droite; exemple: Carré, Rectangle
 - Courbe; exemple: Cercle, Ellipse
 - Symétrie de la base (3): Base invariante sous
 - Rotation et réflexion; exemple: Carré, Cercle
 - Réflexion; exemple: Ellipse, Rectangle
 - Asymétrique; exemple: Triangle avec côtés inégaux
 - Changement de taille de la base (3)
 - Constante; exemple: Cylindre
 - Expansion; exemple: Cone
 - Expansion et contraction; exemple: Ballon de football américain
 - Courbure de l'axe perpendiculaire à la base (2)
 - Droit; exemple: Cone
 - Courbe; exemple: Corne
- Il y a donc $2 \times 3 \times 3 \times 2 = 36$ géons différents

Analogie avec le langage

- Un petit nombre de phonèmes (44) permet de dire tous les mots du vocabulaire,
 - Ex : i, o, p, en etc.
- Les phonèmes résultent de la combinatoires de traits élémentaires (voisement, place de l'articulation, manière),
 - Ex : ba-pa, da-ta, ga-ka
- Selon Biederman, les géons résultent de la combinatoires de traits (propriétés) élémentaires.
- Un petit nombre de géons (36) permet de représenter toutes les formes de l'environnement.

Les relations entre géons (aimer vs marie)



Les relations entre deux géons

- Taille relative de G1 et G2 (3):
 - $G1 > G2$
 - $G1 < G2$
 - $G1 = G2$
- Position relative de G1 par rapport à G2 (2.4 car utile pour 80% des objets)
 - Sur le dessus
 - Sur le dessous
 - Sur le côté
- Nature de la jonction entre G1 et G2 (2)
 - Bout à bout
 - Bout à côté
- Jonction sur quelle surface de G1 (2)
 - Longue surface
 - Courte surface
- Jonction sur quelle surface de G2 (2)
 - Longue surface
 - Courte surface

La puissance de représentation

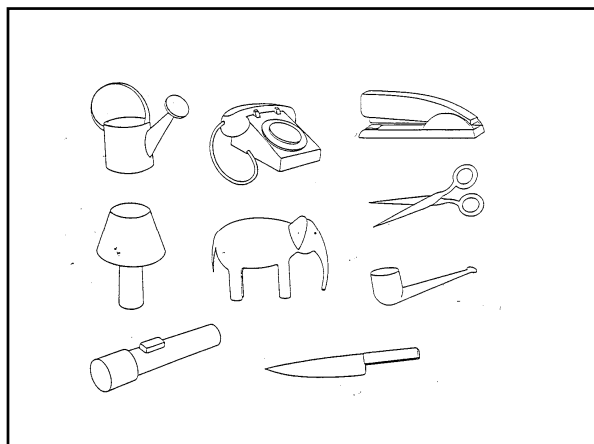
- Nombre de relations possibles:
 - $3 \times 2.4 \times 2 \times 2 = 57.6$
- Nombre d'objets possibles
 - deux géons
 - $36(G1) \times 36(G2) \times 57.6 = 76,649$
 - trois géons
 - $76,649 \times 36(G3) \times 57.6 = 154 \text{ millions}$

Nombre d'objets représentés en mémoire

- Dans le dictionnaire, il y a environ 1500 mots désignent des objets de base. Disons 3000, en cas d'erreur.
- Supposons qu'il y ait 10 formes différentes pour chaque objet de base
 - Différentes perspectives sur objets naturels
 - Différentes formes pour artefacts
- Un totale de 30,000 objets représentés en mémoire (ce qui correspond à l'apprentissage de 13.5 objets par jour jusqu'à 6 ans)

Suffisance de la théorie

- Possible de reconnaître des objets, même complexes, à partir de seulement quelques géons.
- Quand tous les géons sont présentés, la reconnaissance est aussi rapide qu'avec une photographie en couleur de l'objet.

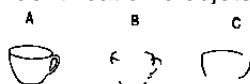


Nécessité de la théorie

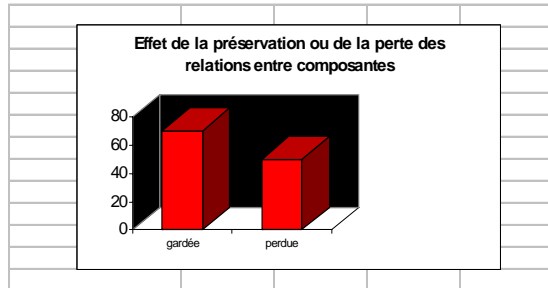
- Selon la théorie de Biederman,
 - si on empêche d'extraire les relations entre les géons, on devrait diminuer la capacité des sujets à reconnaître la forme.

Biederman (1985)

- 65% du contour des objets sont dégradés :
 - Conditions :
 - A. Les objets non-dégradés (tous au début).
 - B. La relation entre les geons est préservée.
 - C. La relation entre les geons n'est pas préservée.
- Tâche : Identification d'objets



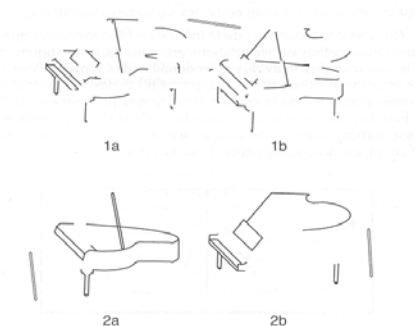
Résultats



Biederman & Cooper (1991)

- Suppression de 50% des contours.
- Conditions :
 - 2 images complémentaires avec les mêmes géons (un même cylindre $\frac{1}{2}$, $\frac{1}{2}$).
 - 2 images complémentaires avec des géons différents.

Amorçage d'objet



Prédictions et résultats

- Identifier correctement 1a facilitera l'identification subséquente de 1b.
- Par contre, identifier correctement 2a ne facilitera pas l'identification subséquente de 2b.
- C'est ce qui est retrouvé!

Les théories structurales: Conclusion

- S'intéressent aux objets en 3D.
- Il faudrait voir si les théories peuvent rendre compte de la reconnaissance d'objets plus complexes, ainsi que de la reconnaissance de scènes visuelles.

Traitement descendant et reconnaissance de formes

- Définitions
- Démonstrations
- Effet de supériorité du mot
 - Modèle de McClelland & Rumelhart
- Effet de supériorité de l'objet
 - Weisstein & Harris
 - Palmer
- Effet de contextes plus larges

Types de traitements

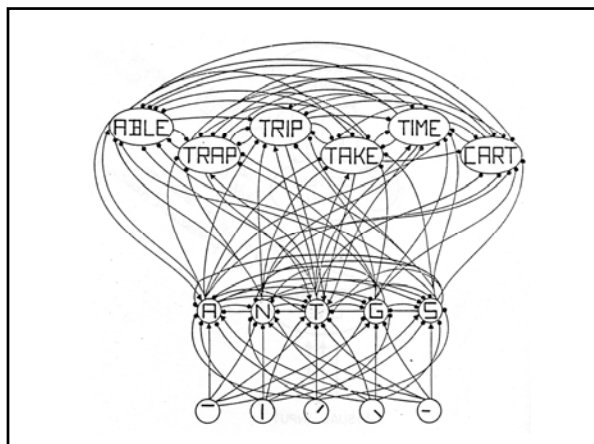
- Traitement ascendant : Traitement dirigé par les données (bottom-up).
- Traitement descendant : Traitement dirigé par les concepts (top-down).

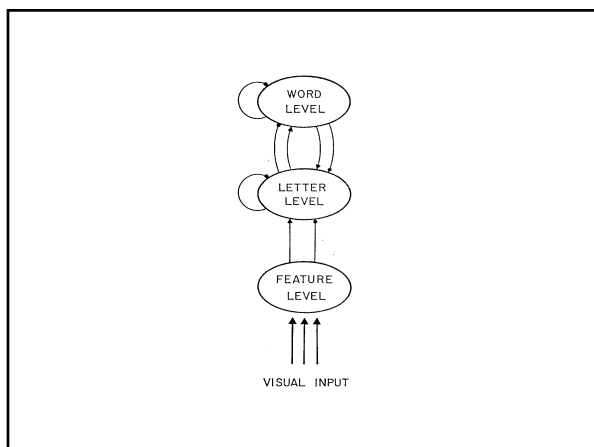
Traitement descendant et reconnaissance de formes

- Notre connaissance de la façon dont le monde est organisé nous aide à identifier les formes.
 - On reconnaît plus rapidement et mieux les formes ambiguës lorsqu'elles sont dans un contexte cohérent.
 - Désavantage: risque d'erreurs.

Effet de supériorité du mot

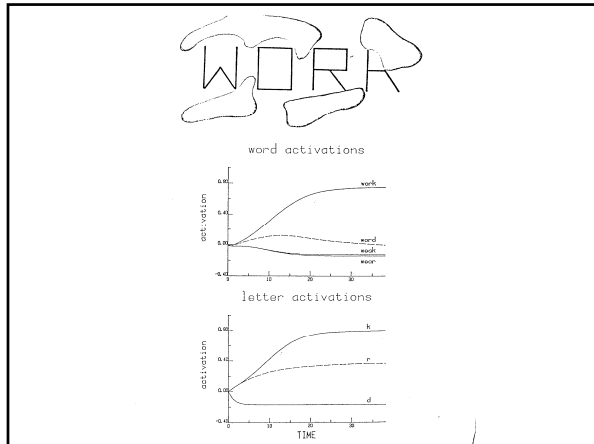
- On identifie mieux et plus rapidement une lettre lorsqu'elle est dans un mot que lorsqu'elle apparaît dans une suite aléatoire de lettres.

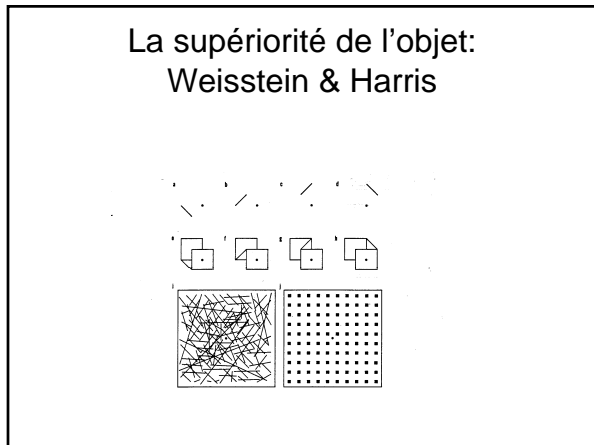


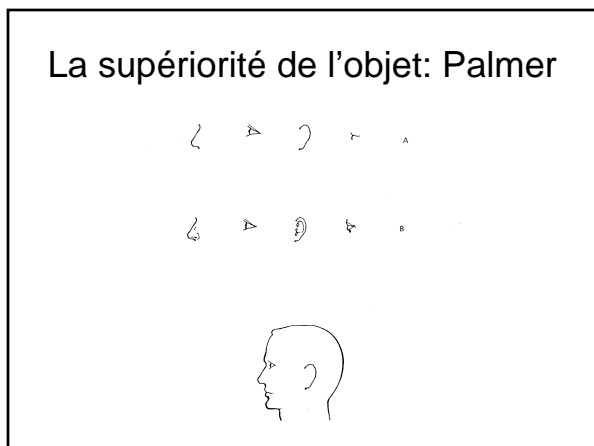


Le modèle de McClelland & Rumelhart

- Réseau neuronal
- Trois niveaux de représentation
 - caractéristiques
 - lettres
 - mots
- Connexions excitatrices et inhibitrices
- Connexions ascendantes, latérales et descendantes
- Activation des nœuds :
 - Activation + Inhibition = état d'activation
- Seuil de décision



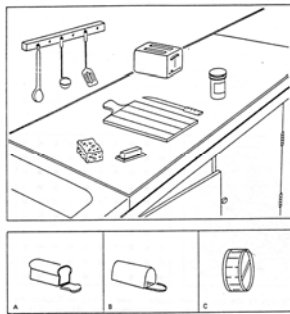




Effets de contexte plus larges

- Démonstrations
- Influence de la phrase sur la reconnaissance des mots
- Influence de la scène sur la reconnaissance des objets
 - Palmer

Influence de la scène... Palmer (1975)



Palmer (1975)

