

## Les théories connexionnistes

Le connexionnisme constitue une approche récente pour comprendre le développement cognitif. Comme les systèmes de production, les théories connexionnistes permettent de simuler sur ordinateur le fonctionnement de la pensée. La principale raison de la popularité croissante des modèles connexionnistes est leur analogie générale avec le fonctionnement neurobiologique du cerveau (on parle de *réseaux neuro-mimétiques*, censés mimer le fonctionnement des neurones du cerveau).

Les modèles connexionnistes possèdent plusieurs caractéristiques fondamentales (Plunkett, 1996) :

1. Ils sont composés d'un grand nombre d'unités simples de traitement qui ressemblent aux neurones du cerveau.
2. Les unités de traitement sont constituées de deux ou plusieurs couches organisées hiérarchiquement (cf. Fig. 1). Typiquement, celles-ci incluent une couche d'entrée, dont les unités de traitement encodent la représentation initiale de la situation, une ou plusieurs couches cachées, dont les unités combinent l'information à partir des unités d'entrée, et une couche de sortie, dont les unités produisent la réponse du système à la situation.
3. Les unités simples de traitement sont connectées à d'autres unités de traitement dans différentes couches (et quelquefois à l'intérieur de la même couche). Le poids des connexions varie avec l'expérience du système et est cruciale pour déterminer le traitement effectué.
4. Comme dans le cerveau, une unité de traitement donnée s'active lorsque le niveau de stimulation qu'elle reçoit de toutes les autres unités de traitement auxquelles elle est connectée dépasse un certain seuil. Le niveau de stimulation qu'une unité reçoit de chaque unité à laquelle elle est connectée est déterminé, d'une part, par le degré d'activation de l'unité de traitement qui envoie l'activation et, d'autre part, par le poids de la connexion entre les unités.
5. Comme dans le cerveau, l'activité de la plupart des unités simples de traitement se produit en parallèle (simultanément).
6. La connaissance est représentée par le poids des connexions au sein de toutes les unités du système. Une partie, ou un élément du système, ne correspond pas à une connaissance particulière. Au contraire, la connaissance est distribuée sur toutes les unités et leurs interconnexions. C'est pour cette raison, et parce que le traitement se produit sur de nombreuses unités en parallèle, que ces systèmes sont souvent désignés comme des systèmes de traitement distribué parallèle ou, en anglais, *Parallel distributed processing* (PDP).
7. L'apprentissage se produit à partir du système qui reçoit les données, élabore une réponse, observe la différence entre la réponse fournie et la réponse correcte et ajuste le poids des connexions entre les unités de traitement afin de produire une meilleure réponse. Les ajustements incluent le renforcement de certaines connexions et l'affaiblissement d'autres. Par ce processus d'ajustement, le système apprend implicitement les règles sous-jacentes aux réponses correctes du problème, bien qu'il n'y ait pas un seul endroit où les règles soient explicitement représentées.
8. La généralisation de la connaissance du système est fondée sur la similitude de nouvelles situations à celles que le système a rencontrées précédemment. Quand les mêmes types de règles implicites s'appliquent à de nouveaux problèmes, les systèmes connexionnistes sont très efficaces pour généraliser les expériences antérieures.

De nombreux chercheurs défendent des modèles connexionnistes du développement : McClelland (1995) ; Shultz, Schmidt, Buckingham et Mareschal (1995) ; Plunkett (1996) et Marchman (1992), entre autres. Un modèle connexionniste particulièrement impressionnant, et qui illustre la force de cette approche, est celui de Mac Whinney, Leinbach, Taraban et McDonald (1989) pour rendre compte de l'apprentissage de l'article défini (en fonction du mot qui suit) chez des enfants allemands (cf. fig. 1).

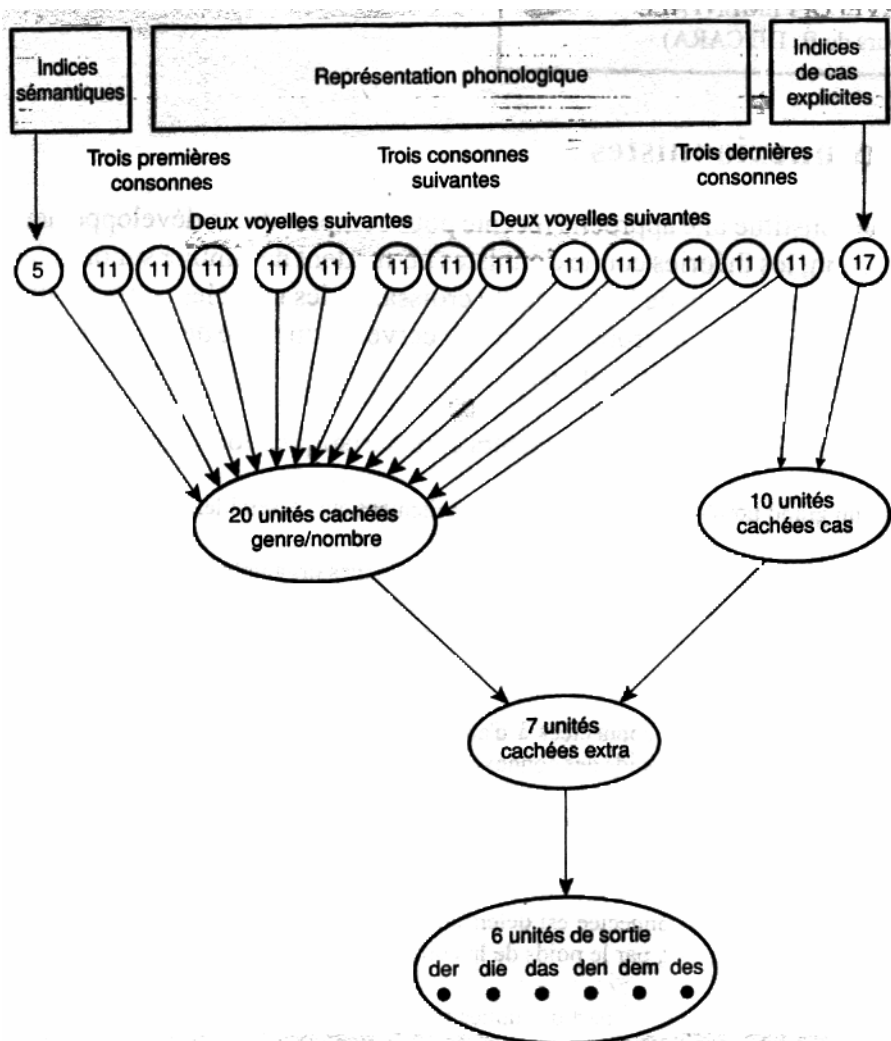


Fig. 1. Modèle connexionniste de MacWhinney et al. (1989) sur la façon dont les enfants apprennent le système des articles en langue allemande. Remarquez qu'au niveau supérieur (entrée), le modèle encode cinq caractéristiques sémantiques du nom (correspondant au « 5 » situé à l'extrême gauche du niveau qui se trouve en dessous du niveau supérieur), la présence ou l'absence de 11 caractéristiques phonologiques aux 13 localisations possibles dans le mot (également représentées au niveau qui se trouve en dessous du niveau supérieur) et 17 indices de cas explicites, qui indiquent la fonction du nom dans la phrase. Ces unités du niveau d'entrée transfèrent l'activation aux unités cachées qui, elles-mêmes, la transmettent aux 6 unités de sortie correspondant aux six articles accompagnant les noms en langue allemande. L'article qui accompagne l'unité de sortie la plus activée est proposé comme étant la réponse par le système (d'après MacWhinney et al., 1989).

Le modèle de MacWhinney et al. (Fig. 1) décrit l'apprentissage par des enfants allemands des articles définis dans leur propre langue. Ces articles définis sont des termes qui remplissent en allemand la même fonction que mot « le » en français. La tâche est intéressante précisément parce que le système des articles en allemand est très difficile à apprendre. Le choix de l'article à utiliser pour déterminer un nom donné dépend du genre du nom (masculin, féminin ou neutre), de son nombre (singulier ou pluriel) et de son rôle dans la phrase (sujet, possesseur, objet direct ou objet indirect). Pour compliquer les choses, l'attribution des noms aux catégories de genre est souvent contre intuitive. Par exemple, le mot *'fourchette'* est féminin, le mot *'cuillère'* est masculin et le mot *'couteau'* est neutre. Les relations sont si complexes qu'elles semblent presque impossibles à apprendre. Cependant, MacWhinney et al. ont élaboré un modèle connexionniste montrant comment les enfants pouvaient les apprendre.

Le modèle de MacWhinney et al. comme la plupart des modèles connexionnistes comprend une couche d'entrée, plusieurs couches cachées et une couche de sortie (cf. Fig. 1). Chacune de ces couches contient un nombre d'unités discrètes. Par exemple, dans le modèle, les 35 unités dans la couche d'entrée représentent les caractéristiques du nom particulier auquel on cherche à associer un article. Chaque couche cachée inclut des unités qui représentent des combinaisons de ces caractéristiques du niveau d'entrée. Les six unités de sortie représentent les six articles définis possibles en langue allemande.

Une caractéristique centrale des modèles connexionnistes est le très grand nombre de connexions entre les unités de traitement. Dans le modèle de MacWhinney et al., chaque unité de la couche d'entrée est connectée aux unités cachées du premier niveau ; chaque unité cachée du premier niveau est connectée aux unités cachées du second niveau ; et chaque unité cachée du second niveau est connectée à chacune des six unités de sortie. L'apprentissage se produit à travers un cycle du système qui (1) reçoit l'entrée initiale (dans ce cas, un nom dans un certain contexte) ; (2) projette quelle sortie produire sur la base du poids de ses diverses connexions (qui reflète l'expérience passée) ; (3) active cette réponse ; et (4) ajuste le poids des connexions entre les unités afin que celles qui ont suggéré la réponse correcte soient renforcées et celles qui ont suggéré la réponse incorrecte soient diminuées.

MacWhinney et al. ont testé la capacité de ce modèle à maîtriser le système des articles en langue allemande en lui présentant de manière répétée 98 noms communs allemands. La simulation nécessitait à choisir l'article à utiliser pour un nom donné dans un contexte particulier, c'est-à-dire, dans un contexte consistant à vouloir exprimer une signification particulière avec des mots particuliers. Après que ce choix ait été effectué, la réponse correcte était présentée et la simulation ajustait le poids des connexions afin d'optimiser sa précision pour le futur.

Grâce à l'expérience acquise avec cette série d'entraînement, la simulation de MacWhinney et al. a permis de choisir l'article correct pour plus de 90 % des noms de la série originale. Ce résultat ne pouvait pas être attribué à un apprentissage par coeur associant chaque nom à l'article qui l'accompagne. Lorsqu'un nom rencontré précédemment était à nouveau présenté à la simulation mais dans un contexte légèrement différent, le modèle avait tendance à fournir la même réponse que celle donnée dans le contexte original.

Par ailleurs, la simulation s'est également montrée capable de généraliser l'apprentissage à des noms nouveaux. Même lorsqu'elle n'avait jamais rencontré le terme présenté, la simulation pouvait utiliser le son et la signification du mot pour estimer raisonnablement l'article à utiliser.

L'apprentissage par simulation connexionniste est équivalent sur plusieurs aspects à l'apprentissage des enfants. Tôt dans le processus d'apprentissage, la simulation, tout comme les enfants dont la première langue est l'allemand, tend à utiliser exagérément les articles qui accompagnent les noms féminins, vraisemblablement parce que cette forme d'article est utilisée le plus souvent dans la langue. De plus, les combinaisons article-nom les plus difficiles pour les enfants allemands sont également les plus difficiles à apprendre pour la simulation. Les erreurs commises par la simulation ressemblent à celles des enfants.

Les modèles connexionnistes ont aussi décrit avec succès d'autres acquisitions développementales, comme la permanence de l'objet (Munakata, McClelland, Johnson, & Siegler, sous presse), la compréhension du problème d'équilibre de la balance, les problèmes de rapport temps vitesse, ainsi que le raisonnement causal (Shultz et al., 1995), l'acquisition de la lecture (Plaut, McClelland, Seidenberg & Patterson, 1995), l'apprentissage d'une langue seconde (MacWhinney, 1996) et l'acquisition du sens des mots et de la compréhension grammaticale (Elman, 1993 ; MacWhinney & Chang, 1995 ; Marchman, 1992 ; Plunkett & Sinha, 1992).

Comme toutes les théories, les approches connexionnistes sont sujettes à la critique. Un reproche fréquent est que les approches connexionnistes se prétendent « biologiquement plausibles ». Toutefois, les théories connexionnistes n'évoquent pas l'activité chimique qui est cruciale dans le fonctionnement du cerveau. Par ailleurs, le fonctionnement des unités simples de traitement ne possède qu'une similitude abstraite à celui des neurones.

Une autre limite est que l'apprentissage est *extrêmement lent* et nécessite beaucoup plus d'expériences que chez les êtres humains. On n'observe pas le type de compréhension soudaine que présentent parfois les individus (Raijmakers, Koten, & Molenaar, 1996).

Une troisième limite, liée à la seconde, est que ces modèles n'apprennent pas les règles symboliques formulées de façon formelle, comme les formules mathématiques ou les règles de l'orthographe enseignées à l'école (Pinker & Prince, 1988). D'un autre côté, les modèles connexionnistes ont permis de modéliser les nombreux développements qui ne dépendent pas de l'apprentissage de règles explicites. Bien que le fonctionnement de tels systèmes diffère clairement de celui du cerveau, il lui ressemble davantage que les autres approches qui utilisent, elles aussi, les simulations sur ordinateur comme les systèmes de production. Les modèles connexionnistes ont été particulièrement utiles pour modéliser des domaines comme la perception et le langage, dans lesquels de nombreuses sources d'information partiellement valides doivent être intégrées pour obtenir des performances réussies. Du fait de ces avantages, il n'est pas surprenant que ces modèles connaissent une popularité grandissante.

Source: Siegler, R. S. (2001). *Enfant et raisonnement : Le développement cognitif de l'enfant*. Traduction de Béatrice Bourdin et Clara Martinot (pp. 103-107). Bruxelles : De Boeck Université.