

LE ROLE DE L'ATTENTION DANS LES APPRENTISSAGES IMPLICITES

Pierre Perruchet

Laboratoire d'Etude des Apprentissages et du Développement

Université de Bourgogne

1.- Le paradoxe

Tous les pédagogues, suivant en cela le sens commun, attribuent un rôle favorable à l'attention dans les processus de mémoire et d'apprentissage. Qu'il s'agisse d'apprendre une leçon d'histoire ou une formule mathématique, la condition primordiale semble être d'accorder le maximum d'attention au matériel à apprendre.

A examiner la question de plus près, toutefois, il est facile de trouver des exemples où un apprentissage semble s'opérer en dehors de toute attention. Un exemple privilégié est celui de l'acquisition de la langue maternelle. Bien que certaines règles grammaticales soient apprises à l'école, il reste que l'enfant sait normalement parler avant de porter attention à ces règles, et que la compréhension ou la production du langage par l'adulte obéit à des règles qu'il n'a jamais étudiées. Ainsi, pour la plupart des adultes de langue française, la différence entre un savant inconnu et un savant-inconnu (le trait d'union notifiant une liaison audible) est immédiate: il s'agit dans le premier cas d'un savant ignoré du monde, et dans le second, d'un inconnu qualifié de savant. La règle qui sous-tend cette compréhension, à savoir, selon la grammaire de Grévisse, que la liaison ne se fait pas après la consonne finale d'un nom au singulier, n'est jamais un objet d'étude.

Des exemples de ce genre peuvent être multipliés. Ainsi dans le domaine de la physique, chacun de nous est capable d'anticiper plus ou moins correctement le point de chute d'une balle en voyant le début de son parcours, et cette performance est complètement indépendante de l'étude formelle de la balistique. Prenons un dernier exemple dans le domaine de la musique. N'importe quel individu de culture occidentale est capable d'évaluer correctement, en regard des règles de la musique tonale, si une note donnée complète de façon cohérente un fragment musical, sans que l'attention consacrée aux cours de solfège pèse beaucoup dans l'exactitude de sa réponse. Tous ces cas, tirés presque au hasard d'une multitude d'exemples possibles, sont intrigants, car ils suggèrent que la plupart des "savoir-faire" fondamentaux, à l'inverse de nombreuses connaissances, s'acquerraient sans attention.

Mon objectif est de tenter de résoudre cet apparent paradoxe. Est-il vrai que l'attention, qui semble si utile dans la plupart des situations d'apprentissage intentionnellement dirigé, soit inutile, ou du moins facultative, pour les apprentissages qui fondent notre adaptation comportementale à la structure complexe du monde, qu'il s'agisse de ses aspects linguistiques, physiques, ou culturels? La solution que je proposerai est celle que nous offre un domaine de recherche récent,

centré sur ce qu'il est convenu d'appeler depuis une vingtaine d'années: l'apprentissage implicite". Pour anticiper, l'idée défendue sera que l'attention est bien le moteur qui sous-tend toutes formes d'apprentissage, et que les apparences contraires sont dues à une croyance erronée portant sur la nature de ce qui est appris.

2- L'apprentissage implicite

L'ambition de ceux qui participent aux recherches sur l'apprentissage implicite est de simuler en laboratoire les apprentissages auxquels nous venons de faire référence, ceux qui constituent nos savoir-faire les plus fondamentaux en dehors de toute attention apparente portée à leur objet. Les sujets sont placés face à une situation structurée par des règles trop complexes pour qu'elles puissent être découvertes durant les quelques dizaines de minutes ou les quelques heures d'une session expérimentale. Les sujets n'ont pas pour consigne de rechercher ces règles, dont ils ignorent même l'existence. Les règles en question sont arbitraires, de façon à être totalement nouvelles pour le sujet. Cette caractéristique est imposée par l'objectif poursuivi, qui est d'analyser les processus d'adaptation en temps réel.

Plusieurs grandes catégories de situations sont utilisées, parmi lesquelles nous n'en examinerons qu'une, désignée sous le nom de "tâches de temps de réaction séquentiel". Une cible est présentée à l'écran d'un ordinateur, sous la forme d'une étoile par exemple, et la tâche du sujet consiste à pister cette cible le plus rapidement possible sur le clavier. La cible peut occuper un certain nombre de positions prédéfinies, et certaines touches du clavier sont choisies de façon à être en correspondance spatiale avec la position des cibles à l'écran. Généralement, il y a quatre positions possibles, disposées en ligne. A chaque essai, une cible apparaît, et le sujet doit appuyer le plus rapidement possible sur le bouton correspondant à cette cible. La cible disparaît, et une nouvelle cible apparaît après un temps bref, de l'ordre de 250 msec.

Le déplacement de la cible est apparemment aléatoire. En réalité, il est contraint par des règles, qui sont de nature variée selon les expériences. Dans bien des cas, la règle ne consiste qu'à répéter la même séquence d'une dizaine de positions (e.g. Verlut, intra). Dans d'autres cas, les règles définissant les positions sont plus complexes. Nous en verrons un exemple plus loin. Quel que soit le type de règles, les temps de réaction témoignent d'une prise en compte précoce, de la part du sujet, du caractère structuré de la séquence. Concrètement, l'adaptation du sujet est mesurée par l'amélioration du temps de réaction, estimée par comparaison avec une situation contrôle ou la séquence est réellement aléatoire. On observe, dans ce paradigme, que les temps de réactions décroissent pour tous les sujets, signe d'une adaptation générale à la situation : mais ils décroissent davantage pour les sujets réagissant aux séquences structurées que pour les sujets réagissant aux séquences aléatoires. Pourtant, interrogés en fin d'expérience, la plupart des sujets confrontés aux séquences structurées avoue n'avoir rien remarqué de particulier, et sont surpris d'apprendre que des règles sous-tendaient l'organisation de ces séquences.

Bien évidemment, ce genre de résultat n'est pas isolé. Il est retrouvé dans les autres tâches de laboratoire, où les sujets sont familiarisés avec une tâche complexe

en condition incidente d'apprentissage (pour des ouvrages de synthèse, on peut consulter Cleeremans, 1993, Berry et Dienes, 1993, et Reber, 1993). Certains résultats tendent même à montrer que si l'on compare les performances de sujets qui apprennent en conditions incidentes, comme ci-dessus, à celle de sujets dont l'attention est orientée vers la présence de règles, et qui sont incités à les rechercher, la performance des premiers s'avère meilleure que celle des seconds. Ceci tendrait à suggérer que, si une situation est de toute façon trop complexe pour qu'une recherche intentionnelle des règles ait une chance d'aboutir, il est préférable d'adopter une attitude passive, en se laissant "imprégner" par les règles.

3- Les situations de simple répétition

Considérons tout d'abord les situations où une même séquence, comprenant entre 5 et 16 essais selon les études, est continuellement répétée. Les performances des sujets s'améliorent, sans que leur attention ne se soit portée sur le caractère cyclique des événements. Certes, on sait depuis l'introduction de ces situations sur la scène expérimentale que l'attention portée à la tâche est nécessaire, ou du moins exerce un effet favorable, à l'obtention d'une amélioration sélective des performances. Nissen et Bullemer (1987), qui sont à l'origine de ce paradigme, ont démontré cela par l'usage d'une double tâche. Les sujets avaient pour tâche principale de réagir à la localisation des cibles, qui suivaient la séquence DBCACBDCBA (où A, B, C et D représentent les positions de gauche à droite). De plus, ils entendaient des sons, soit graves, soit aigus, après chaque essai, et ils devaient compter le nombre de sons graves. Cette tâche ajoutée a eu pour effet de supprimer, ou diminuer, l'avantage que les sujets tiraient de la répétition de la séquence, par rapport à une séquence aléatoire. Ce résultat a été maintes fois répliqué depuis.

Mais le paradoxe demeure. En effet, il reste que les sujets, en dépit de l'attention accordée à la tâche en général, restent toujours inconscients, dans leur majorité, du caractère répétitif de la séquence. Tout se passe comme si l'attention était nécessaire, sans qu'il soit besoin qu'elle se porte sur l'objet même de l'apprentissage. L'attention se porterait sur les caractéristiques de surface du matériel pour permettre à un processeur inconscient d'analyser ses caractéristiques structurales. Ce genre d'interprétation, conférant aux processus inconscients un certain côté magique et tout-puissant, a eu quelque succès, et continue parfois d'être défendu.

Il apparaît pourtant probable aujourd'hui que les mécanismes en jeu ne soient pas de ce type. L'erreur est de croire que les sujets doivent apprendre qu'une même séquence est continuellement répétée. Or cette connaissance est nullement indispensable à l'amélioration sélective des performances. En effet, dans les séquences utilisées, la position de la cible à un essai donné est généralement prédictible, au moins de façon probabiliste, à partir de l'essai précédent, et peut toujours être déterminée avec certitude si l'on considère les deux, voire les trois essais immédiatement antérieurs. Le fait que les sujets n'aient pas conscience du caractère répétitif d'une longue séquence n'a donc aucune pertinence: il suffit qu'ils soient sensibles à des règles de dépendance séquentielle d'ordre limité. A cette étape

du raisonnement, on pourrait arguer que cette éventualité ne résout toujours pas le paradoxe initial selon lequel l'apprentissage ne semble pas porter sur l'objet de l'attention. En effet, l'attention des sujets ne semble pas être davantage focalisée sur l'analyse des règles séquentielles que sur la recherche de la répétition d'une longue séquence.

Si l'on s'interroge, à ce point, sur l'objet de l'attention des sujets, c'est à dire sur la façon qu'ils ont de traiter attentionnellement la tâche qui leur est présentée, il semble vraisemblable de supposer, compte tenu de ce que l'on sait des mécanismes attentionnels, qu'ils perçoivent la séquence comme une succession de fragments, composés de quelques localisations successives décrivant, par exemple, des sauts, des alternances, des effets de symétrie, ou des suites régulières. Dans la séquence utilisée par Nissen et Buller (1987) et beaucoup d'autres à leur suite, il est probable que beaucoup de sujets aient leur attention captée par la suite DCBA, qui représente un déplacement régulier de droite à gauche. Supposons que les sujets perçoivent cette suite comme une unité, en conséquence directe du traitement attentionnel qu'elle suscite, et se familiarisent avec cette unité lors de sa répétition. Le point crucial consiste à remarquer que ceci revient à apprendre plusieurs règles séquentielles d'ordre 1: D est suivi de C, C est suivi de B, et B est suivi de A. En fait, chacune des positions D, C et B peut être aussi suivie par une autre position dans le restant de la séquence, mais puisque l'éventail des suites possibles est néanmoins restreint, cet apprentissage peut suffire à améliorer le temps de réponse par rapport à une séquence totalement aléatoire. De plus, la suite DCBA intègre deux règles séquentielles d'ordre deux, qui permettent d'anticiper B et A, respectivement, de façon certaine.

L'hypothèse selon laquelle le traitement attentionnel de la séquence aboutit à un découpage en fragments incorporant un certain nombre des règles séquentielles structurant la situation résout notre paradoxe: dans ce scénario, c'est bien l'objet de l'attention du sujet qui constitue l'objet de l'apprentissage. Cette hypothèse a divers supports expérimentaux, parmi lesquels on peut citer les résultats d'une expérience récente de Stadler (sous presse). Dans la tâche secondaire la plus souvent utilisée, les sujets entendent un son après chaque essai, mais ne doivent incrémenter le nombre de sons graves que sur une partie des essais. Stadler remarque que la présentation de sons graves et aigus étant aléatoire, cette interruption du traitement dévolu à la tâche principale n'arrive pas toujours au même point dans la série. Ainsi la séquence DCBA peut être successivement découpée par le son grave comme D/CBA, puis DCB/A, etc.. De ce fait, le sujet ne peut former une unité perceptive "DCBA", à laquelle il deviendrait de plus en plus familier. D'une certaine façon, il se trouve devant la tâche d'apprendre un numéro de téléphone lu une fois 85 48 96 52, et une autre fois 854 89 65 2. Stadler compare les performances de deux groupes de sujets. Pour un premier groupe, la tâche ajoutée demande de l'attention, c'est-à-dire détourne une partie des capacités de traitement, sans toutefois nuire au découpage de la séquence. Il s'agit de maintenir en mémoire à court terme une suite de lettres lue au début de chaque bloc, pour la restituer en fin de bloc. Pour un second groupe, la modification de procédure ne demande aucun détournement d'attention, mais elle interfère avec un découpage cohérent de la séquence. Cet objectif est atteint en

allongeant à 2 secs certains intervalles entre essais, habituellement fixés à 400 ms. Les résultats sont nets: c'est la condition dans laquelle les variations d'intervalles introduisent un découpage variable qui affecte négativement les performances.

La conclusion qui se dégage est que, paradoxalement, l'attention est au coeur des apprentissages implicites de séquences répétées. C'est la répétition du codage attentionnel, celui là même qui forge notre appréhension phénoménale de la réalité, qui induit sa facilitation. Certes, ces apprentissages demeurent "implicites", et cela en plusieurs sens. Tout d'abord, ils n'impliquent aucune intention d'apprendre de la part du sujet. D'autre part, la facilitation liée à la répétition est entièrement automatique, et n'implique pas que la répétition soit consciemment détectée et traitée comme telle. Mais ce sont néanmoins les processus attentionnels qui jouent le rôle déterminant: lorsque le sujet apprend, c'est ce qui forme l'objet même de sa perception phénoménale du monde qui se modifie, sous l'effet de la répétition.

4- Les situations structurées par des règles complexes

A première vue, le raisonnement précéderait ne s'applique qu'aux situations où une même suite d'événements est répétée, et ne concerne pas les situations où les événements sont déterminés par des règles génératives abstraites, qui n'induisent pas de répétition à grande échelle. Or il y a des raisons de penser que la même interprétation s'applique aux deux cas de figures. Pour comprendre comment une telle généralisation est possible, examinons en détail un exemple.

Dans cet exemple (Lewicki, 1988), les sujets doivent, comme précédemment, "pister" une cible pouvant prendre quatre positions. Celles-ci sont disposées en carré, et les sujets doivent utiliser quatre touches du pavé numérique du clavier, elles aussi disposées en carré, pour leurs réponses. Du point de vue du sujet, la série des positions apparaît comme une longue suite continue de 3600 essais. Cette suite est en fait composée d'une succession de "blocs logiques" de cinq essais, au sein desquels des règles spécifiques s'appliquent. Dans chaque bloc logique, les deux premiers essais sont toujours aléatoires. Il y a 4X3, soient 12 possibilités, car une cible n'apparaît jamais deux fois à la même position. Puis chacun des trois derniers essais est complètement déterminé à partir des deux essais qui le précèdent immédiatement. Ainsi, la localisation du 5^e essai de chaque bloc logique est définie par une suite de règles du genre: "si la cible effectue un mouvement horizontal en passant du 3^e au 4^e essai, elle doit effectuer un mouvement diagonal en passant du 4^e au 5^e essai". Puisque trois types de déplacements sont possibles, horizontal, vertical, et diagonal, trois règles s'appliquent en fait, et ces règles diffèrent pour chacune des trois dernières positions. Il est inutile d'ajouter que la découverte de ces règles, sans aucune indication préalable sur leur nature, apparaît comme une tâche effroyablement complexe. La première étape de segmentation en blocs de cinq essais, étape nécessaire à la découverte des règles de dépendances séquentielles, apparaît déjà difficilement surmontable.

Les temps de réaction tendent à décroître avec les essais, un effet sans doute imputable à une familiarisation non spécifique avec le dispositif et la tâche. Mais, plus intéressant, le TR aux essais 3, 4, et 5, prédictibles à partir des règles, décroît

plus rapidement que les TR aux essais 1 et 2, aléatoires. Ce résultat surprenant a été obtenu dans l'expérience originelle de Lewicki et al (1988), et reproduit par Perruchet, Gallego et Savy (1990), ainsi qu'il est montré en Figure 1. Les sujets sont donc sensibles à la structure de la situation. Ils ont acquis cette sensibilité sans prêter, apparemment, la moindre attention à cette structure, dont ils sont d'ailleurs incapables de formuler quoi que ce soit au terme de l'expérience.

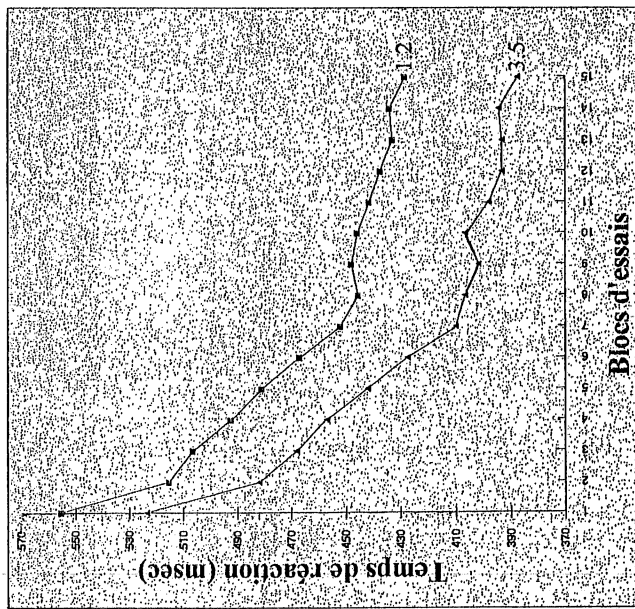


FIGURE 1 : Chaque bloc d'essais comprend 48 séries de 5 essais successifs. Au sein de chaque série, les temps de réactions décroissent moins vite pour les essais 1 et 2, aléatoires, que pour les essais 3 à 5, où la position de la cible est prédictible à partir de règles complexes dont les sujets n'ont aucune connaissance consciente (d'après Perruchet et al., 1990)

Si l'on s'interroge à nouveau sur l'objet effectif de l'attention des sujets dans ce type de tâche, il semble vraisemblable de supposer que la séquence est subjectivement perçue, non pas comme un cheminement continu de la cible, mais comme une succession d'"unités", telles que des mouvements horizontaux, verticaux, ou diagonaux, ou des enchaînements de ces mouvements de base, définissant des allers-retours, des déplacements triangulaires, ou d'autres configurations saillantes. Or le point capital est qu'il suffit de postuler que les sujets perçoivent les séquences sous cette forme, c'est-à-dire comme une succession d'unités discrètes, et qu'ils sont

sensibles à la fréquence avec laquelle ces unités apparaissent, pour rendre compte des résultats présentés plus haut.

Pour comprendre ce curieux phénomène, il est nécessaire de revenir à la nature particulière des règles sous-tendant la situation. L'examen de ces règles montre, par exemple, qu'elles prohibent l'apparition de mouvements d'aller et retour dans les trois derniers essais de chaque bloc logique (ainsi, aucune règle n'est du genre: si le mouvement de n-2 à n-1 est horizontal, alors le mouvement de n-1 à n est horizontal). Des mouvements d'aller et retour sont par contre présents dans les deux premiers essais de chaque bloc. Puisque ceux-ci sont aléatoires, à ceci près qu'il n'y a pas de répétition, ils représentent 33% des déplacements possibles sur ces essais. Le résultat général de cette distribution est que, sur l'ensemble des essais, les mouvements d'aller-retour sont moins nombreux que ce que le hasard permettrait d'attendre (sur l'ensemble de la série, le pourcentage effectif n'est que de 10.8 %). Compte tenu de ce que l'on sait sur les variations de TR en fonction de la fréquence de la cible en situation de choix multiple, il est permis d'anticiper qu'avec l'entraînement sur une structure de ce type, les TR lors d'allers-retour seront plus longs que les TR correspondant à d'autres séquences, nécessairement plus fréquentes, par complémentarité. Or ces TR longs, par nécessité logique, seront localisés sur les essais 1 et 2 de chaque bloc. En effet, l'application des règles rend à la fois les allers-retour peu fréquents sur l'ensemble de la série, et présent uniquement sur les essais aléatoires. Les résultats empiriques ne sont pas dus qu'à la distribution des allers-retour, car d'autres facteurs entrent en jeu. Mais ces facteurs sont du même ordre. Démontrer que les résultats empiriques sont entièrement dus à des facteurs de ce genre impliquerait une analyse plus poussée, et le lecteur intéressé pourra se reporter à l'article original (Perruchet, Gallego, & Savy, 1990). Il y est montré que la familiarisation avec de petites unités correspondant au codage attentionnel de la série rend compte de l'ensemble des performances, plus finement encore que l'hypothèse d'une abstraction passive des règles.

Ainsi, dans cette expérience au moins, l'apparente sensibilité des sujets à des règles abstraites s'explique par l'effet de la fréquence d'occurrence d'événements qui correspondent aux unités subjectives de codage. Il convient de noter que ce résultat est obtenu alors même qu'apparemment il n'existe pas de relations entre les règles déterminées par l'expérimentateur, et la fréquence des événements consciemment perçus comme tels par les sujets, parce que faisant l'objet de focus attentionnels successifs.

5- Conclusion

Le paradoxe soulevé en introduction était que certaines formes d'apprentissage, parmi les plus fondamentales, semblaient s'opérer en dehors de toute attention, alors que l'attention semble nécessaire pour la moindre opération de mémorisation, en situation scolaire par exemple. En fait, les données issues des expériences d'apprentissage implicite suggèrent qu'il s'agit là d'une illusion. L'attention est sans doute nécessaire à toute forme d'apprentissage. Ce que les données expérimentales nous montrent, c'est que des phénomènes initialement considérés comme témoignant d'une capacité d'apprentissage sans attention, sont en

fait réductibles à des changements induits dans le codage attentionnel de l'information. Ceci est vrai lorsque la structure de l'environnement consiste en la simple répétition des événements passés, mais également, bien que cela ne soit pas immédiatement évident, lorsque la structure de l'environnement répond à des règles abstraites extrêmement complexes. Ce chapitre ne concernait qu'une situation d'apprentissage implicite particulière, dite d'apprentissage séquentiel, mais il serait possible de montrer que cette conclusion se généralise aisément aux autres situations d'apprentissage implicite communément utilisées en laboratoire (pour une analyse développée, voir Perruchet et Gallego, sous presse).

Tout le problème consiste maintenant à se demander si ces conclusions, issues de données de laboratoire, sont généralisables au monde réel. Bien des arguments militent pour le contraire. Ainsi, on peut remarquer que la complexité des situations expérimentales et naturelles n'a pas de commune mesure. Même si la situation décrite plus haut est hors de portée des capacités d'analyse d'un sujet dans le contexte où elle est présentée, on aurait peine à soutenir que sa complexité est du même ordre que celle d'une langue naturelle par exemple. Les durées durant lesquelles l'apprentissage est censé s'opérer sont aussi très différentes. Enfin, et peut-être plus fondamentalement, on peut arguer que les situations de laboratoire ne sont pas pertinentes, parce qu'elles impliquent des règles arbitraires, alors que les règles "naturelles", celles qui structurent notre environnement physiques par exemple, peuvent répondre à des nécessités intrinsèques, pour lesquelles l'organisme serait préparé par l'histoire évolutive. Chacun de ces arguments mériterait une longue discussion.

A défaut, il convient de souligner le caractère spéculatif de notre conclusion, dès lors qu'elle se généralise du laboratoire au monde réel. Mais cette spéculation peut s'alimenter à certaines sources de données. Parce que la plupart des apprentissages fondamentaux s'opèrent durant l'enfance, Perruchet et Vinter (sous presse) ont examiné comment la conception de l'apprentissage ici développée peut orienter une réinterprétation des données de la littérature concernant le développement de l'enfant. Le résultat est encourageant: bien des phénomènes semblant dénoter l'existence de connaissances implicites précoces, acquises sans qu'aucune attention n'ait jamais été apportée à leur contenu, apparaissent interprétables en termes similaires à ceux qui permettent la réinterprétation des données expérimentales collectées chez l'adulte placé en situation d'apprentissage implicite. En résumé, dans cette perspective, le fait que l'acquisition des savoir-faire fondamentaux semble s'opérer en dehors de toute attention est une illusion due à une erreur sur l'objet même de l'apprentissage. Ce ne sont pas des règles, ou des connaissances abstraites sur la nature du réel qui sont acquises. Le fruit de l'expérience est une capacité à coder le monde de façon efficace, et l'attention est au cœur du mécanisme qui permet cette adaptation.

Références

- Berry, D.C., & Dienes, Z. (1993). *Implicit Learning: Theoretical and empirical issues*. Hillsdale, N.J.: Lawrence Erlbaum Associates (pp. 197).
- Cleeremans, A. (1993). *Mechanisms of Implicit Learning: A connectionist model of sequence processing*. MII press: Bradford Books (pp.227)
- Lewicki, P., Hill, T. & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, 20, 24-37.
- Perruchet, P., & Gallego, J. (sous presse). A subjective unit formation account of implicit learning. In D. Berry (Ed.). *How implicit is implicit learning*. Oxford: Oxford University Press.
- Perruchet, P., Gallego, J. & Savy, I. (1990). A critical reappraisal of the evidence for unconscious abstraction of deterministic rules in complex experimental situations. *Cognitive Psychology*, 22, 493-516.
- Perruchet, P., & Vinter, A. (sous presse). Learning and development: The implicit knowledge assumption reconsidered. In M. I. Stadler & P. A. Frensch (Eds), *Handbook of implicit learning*, Thousand Oaks, CA: Sage publications.
- Reber, A.S. (1993). *Implicit learning and tacit knowledge*. Oxford: Oxford University Press (pp. 188).
- Stadler, M.A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 674-685.
- Adresse de l'auteur : Laboratoire d'Etude des Apprentissages et du Développement
Faculté des Sciences
Université de Bourgogne
6 Bd. Gabriel
21000 DIJON
e-mail: perruche@u-bourgogne.fr