

Effet du type d'agencement temporel des répétitions d'une information sur la récupération explicite

Par Émilie GERBIER

Thèse de doctorat en Sciences cognitives

Sous la direction d'Olivier KOENIG

Présentée et soutenue publiquement le 20 mai 2011

Membres du jury : Olivier KOENIG, Professeur des universités, Université Lyon 2 Liliane MANNING, Professeur des universités, Université de Strasbourg Serge CARBONNEL, Professeur des universités, Université de Savoie Rémy VERSACE, Professeur des universités, Université Lyon 2 Denis BROUILLET, Professeur des universités, Université Montpellier 3

Table des matières

Contrat de diffusion . . .	5
[Dédicaces] . . .	6
Remerciements . . .	7
Résumé . . .	9
Abstract . . .	10
Préambule . . .	11
Première partie : Théorie . . .	15
Chapitre 1 : Les effets de pratique distribuée dans la littérature . . .	15
1. 1. Les travaux pionniers d'Ebbinghaus sur la répétition . . .	15
1. 2. L'effet d'espacement . . .	15
1. 3. L'effet d'intervalle . . .	17
1. 4. Effets de différents agencements temporels . . .	22
Chapitre 2 : Les théories psychologiques de l'effet de pratique distribuée . . .	29
Introduction . . .	29
2. 1. Théorie de la variabilité de l'encodage . . .	31
2. 2. Théorie de la récupération en phase d'étude . . .	41
2. 3. Modèles mixtes combinant variabilité de l'encodage et récupération en phase d'étude . . .	47
2. 4. Traitement déficitaire, amorçage, accessibilité et reconstruction . . .	49
2. 5. Explications métacognitives . . .	52
2. 6. Conclusion . . .	53
Chapitre 3 : Neurosciences cognitives de l'effet de pratique distribuée . . .	53
3. 1. Apport de l'imagerie cérébrale pour l'étude de l'effet de pratique distribuée . . .	54
3. 2. Modélisation mathématique . . .	58
3. 3. Consolidation et taux d'oubli . . .	64
Conclusion et problématique . . .	74
Deuxième partie : Expériences . . .	76
Introduction . . .	76
Chapitre 4 : Effet de l'agencement temporel des répétitions à long terme sur la mémoire . . .	78
4. 1. Création du matériel expérimental . . .	78
4. 2. Expérience 1 . . .	80
4. 3. Expérience 2 . . .	88
4. 4. Expérience 3 . . .	109
Chapitre 5 : Effet de la variation du contexte . . .	130
5. 1. Expérience 4 . . .	130
Troisième partie : Discussion générale . . .	146
Chapitre 6 : Discussion générale . . .	146
6. 1. Synthèse des résultats . . .	146
6. 2. Implications pour les théories de l'effet de pratique distribuée . . .	147

6. 3. Délai de rétention et taux d'oubli . .	152
6. 4. Effet de l'agencement selon l'échelle de temps . .	157
6. 5. Généralisation des résultats . .	158
6. 6. Discussion des concepts utilisés . .	161
6. 7. Perspectives de recherche . .	162
6. 8. Applications potentielles . .	163
Conclusion générale . .	165
Références . .	166
Notes bibliographiques . .	176
Annexes . .	177
<u>Annexe n°1</u> : Caractéristiques des mots utilisés au sein des paires, selon Bonin et al. (2003). . .	177
<u>Annexe n°2</u> : Liste des paires distractrices utilisées dans les Expériences 1 et 2 . .	178
<u>Annexe n°3</u> : Feuille de route de l'Expérience 1. . .	178
<u>Annexe n°4</u> : Items présentés aux participants du groupe A dans l'Expérience 1 . .	179
<u>Annexe n°5</u> : Items présentés aux participants du groupe A dans l'Expérience 2 . .	180
<u>Annexe n°6</u> : Items présentés aux participants du groupe A dans l'Expérience 3 . .	181
<u>Annexe n°7</u> : Items présentés dans l'Expérience 4 . .	182
<u>Annexe n°8</u> : Formulaire utilisé dans l'Expérience 2 . .	183
<u>Annexe n°9</u> : Scores de motivation / concentration / performance en fonction des groupes de DR dans l'Expérience 3 . .	184
<u>Annexe n°10</u> : Questionnaire de débriefing de l'Expérience 4. . .	185
<u>Annexe n°11</u> : Détails sur les activités réalisées dans l'Expérience 4 . .	186

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas d'utilisation commerciale - pas de modification](http://creativecommons.org/licenses/by-nc-nd/2.0/fr/) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer, l'adapter ni l'utiliser à des fins commerciales.

[Dédicaces]

À mon père

À mon grand-père

Remerciements

Je souhaite remercier l'ensemble des personnes qui, d'une façon ou d'une autre, m'ont aidée à faire aboutir ce travail.

En premier lieu, je tiens à remercier chaleureusement le professeur Olivier Koenig, pour avoir encadré mon travail pendant maintenant six ans depuis le Master 1. Merci pour votre optimisme permanent, pour m'avoir toujours encouragée, et pour m'avoir laissée toute liberté dans mon travail. Merci d'avoir pris du temps pour me faire partager vos connaissances, vos compétences et votre vision des choses. J'ai énormément appris à vos côtés. Plus largement, je remercie le laboratoire EMC et l'Université Lyon 2, ainsi que le ministère de la Recherche pour m'avoir permis de bénéficier d'une certaine sécurité financière pendant mes trois années d'allocation ministérielle.

Je remercie chaleureusement les membres du jury, Madame Lilianne Manning, Messieurs Denis Brouillet, Serge Carbonnel, et Rémy Versace, qui m'ont fait l'honneur de lire et d'évaluer mon travail, et ce dans des délais particulièrement serrés. Je remercie également François Pellegrino et Serge Carbonnel pour leurs commentaires éclairés dans le cadre de mon comité de suivi de thèse.

Je remercie également la société Climats Médias pour m'avoir permis de dédommager les participants des Expériences 1 à 3, sans quoi ces expériences n'auraient pas pu être conduites. Je remercie Eric Ortéga et Joël Brogniart pour leur aide technique pour la mise au point du site internet de l'Expérience 5 qui n'a malheureusement jamais pu voir le jour. Merci également à Jonathan Fayard et Carolina Saletti pour m'avoir accompagnée dans les passations de l'Expérience 4. Merci également à tous les participants des études qui ont donné de leur personne pour faire « avancer la science ».

Au plan personnel, je remercie ma famille et en premier lieu je tiens à remercier ma mère pour m'avoir permis d'étudier, depuis toujours, dans les conditions matérielles les plus favorables, et d'avoir toujours respecté mes choix d'études. J'embrasse également ma sœur Julie et ses futurs enfants, et également ma cousine Célia. Je remercie et je rends hommage à ma grand-mère, mon grand-père et mes tantes Martine et Françoise, pour leur affection sans faille, leur soutien de tous les instants, leurs encouragements à étudier dur, leur goût du travail bien fait, et leur intérêt pour le bien commun. Tout ce sans quoi je ne serais pas arrivée jusque là !

J'embrasse affectueusement mon fiancé Sébastien, et je le remercie particulièrement pour son soutien pendant la période difficile de la fin de thèse. Merci à Nicole, Guy, pour leur accueil si chaleureux. Des bisous à Karine, Laurent, et aux enfants.

Un merci général à tous mes amis du laboratoire, pour leur écoute et leur soutien. Je remercie infiniment Elodie pour son amitié et son aide énorme apportée à la fin de ma thèse, tant psychologiquement que par ses corrections éclairées. Merci à tous les doctorants pour l'ambiance festive qui a régné au laboratoire pendant toutes ces années, tout en permettant une émulation scientifique tellement riche : Mathieu, Damien, Benoît, Lionel L., Lionel B., Hélène, Sarah-Lise, Jordan, Elodie P., Céline, Mélaïne, Audric, Claire C. Merci à tous pour avoir fait que nos conditions de travail soient si agréables ! Remerciements spéciaux à Filio, Adrien, Guillaume, Arnaud, Anna, Nina, et German pour leur aide et leurs conseils qui m'ont beaucoup aidée. Merci aux huit participants de l'étude contraignante sur l'effet du sommeil, qui n'a pas pu être incluse dans le présent document : Nina, Jordan, Adrien, Elodie P., Claire, Anna, Arnaud et Sarah-Lise.

Merci aux anciens doctorants du laboratoire et aux membres (très) actifs d'Estigma, pour nous avoir montré la voie et inspirés, en particulier Norbert, François-Xavier, Benjamin P., Adel et Elise. Longue vie scientifique à vous !

Merci aux enseignants-chercheurs du laboratoire, pour tous les échanges que j'ai pu avoir avec eux, je pense en particulier à Stéphanie Mazza, Emmanuelle Reynaud, Jean-Claude Bougeant, Georges Michael, et Rémy Versace.

Je remercie enfin mes amis pour leur écoute et leur soutien : Anne-Gaëlle, Sandrine, Léa, Bérengère, Carine, Claire B., Lisianne, et Benjamin S. Que la fin de la thèse me permette de vous voir davantage !

Merci enfin à tous ceux que j'ai oubliés !

Résumé

La façon dont une information se répète au cours du temps a une influence sur la façon dont nous nous souviendrons de cette information. Les recherches en psychologie ont mis en évidence l'effet de pratique distribuée, selon lequel on retient mieux les informations qui se répètent avec des intervalles inter-répétitions longs que celles qui se répètent avec des intervalles inter-répétitions courts. Nos travaux ont porté spécifiquement sur les situations où l'information se répète sur plusieurs jours, et nous avons comparé l'efficacité relative de différents types d'agencement temporel des répétitions. Un agencement uniforme consiste en des répétitions se produisant à intervalles réguliers, un agencement expansif en des répétitions se produisant selon des intervalles de plus en plus espacés, et un agencement contractant en des répétitions se produisant selon des intervalles de plus en plus rapprochés. Les Expériences 1 et 2 consistaient en une phase d'apprentissage d'une semaine et ont révélé la supériorité des agencements expansif et uniforme après un délai de rétention de deux jours. L'Expérience 3 consistait en une phase d'apprentissage de deux semaines, et les sujets étaient ensuite testés lors de trois délais de rétention différents (2, 6 ou 13 jours). La supériorité de l'agencement expansif sur les deux autres agencements est apparue progressivement, suggérant que les différents agencements induisaient des taux d'oubli différents. Nous avons également tenté de tester différentes théories explicatives des effets de l'agencement temporel des répétitions sur la mémorisation, en particulier les théories de la variabilité de l'encodage (Expérience 4) et de la récupération en phase d'étude (Expérience 2). Les résultats observés tendent à confirmer la théorie de la récupération en phase d'étude. Dans les conclusions de notre travail, nous insistons sur l'importance de la prise en compte des apports des autres disciplines des sciences cognitives, en particulier les neurosciences et la modélisation, dans l'étude de l'effet de pratique distribuée.

Abstract

How information is repeated over time determines future recollection of this information. Studies in psychology revealed a distributed practice effect, that is, one retains information better when its occurrences are separated by long lags rather than by short lags. Our studies focused specifically on cases in which items were repeated upon several days. We compared the efficiency of three different temporal schedules of repetitions: A uniform schedule that consisted in repetitions occurring with equal intervals, an expanding schedule that consisted in repetitions occurring with longer and longer intervals, and a contracting schedule that consisted in repetitions occurring with shorter and shorter intervals. In Experiments 1 and 2, the learning phase lasted one week and the retention interval lasted two days. It was shown that the expanding and uniform schedules were more efficient than the contracting schedule. In Experiment 3, the learning phase lasted two weeks and the retention interval lasted 2, 6, or 13 days. It was shown that the superiority of the expanding schedule over the other two schedules appeared gradually when the retention interval increased, suggesting that different schedules yielded different forgetting rates. We also tried to test major theories of the distributed practice effect, such as the encoding variability (Experiment 4) and the study-phase retrieval (Experiment 2) theories. Our results appeared to be consistent with the study-phase retrieval theory. We concluded our dissertation by emphasizing the importance of considering findings from other areas in cognitive science—especially neuroscience and computer science—in the study of the distributed practice effect.

Préambule

La **thématique générale** du présent travail est celle de comprendre l'effet de la répétition d'une information sur la mémoire. Il est de sens commun que le fait de rencontrer ou d'étudier une même information à plusieurs reprises est plus utile à sa rétention ultérieure qu'une unique rencontre. De plus, nous voudrions tous améliorer ou optimiser nos apprentissages, dans le cadre sociétal actuel qui pousse l'individu à réaliser des performances toujours plus élevées, et ce en un temps minimal. Un grand nombre de travaux ont étudié les situations dans lesquelles l'information est présentée, puis présentée à nouveau, puis testée, au sein d'une session expérimentale unique. Il a été découvert que la durée de l'intervalle temporel séparant les présentations successives d'une même information a une grande influence sur la rétention ultérieure. Des effets robustes relatifs à l'effet de pratique distribuée ont été mis en évidence depuis les années 1950. Les effets de pratique distribuée seront décrits dans le premier chapitre. Plus précisément, le présent travail porte sur un sous-ensemble de la problématique de l'effet de pratique distribuée. Nous nous sommes intéressée à la question de la répartition de trois occurrences ou plus d'une information verbale sur une période de plusieurs jours. Par exemple, un étudiant sérieux qui apprend et révise une liste de vocabulaire dans une langue étrangère va étendre son travail d'apprentissage sur plusieurs jours et non pas sur quelques minutes avant l'examen. La particularité de notre travail est donc, d'une part, liée au nombre d'occurrences, qui est plus élevé que ce qui est étudié classiquement dans la littérature, et d'autre part, au fait que l'échelle temporelle est également plus longue que dans la majorité des études existantes.

Le présent document est scindé en trois grandes parties : une partie théorique composée de trois chapitres, une partie expérimentale composée de deux chapitres, et une dernière partie consacrée à la discussion générale. Dans la partie théorique, nous nous sommes attachée à décrire de la façon la plus exhaustive possible les effets de pratique distribuée ; cette recherche nous a menée vers divers champs disciplinaires. Bien entendu, le champ auquel nous serons le plus fortement rattachée est celui de la psychologie cognitive.

Le **premier chapitre** sera consacré à une revue de littérature relative à l'effet de la pratique distribuée en général. Après un rapide historique relatif aux travaux d'Ebbinghaus (1885), nous décrirons brièvement la notion d'*effet d'espacement* (e.g., Melton, 1967), loi selon laquelle il est plus bénéfique pour la mémoire d'espacer les présentations répétées d'un item que de les masser, c'est-à-dire de les répéter en succession immédiate. Au-delà de l'effet d'espacement, nous aborderons également la notion d'*effet d'intervalle* (ou lag effect) qui s'attache à décrire, toujours dans le cas où un item est présenté à deux reprises, la performance mnésique en fonction de la durée de l'intervalle inter-répétition. Les études ont montré globalement qu'il existe un intervalle inter-répétition optimal. En d'autres termes, la performance mnésique décrit une courbe en U inversé en fonction de l'intervalle inter-répétition (IIR) et ce, que les répétitions aient lieu sur une courte (e.g., Glenberg, 1976) ou une longue (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008) échelle de temps. Nous aborderons ensuite les situations dans lesquelles trois occurrences d'un item (ou plus) surviennent. La situation se complexifie alors puisque plusieurs paramètres peuvent varier et se combiner entre eux : le nombre d'occurrences, la durée entre la première et la dernière occurrence, la durée des IIR successifs, et le délai de rétention (DR). D'une façon générale, les études maintiennent constant le délai de rétention et comparent l'effet de trois types d'*agencement temporel* des répétitions sur la mémoire : a) un agencement uniforme dans lequel les IIR successifs sont de même durée (par exemple X-X-X-X où les 'X' représentent les jours avec présentation

de l'item et les tirets les jours sans présentation) ; b) un agencement expansif dans lequel les IIR sont initialement courts puis de durée croissante (par exemple XX-X--X) ; et c) un agencement contractant dans lequel les IIR sont initialement longs puis de durée décroissante (par exemple X--X-XX). Dans les études réalisées sur une courte échelle de temps, c'est-à-dire lorsque apprentissage et récupération se déroulent au sein d'une même journée, il a été montré que l'effet varie selon le DR. L'agencement expansif est le plus avantageux lorsque le DR est court, tandis que l'agencement uniforme est le plus avantageux lorsque le DR est long, (e.g., Logan & Balota, 2008). Dans les rares études ayant été réalisées sur une longue échelle de temps, c'est-à-dire lorsque apprentissage et récupération se déroulent sur plusieurs jours (e.g., Tsai, 1927), c'est l'agencement expansif qui, globalement, était le plus avantageux pour la rétention ultérieure.

Le **deuxième chapitre** sera consacré à la description des théories évoquées dans la littérature de la psychologie cognitive pour expliquer les effets de pratique distribuée. Tout d'abord, la théorie de la *variabilité de l'encodage* (e.g., Glenberg, 1979) est une théorie relativement complexe qui stipule que le passage du temps entre les occurrences d'un item favorise un encodage varié de cet item lors des différentes occurrences. En effet, les éléments contextuels présents lors de l'encodage seraient inclus dans les traces mnésiques alors créées. D'une façon générale, un item répété qui donne lieu à un encodage varié a pour conséquence une trace mnésique plus riche, ce qui en favorise la récupération ultérieure. Il s'ensuit qu'un long intervalle entre les occurrences répétées est favorable au rappel final. En outre, le moment où se produit la récupération finale, qui dépend de la durée du délai de rétention, influence également la récupération et, donc, devrait moduler les effets de l'IIR. Deuxièmement, la théorie de la *récupération en phase d'étude* (e.g., Hintzman, Summers, & Block, 1975) propose que, pour qu'une répétition soit efficace, elle doit déclencher chez le sujet la récupération des traces mnésiques associées à la (ou aux) occurrence(s) antérieure(s) de cet item. Si cette récupération n'a pas lieu, la rétention ultérieure ne sera pas optimale. Bien que ces deux premières hypothèses soient classiquement opposées dans la littérature, quelques auteurs ont proposé de les combiner dans ce que nous appellerons les *modèles mixtes*. Par ailleurs, d'autres hypothèses ont été proposées, quoique étant moins souvent évoquées : les théories de *l'accessibilité et de la reconstruction* (e.g., Jacoby, 1978), et les théories du *traitement déficitaire* et de *l'amorçage* (e.g., Challis, 1993). Enfin, des hypothèses en termes de *métacognition* ont également été proposées récemment (e.g., Bahrick & Hall, 2005).

Le **troisième chapitre** abordera des conceptions issues d'autres champs disciplinaires des sciences cognitives pouvant rendre compte des effets de pratique distribuée. Nous décrirons tout d'abord les récentes études d'imagerie cérébrale qui ont porté sur les effets de pratique distribuée et nous constaterons que leurs résultats vont dans le sens des théories du *traitement déficitaire et de l'amorçage*. En effet, une répétition se produisant trop précocement serait associée à des effets d'amorçage comportementaux et neuraux qui empêcheraient un ré-encodage optimal de l'information (e.g., Wagner, Maril, & Schacter, 2000). Nous décrirons ensuite les principales tentatives de modélisation mathématique réalisées pour rendre compte des effets de pratique distribuée. Nous verrons que ces modèles, souvent, implémentent une combinaison entre plusieurs des hypothèses existantes, et en particulier celles de la variabilité de l'encodage et de la récupération en phase d'étude, ce qui s'apparente aux modèles dits *mixtes*. Nous verrons également qu'un de ces modèles (Mozer, Pashler, Cepeda, Linsdsey, & Vul, 2009) émet une prédiction relative à l'effet de l'agencement temporel à long terme : au fur et à mesure que le DR augmente, l'agencement expansif verrait son avantage sur les autres agencements augmenter. Dans la suite du troisième chapitre, nous évoquerons des concepts issus des neurosciences, en particulier les notions de

consolidation (e.g., Frankland & Bontempi, 2005), de taux d'oubli (e.g., Litman & Davachi, 2008), de reconsolidation (e.g., Hupbach, Gomez, Hardt, & Nadel, 2007), ainsi que l'influence du sommeil sur les effets de répétition (e.g., Ellenbogen, Hulbert, Jiang, & Stickgold, 2009). Les phénomènes de consolidation et d'oubli sont intéressants dans la mesure où ils peuvent ouvrir une nouvelle perspective sur les effets de pratique distribuée : Alors que l'étude des effets de pratique distribuée a relativement négligé la notion de taux d'oubli, l'interaction entre l'IIR et le DR pourrait s'expliquer par des taux d'oubli différents. Nous terminerons par des considérations de nature évolutionniste sur ces effets (e.g., Anderson & Schooler, 1991).

Le **quatrième chapitre** sera consacré aux Expériences 1 à 3, qui avaient pour objectif d'étudier l'effet de l'agencement temporel sur une longue échelle de temps. Les Expériences 1 et 2 comparaient des agencements uniforme (X-X-X-X), expansif (XXX---X) et contractant (X---XXX) comportant quatre occurrences d'un même stimulus réparties sur une phase d'apprentissage de sept jours. Les stimuli répétés et à apprendre étaient un ensemble de paires associant un pseudomot et un mot. La rétention de ces paires était mesurée lors du 9ème jour par un ensemble de tâches de mémoire explicite, dont une tâche de rappel indicé. L'Expérience 1 était réalisée par les participants par le biais d'Internet, tandis que l'Expérience 2 constituait une réplication de l'Expérience 1 en laboratoire. L'Expérience 1 a révélé que l'agencement expansif était associé à un rappel indicé plus élevé que les agencements uniforme et contractant. L'Expérience 2 a révélé que les agencements expansif et uniforme étaient associés à un rappel indicé plus élevé que l'agencement contractant. Nous discuterons des raisons de cette différence de résultats. Ces deux études répliquaient donc globalement les observations antérieures (e.g., Gay, 1973 ; Tsai, 1927). De plus, dans l'Expérience 2, une tâche de reconnaissance réalisée pendant la présentation des items a mis en évidence que les paires de l'agencement expansif étaient davantage reconnues lors de leur seconde présentation (i.e., P2) que celles des autres agencements. Or, tous agencements confondus, les paires reconnues lors de P2 étaient davantage rappelées ultérieurement, ce qui va dans le sens de la théorie de la récupération en phase d'étude. En effet, les items reconnus à P2, c'est-à-dire récupérés pendant l'apprentissage, bénéficieraient de cette récupération dans la mesure où leurs traces mnésiques seraient enrichies, ce qui favoriserait à son tour leur rétention finale.

L'Expérience 3 comparait les trois types d'agencements suivants : uniforme (X----X----X), expansif (XX-----X), et contractant (X-----XX), comportant 3 occurrences d'un même stimulus sur une période d'apprentissage de 13 jours. La durée du DR était manipulée, puisque certains participants étaient testés à J15, d'autres à J19, et d'autres encore à J26. Les résultats n'ont pas révélé d'effet de l'agencement à J15, un effet tendanciel à J19, et un effet significatif à J26, avec supériorité de l'agencement expansif par rapport aux autres agencements. Ainsi, la durée du délai de rétention modulait l'effet de l'agencement temporel. Ces résultats suggèrent également que l'agencement expansif est associé à un taux d'oubli plus faible que les autres agencements, c'est-à-dire que les souvenirs acquis par l'agencement expansif seraient plus résistants à l'oubli que les autres souvenirs.

Le **cinquième chapitre** sera consacré à l'Expérience 4 et plus généralement à la théorie de la variabilité de l'encodage. L'Expérience 4 avait pour objectif de tester cette théorie en utilisant un protocole consistant à présenter les items à deux reprises. Les mêmes items et les mêmes tâches mnésiques que dans les expériences précédentes étaient utilisés. La théorie de la variabilité de l'encodage prédit que si les deux occurrences ont lieu dans des contextes différents, alors la rétention finale est favorisée. Nous avons tenté d'induire une fluctuation plus ou moins importante des contextes internes dans deux groupes de sujets : le groupe dit de *Fluctuation Faible* réalisait une

activité monotone pendant l'IIR, tandis que le groupe dit de *Fluctuation Forte* réalisait un ensemble d'activités variées pendant l'IIR. La théorie de la variabilité de l'encodage prédisait que les sujets du groupe *Fluctuation Forte* seraient dans un état très différent à P2 par rapport à P1, tandis que les sujets du groupe *Fluctuation Faible* seraient dans un état à P2 relativement similaire à celui de P1. Par conséquent le groupe *Fluctuation Forte* devrait présenter des performances finales plus élevées que le groupe *Fluctuation Faible*. Les résultats n'ont pas révélé de différence significative, bien que le sens de la différence entre les moyennes pour les deux groupes aille dans le sens de la théorie.

La **dernière partie** sera consacrée à la discussion générale. Nos expériences ont montré, dans la continuité de Tsai (1927) et de Gay (1973), que l'agencement expansif était le plus à même de favoriser la mémorisation et la rétention ultérieures. Elles constituent donc des avancées empiriques pour cette problématique relativement peu explorée. Le résultat réellement nouveau est cependant d'avoir montré que l'avantage de l'agencement expansif sur les autres était de plus en plus important avec le passage du temps après la phase d'apprentissage. Ce résultat suggère que l'agencement expansif permettrait aux items de bénéficier d'un taux d'oubli plus faible que les autres.

Nous discutons tout d'abord de l'apport théorique de nos résultats. La théorie de la variabilité de l'encodage s'accommode relativement mal de ces résultats dans la mesure où elle prédirait que le meilleur agencement devrait être l'agencement contractant après un court DR et un agencement uniforme après un long DR. La théorie de la récupération en phase d'étude s'accommode relativement mieux de ces résultats, sans avoir toutefois prédit la modulation par le délai de rétention. Un modèle mixte combinant les propriétés des deux théories pourrait être cohérent. Il faut en effet noter que nos résultats concordent avec les prédictions du modèle mathématique de Mozer et al. (2009) qui implémente une combinaison des théories de la variabilité de l'encodage et de la récupération en phase d'étude.

Les considérations sur le taux d'oubli nous ont amenée à considérer les effets du délai de rétention dans les études antérieures (Tsai, 1927 ; Cull, 2000 ; ainsi que dans notre Expérience 2 pour les données de la rétention deux mois après l'apprentissage). Ces observations ont révélé que, dans ces études, l'agencement expansif était soit responsable des meilleures performances mnésiques ultérieures, soit associé au taux d'oubli le plus faible, soit les deux. Nous défendons ensuite l'idée que, dans les études sur l'effet de la pratique distribuée, il pourrait être fructueux d'adopter une nouvelle perspective vis-à-vis des facteurs IIR et DR. En effet, en général les auteurs considèrent la courbe de performance en fonction de l'IIR, et ce, en prenant en compte séparément chaque condition de DR. Ces courbes représentent donc en général une courbe en U inversé, et l'IIR optimal est de plus en plus long lorsque le DR augmente. Ce type de données est relativement difficile à interpréter. Or, une façon plus intuitive de représenter et d'interpréter les données serait de présenter les performances en fonction du DR, en considérant chaque IIR séparément, autrement dit de représenter la courbe d'oubli induite par les différents IIR. Nous poursuivons la discussion sur la question de la généralisabilité des résultats, sur les applications possibles des effets de l'agencement pour l'apprentissage et sur la poursuite des études à conduire sur ce sujet.

Première partie : Théorie

Chapitre 1 : Les effets de pratique distribuée dans la littérature

Defining the conditions that determine the effectiveness of a repetition is a major concern for the psychology of learning and memory. Glenberg (1976)

La question de l'effet de la répétition d'une information sur la mémoire est une question que se sont posée nombre d'éducateurs et de philosophes, soucieux pour certains d'améliorer leur façon d'enseigner par le biais de la répétition, et pour d'autres de comprendre les mécanismes du fonctionnement de la mémoire.

Nous nous attacherons dans ce chapitre à décrire les différents effets observés depuis plus d'un siècle relatifs à l'effet de la répétition et de l'intervalle temporel entre les répétitions. Les chapitres ultérieurs seront consacrés à la description et à l'évaluation des différentes théories visant à les expliquer.

1. 1. Les travaux pionniers d'Ebbinghaus sur la répétition

La question de la répétition est une des premières questions abordées par la psychologie expérimentale puisque l'instigateur de celle-ci, Hermann Ebbinghaus, mena une série d'expériences visant à mettre au jour les lois régissant l'effet des répétitions sur la rétention des informations. En 1885, Ebbinghaus partit du postulat qu'il existe un nombre de répétitions nécessaire à l'apprentissage d'une série de syllabes sans signification, et s'interrogeait sur ce qui se passerait si le nombre de répétitions ayant réellement lieu était en-deça de ce seuil ou, au contraire, au-delà.

Comme la « profondeur d'impression » du souvenir -notion que nous exprimons aujourd'hui par celle de *force* de la trace mnésique- n'est pas directement mesurable selon l'auteur, celui-ci la définit expérimentalement par la facilité avec laquelle le souvenir est reproduit à un moment donné après son apprentissage. Il mesurait cette facilité par la quantité de travail, ou d'effort, économisée lors du réapprentissage d'une série, en comparaison avec le travail normal à réaliser pour apprendre une toute nouvelle série. Ainsi, l'économie au réapprentissage est censée refléter la force du souvenir. Ebbinghaus choisit arbitrairement 24 h comme la durée entre l'apprentissage initial et le moment du réapprentissage où l'on évalue la rétention du matériel. Il découvrit que la rétention était proportionnelle au nombre de répétitions ayant lieu lors de l'apprentissage initial. Ainsi, plus l'information est répétée, plus elle donne lieu à un souvenir fort. De façon intéressante, Ebbinghaus a observé, en comparant les résultats de plusieurs de ses expériences, qu'il était plus avantageux pour le réapprentissage de réaliser l'ensemble des apprentissages sur des jours différents plutôt que de réaliser l'ensemble des apprentissages en une seule fois.

1. 2. L'effet d'espacement

De nombreuses études réalisées depuis celles d'Ebbinghaus ont confirmé la loi générale selon laquelle la répétition favorise la mémorisation (e.g., Peterson, Ellis, Toohill, & Kloess, 1935), ainsi que l'avantage de la répartition dans le temps des apprentissages. De façon empirique, les pédagogues avaient déjà eu l'intuition de ce phénomène. Il a été montré qu'un stimulus présenté deux fois était mieux retenu lorsque les deux présentations étaient séparées par un intervalle temporel, contenant ou non d'autres stimuli, par rapport à une situation où les deux présentations avaient lieu immédiatement à la suite l'une de l'autre (e.g., Melton 1967). La première situation correspond à la condition dite « espacée » ou « distribuée », et la dernière à la condition dite « massée ». Cette constatation a été intitulée l'effet d'espacement. Dans les études classiques portant sur cet effet, la procédure se schématise de la façon suivante, et nous utiliserons cette formalisation dans le reste de notre exposé : la première présentation d'un item, dite P1, est suivie de la seconde présentation de ce même item, dite P2. L'intervalle séparant P1 et P2 est variable et est appelé intervalle inter-répétition (IIR). Ultérieurement a lieu la tâche de rétention, et l'intervalle séparant P2 de la tâche de rétention est appelé intervalle de rétention ou délai de rétention dans le présent travail (DR). La Figure 1 représente cette succession d'événements.

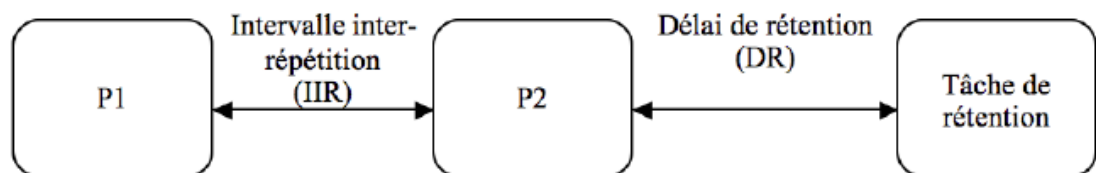


Figure 1 : Schématisation des étapes d'une expérience étudiant l'effet de l'intervalle inter-répétition sur les performances mnésiques.

L'effet d'espacement est devenu un thème de recherche sur la mémoire faisant l'objet de très nombreuses publications dans les années 1960-70 (avec une diminution de l'intérêt depuis les années 1970). En effet, l'effet d'espacement a été observé dans un très grand nombre de situations expérimentales mettant en jeu du matériel varié (syllabes sans signification, mots, phrases, paires de mots, mais aussi images, etc), diverses tâches de mémoire (de la reconnaissance au rappel libre en passant par le rappel indicé et l'estimation de fréquence), des modalités sensorielles variées (visuelle, auditive, et aussi visuelle et auditive ensemble), dans des conditions diverses d'apprentissage (incidentes et délibérées), et également dans des tâches d'apprentissage procédural (pour revue, voir Lee et Genovese, 1988). Plusieurs revues et méta-analyses de cet effet sont disponibles : Hintzman (1974), Cepeda, Pashler, Vul, Wixted, et Rohrer (2006), Crowder (1976, chapitre 9), Janiszewski, Noel, et Sawyer (2003), et Donovan et Radosovich (1999). Elles rapportent toutes le caractère robuste de l'effet d'espacement.

La grande majorité des études a été réalisée dans un cadre dans lequel les présentations des stimuli et le test de mémoire avaient lieu au cours d'une seule session expérimentale, par conséquent sur une fenêtre temporelle de quelques minutes, voire une heure au plus. En général, on présentait au sujet une liste d'items dans laquelle un item donné était présenté deux fois avec un intervalle variable entre les deux occurrences (IIR), puis quelques instants plus tard, une tâche de rétention avait lieu. D'autre part, quelques études ont été menées pour comparer une situation massée à une situation distribuée dans laquelle l'intervalle entre deux répétitions était d'au moins un jour (e.g., Rohrer & Taylor, 2006 ; Landauer & Ross, 1977 ; Bloom & Shuell, 1981 ; Pashler, Zarow, & Triplett, 2003 ; Seabrook, Brown, & Solity, 2005 ; Robinson, 1921). Les résultats ont conforté la loi selon laquelle une condition massée était délétère pour la mémoire ultérieure par rapport

à la condition distribuée. La seule exception à cette règle, à notre connaissance, était l'observation faite par Rohrer et Taylor (2006). Ils ont en effet observé que la condition espacée, correspondant à un intervalle d'une semaine, n'était pas supérieure à la condition massée lorsque les connaissances étaient testées après une semaine, alors qu'elle lui était nettement supérieure lorsque le test avait lieu après quatre semaines. Cet exemple nous permet de noter dès à présent l'importance du délai de rétention sur les effets observés, que nous détaillerons plus tard dans l'exposé.

En somme, l'effet d'espacement, même étudié sur une longue échelle de temps, est un effet robuste. Nous ne nous attarderons pas davantage sur ce sujet et recommandons au lecteur désireux d'en savoir plus de se référer aux revues existantes, en particulier Cepeda et al. (2006). Par ailleurs, des recherches récentes chez l'animal (présentées dans la section 3.3.1) ont apporté des éléments relatifs aux corrélats neurophysiologiques élémentaires impliqués dans l'effet d'espacement.

1. 3. L'effet d'intervalle

L'avantage de la répétition distribuée sur la répétition massée est donc un phénomène robuste, mais on peut s'interroger sur la différence d'efficacité entre deux IIR non nuls. En effet, au sein de la condition espacée, une infinité d'intervalles sont possibles, s'étendant de deux secondes à plusieurs années, et tous ne sont probablement pas aussi bénéfiques vis-à-vis de la mémoire. Cette question -intitulée effet d'intervalle (lag effect)- peut concerner les situations dans lesquelles l'étude a lieu sur une courte échelle de temps (i.e., une session expérimentale unique) ou sur une longue échelle de temps (i.e., plusieurs sessions expérimentales).

1. 3. 1. À court terme

Dans cette partie nous abordons l'effet d'intervalle sur le court terme, c'est-à-dire lorsque les répétitions sont réalisées au sein d'une liste. Un des premiers à avoir étudié cette question fut Melton en 1970, qui postulait que plus l'IIR était grand, meilleure était la rétention ultérieure. Autrement dit, la rétention finale était le résultat d'une fonction monotone croissante de l'IIR. Rapidement, cette hypothèse a été contredite expérimentalement. Par exemple, Glenberg (1976) a observé que la performance de rétention en fonction de l'IIR décrivait une fonction parfois monotone croissante et parfois non-monotone, c'est-à-dire décrivant une courbe en U inversé. La forme de la fonction différait selon plusieurs paramètres comme la durée du délai de rétention et la tâche mnésique utilisée pour mesurer la rétention.

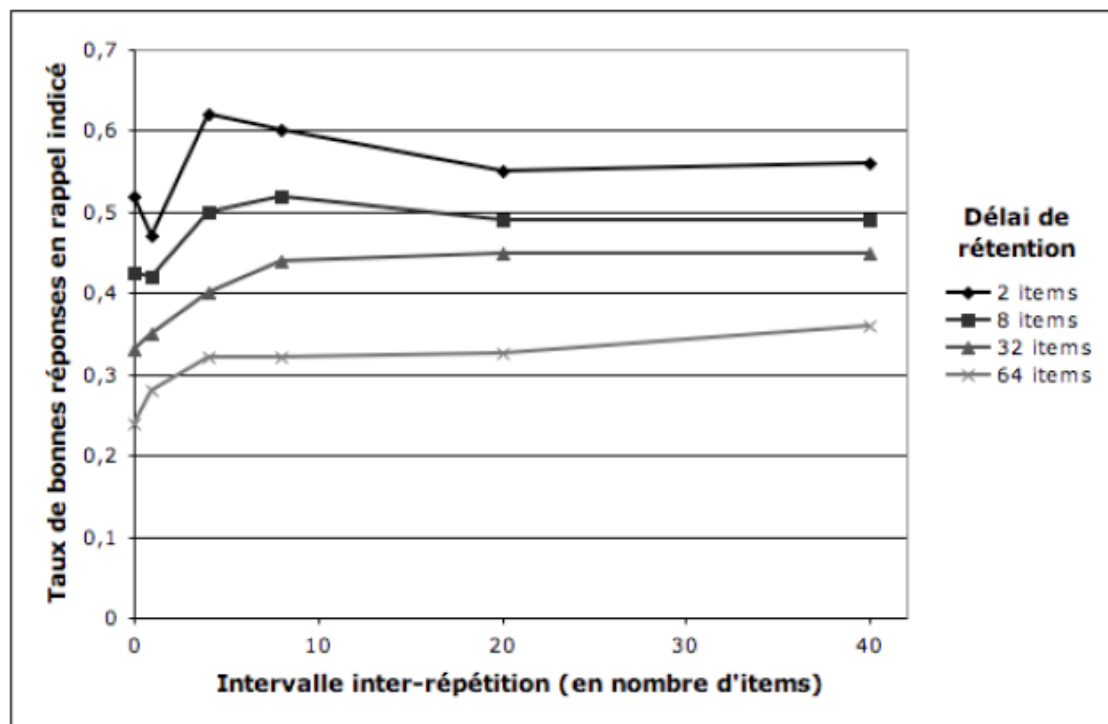


Figure 2 : Reproduction des résultats de l'étude de Glenberg (1976, Expérience 1). Taux de bonnes réponses au rappel indicé en fonction de l'IIR et du DR, tous deux mesurés en nombre d'items.

La Figure 2 représente les résultats de Glenberg 1976 (Expérience 1) dans laquelle les participants devaient apprendre un ensemble de paires de mots, puis devaient répondre à une tâche de rappel indicé en fournissant le mot associé en réponse à la présentation d'un mot seul. Sur la figure est représenté en abscisse l'IIR, mesuré en nombre d'événements intercalés, et en ordonnées la performance à la tâche de rappel indicé. Les quatre courbes représentent la performance après quatre DR différents. Le DR était lui-même mesuré en nombre d'événements intercalés, puisque les essais de présentation des paires et les essais de rappel indicé étaient mélangés au sein de la session expérimentale. Glenberg a observé que la courbe était non monotone, c'est-à-dire décrivait une courbe en U inversé, suite à un délai de rétention court, comme ici 2 ou 8 événements. Et dans ce cas, l'IIR optimal était de 4 à 8 événements. À l'opposé, la courbe était monotone croissante suite à un délai de rétention long, comme ici 32 ou 64 événements. Ainsi, la durée de l'IIR faisait varier la forme de la courbe de performance. Les études ultérieures ont confirmé la variabilité de la forme de cette fonction : celle-ci est parfois monotone croissante (e.g., Hintzman, Summers, Eki, & Moore, 1975 ; Hintzman, et al., 1975), parfois en forme de U inversé (e.g., Verhoeven, Rikers, & Schmidt, 2005 ; Toppino & Bloom, 2002).

1. 3. 2. À long terme

L'effet d'intervalle est-il le même lorsque l'on élargit l'échelle de temps des études et que les répétitions se produisent lors de plusieurs sessions séparées par au moins un jour ? Seules trois études récentes, à notre connaissance, ont étudié cette question sur des échelles de plusieurs jours : Glenberg et Lehman (1980), Cepeda et al. (2009), et Cepeda et al. (2008). De plus, il faut ajouter à cette littérature récente un certain nombre d'études menées par les chercheurs en pédagogie au début du XXème siècle, et qui sont rarement citées : Ausubel

(1966), Bunch (1941), Burt et Dobell (1925), Lyon (1914), Peterson et al. (1935), Reynolds et Glaser (1964), Starch (1912), Strong (1916). Ces dernières études ne bénéficiaient pas de la méthodologie expérimentale ni des outils statistiques actuels, nous conduisant à les considérer comme un ensemble d'indications plutôt que comme des résultats fiables et définitifs. De plus, des différences de protocole par rapport aux études plus récentes rendent les conclusions difficiles (e.g., Strong ainsi que Burt et Dobell n'avaient pas égalisé les DR).

Glenberg et Lehman (1980, Expérience 2) présentaient aux participants des paires de mots à deux reprises, séparées par un IIR de 1 jour ou de 7 jours. Le DR durait ensuite 7 jours et, lors du rappel indicé, les participants devaient donner le mot associé au mot présenté. Les performances étaient plus élevées pour les items répétés suite à l'IIR de 1 jour que suite à l'IIR de 7 jours. La Figure 3 synthétise ces données, ainsi que celles décrites dans les paragraphes suivants.

Dans l'expérience de Cepeda et al. (2009, Expérience 1), les participants apprenaient des paires de vocabulaire swahili-anglais à l'aide d'un protocole informatisé consistant en un rappel indicé avec feedback (ce protocole était similaire à celui utilisé dans Cepeda et al., 2008, décrit plus loin). Les sessions d'apprentissage consistaient en des phases de présentation des paires suivies de phases de test sur ces paires. Pendant ces phases de test, les participants devaient rappeler le mot anglais en réponse au mot swahili, et, à la fin de l'essai, la bonne réponse était présentée (i.e., essais de type *Test-étude*, voir la section 1.4.1). Une paire bénéficiant d'une bonne réponse ne faisait plus l'objet d'un rappel indicé dans le reste de la session. Chaque paire était apprise lors des deux sessions d'apprentissage qui avaient lieu avec un IIR de 1, 2, 4, 7, ou 14 jours selon les sujets. Le délai de rétention était de 10 jours. Les résultats ont montré une courbe de performance en U inversé en fonction de l'IIR ; l'IIR optimum était d'un jour et la courbe était relativement plate pour des IIR plus grands que cet optimum, (i.e., sur la partie droite de la courbe, voir la Figure 3).

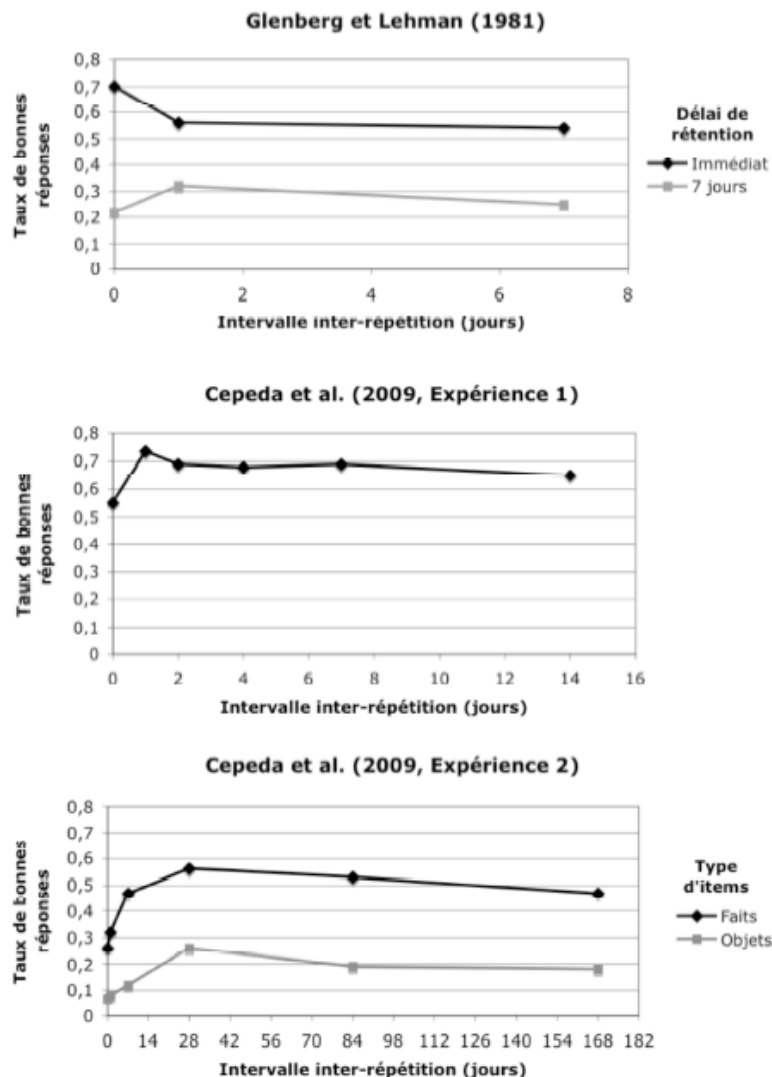


Figure 3 : Reproduction des résultats des études de Glenberg et Lehman (en haut), de Cepeda et al. (2009, Expérience 1 ; au milieu) et de Cepeda et al. (2009, Expérience 2 ; en bas). En abscisses est représenté l'IIR et en ordonnées le taux de bonnes réponses.

Enfin, dans la seconde expérience de Cepeda et al. (2009), on présentait aux participants à la fois des faits peu connus sous forme de question-réponse et des images d'objets peu connus avec le nom de l'objet associé, en suivant le même protocole informatisé que dans l'Expérience 1. Les deux types d'items étaient mélangés dans les sessions d'apprentissage. Chaque item était appris lors de deux sessions d'apprentissage qui avaient lieu avec un IIR de 0, 1, 7, 28, 84, ou 168 jours. Le délai de rétention était de 6 mois (i.e., environ 168 jours) avant le rappel indicé final. De nouveau, les résultats ont montré une courbe de performance en forme de U inversé en fonction de l'IIR, avec un optimum à 28 jours, et une courbe relativement plate à droite de l'optimum.

Ainsi, ces expériences ont mis en évidence, pour un délai de rétention donné, une fonction en U inversé, c'est-à-dire que l'intervalle optimal entre deux sessions semble être de durée moyenne, plus courte que le délai de rétention employé. De plus, cet intervalle optimal ne semble pas avoir une durée fixe.

Cepeda et al. (2008) ont alors fait le constat que dans les études déjà réalisées, un large spectre de durées d'IIR manquait au sein de l'ensemble des IIR possibles et que, de plus, peu d'études avaient étudié l'effet de la durée du DR et l'interaction entre DR et IIR. Ils ont réalisé une expérience à laquelle plus de 1300 sujets ont participé sur Internet. Afin d'observer un grand ensemble de combinaisons possibles entre ces deux intervalles, chaque participant a été réparti dans une condition combinant un IIR (0, 1, 2, 4, 7, 11, 14, 21, 35, 70, ou 105 jours) et un DR (7, 35, 70, ou 350 jours). Lors de la première session, la tâche consistait pour les sujets à répondre à des questions relatant des faits vrais mais peu pertinents, comme par exemple : « Quelle nation européenne consomme le plus de nourriture Mexicaine épicée ? ». La question était posée, puis la réponse était présentée (e.g., la Norvège) dans le cas où la réponse donnée n'était pas exacte. Tant que la bonne réponse n'était pas donnée par le sujet, l'item était programmé pour se présenter de nouveau ultérieurement. Dès que la bonne réponse était donnée, l'item n'était plus présenté pendant la session. Lors de la deuxième session qui avait lieu après l'IIR particulier à chaque sujet, chaque item qui avait été appris lors de la première session était maintenant testé avec feedback à deux reprises, et ce, quelle que soit la réponse du sujet. Lors de la troisième session, qui avait lieu suite à un DR particulier à chaque sujet, les participants devaient répondre aux questions comme dans les deux sessions précédentes mais sans feedback, puis répondaient également à un questionnaire à choix multiple pour les mêmes questions. Les courbes de performance au rappel indicé sont représentées dans la Figure 3. Ces courbes décrivent une courbe en U inversé dont le sommet, c'est-à-dire l'IIR optimal, augmente avec la durée du délai de rétention. Par exemple, pour un délai de rétention de 7 jours, l'IIR optimal était de 1 jour, tandis que pour un délai de rétention de 350 jours, l'IIR optimal était de 21 jours. Cette observation est importante car elle montre que l'IIR et le DR ont des effets joints à long terme tout comme à court terme, et qu'il est absolument essentiel de prendre en considération la durée du DR pour rendre compte des effets de la répétition sur la rétention ultérieure.

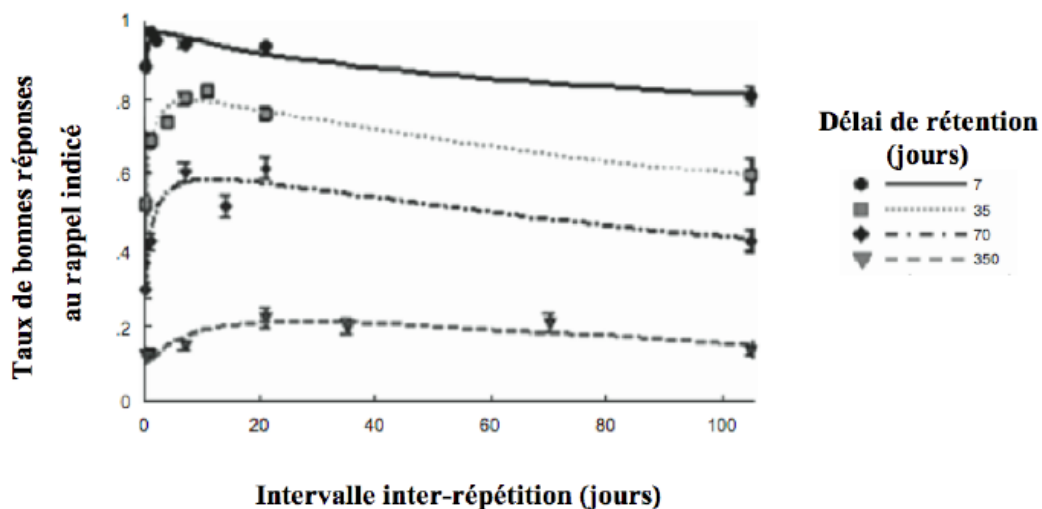


Figure 4 : Reproduction des résultats de l'étude de Cepeda et al. (2008). Taux de bonnes réponses dans la tâche de rappel indicé selon l'IIR (axe des abscisses) et le DR (couleur de la courbe). Les courbes ont été lissées par les auteurs.

Nous avons considéré jusqu'à présent les cas où une information était présentée à deux reprises. Cependant, dans l'environnement, une information peut se présenter un plus

grand nombre de fois. Or, comme nous l'avons vu, la rétention augmente en général avec le nombre de répétitions. Quel est l'effet des IIR lorsqu'une même information est répétée trois fois ou même davantage ? Un ensemble d'études à très long terme, c'est-à-dire sur des semaines, voire des années, a été conduit par Bahrick et ses collaborateurs (Bahrick, 1979 ; Bahrick & Phelps, 1987 ; Bahrick, Bahrick, Bahrick & Bahrick, 1993 ; Bahrick & Hall, 2005). De façon générale, ces études ont montré que pour un nombre de répétitions supérieur à deux, la rétention finale était la plus élevée pour les IIR les plus longs. Par exemple, dans Bahrick (1979, Expérience 2), les participants apprenaient des paires de vocabulaire espagnol-anglais à six reprises, séparées par des intervalles de 1 jour, 7 jours, ou 30 jours. Un participant était confronté à des intervalles ayant toujours la même durée. Après un DR de 30 jours, ce sont les participants avec l'IIR de 30 jours qui ont obtenu les meilleures performances.

Dans les études présentées jusqu'à présent, qu'elles aient comporté deux répétitions ou plus, la durée de la phase d'apprentissage, c'est-à-dire la durée de l'intervalle entre la première et la dernière occurrence de l'item variait selon les conditions. On peut alors s'interroger : pour une même durée de la phase d'apprentissage et pour un même nombre de répétitions des informations, la façon d'agencer temporellement les répétitions au sein de la phase d'apprentissage influence-t-elle la rétention ? Par le terme d'agencement, nous nous référons au rapport entre les durées des IIR successifs (P1-P2, P2-P3, P3-P4, etc). Est-il plus efficace de programmer des intervalles successifs tous de même durée, ou de durée progressivement croissante, ou encore de durée progressivement décroissante ? C'est l'objet de la section suivante et, plus largement, du présent travail.

1. 4. Effets de différents agencements temporels

L'effet de l'agencement temporel des répétitions a fait l'objet d'une dizaine d'articles, et a été étudié d'une part dans le cadre d'expériences sur des échelles de temps courte (jusqu'à plusieurs heures) ou longue (plusieurs jours). Nous examinerons ces deux échelles temporelles successivement, sachant que les études à court terme sont beaucoup plus nombreuses et récentes que les études à long terme. Dans une optique de clarté, certains éléments explicatifs (qui font normalement l'objet des chapitres 2 et 3) seront introduits dans cette section.

1. 4. 1. À court terme

L'étude princeps de l'effet de l'agencement temporel de plusieurs répétitions est celle de Landauer et Bjork (1978). Nous allons la décrire en détail pour permettre au lecteur de bien comprendre le principe des différents agencements temporels.

Études de Landauer et Bjork (1978) et de Cull, Shaughnessy, et Zeichmaster (1996)

Dans la première expérience de Landauer et Bjork (1978), les participants devaient apprendre l'association entre un prénom et un nom. La présentation simultanée du prénom et du nom avait lieu une seule fois au début de l'expérience. Ensuite, lors des trois essais suivants concernant cette association, seul le prénom était présenté et les participants devaient rappeler le nom manquant. Les auteurs ont comparé entre eux de nombreux agencements temporels schématisés de la façon suivante : 0-0-0, 1-1-1, 4-4-4, 5-5-5, 10-10-10, 0-3-10, 1-4-10, 10-3-0 et 10-4-1. Selon cette nomenclature, les trois chiffres successifs représentent les IIR successifs pour un item donné, et sont exprimés en nombre

d'autres items intercalés. Par exemple, la schématisation 0-3-10 signifie que la 2^{ème} présentation (e.g., John ____) avait lieu immédiatement après la 1^{ère} (e.g., John Smith), ce qui correspond au chiffre 0 ; la 3^{ème} présentation avait lieu après que trois autres items soient apparus après la 2^{ème} présentation, ce qui correspond au chiffre 3 ; enfin, la 4^{ème} présentation avait lieu après que 10 items soient apparus après la 3^{ème} présentation, ce qui correspond au nombre 10. Dans ce cas précis, l'agencement est dit expansif : d'une façon générale, les agencements expansifs consistent en des IIR successifs de plus en plus longs. Les agencements dits uniformes consistent en des IIR successifs de même durée, comme par exemple 4-4-4. Les agencements dits contractants consistent en des IIR successifs de plus en plus courts, comme par exemple 10-3-0.

Pendant la phase d'apprentissage, tous les items expérimentaux étaient présentés les uns à la suite des autres, chacun ayant son propre agencement, et chacun servant également à remplir les intervalles séparant les répétitions des autres items. Le test final consistait à donner le nom en réponse au prénom (e.g., John ____), et avait lieu 30 minutes après la 4^{ème} présentation. Les résultats ont montré que pour un IIR moyen identique, l'agencement expansif (comme par exemple 0-3-10) permettait une performance significativement supérieure à celle de l'agencement uniforme (comme 4-4-4) qui, lui-même, n'était pas différent de la condition contractante (comme 10-3-0).

L'Expérience 2 comparait les deux agencements temporels suivants : 3-3-3-3 (uniforme) et 0-1-3-8 (expansif), qui consistaient en cinq essais par item. La première présentation de l'item comportait 3 éléments, c'est-à-dire un visage associé à un prénom et à un nom. Les quatre essais suivants concernant cet item consistaient en la présentation du visage et du prénom, et le sujet devait alors donner le nom associé. Trente minutes après, le test final consistait pour le sujet à rappeler le nom et le prénom en réponse à la présentation du visage. Les résultats ont montré que le bénéfice relatif de l'agencement dépendait du type de présentation des items. Pour la partie de l'item qui était *présentée* à chaque répétition (en l'occurrence, le prénom), le rappel final montrait une tendance non significative en faveur de l'avantage de l'agencement uniforme sur l'agencement expansif. À l'inverse, pour la partie de l'item qui n'était présentée que lors de la première présentation et qui était *testée* lors des présentations suivantes (en l'occurrence, le nom de famille), l'agencement expansif entraînait de meilleures performances de rappel lors du test final, mais aussi tout au long de l'apprentissage, par rapport à la condition uniforme.

Cette expérience montrait donc qu'il est important de considérer la façon de présenter l'information (i.e., présentation ou test) pour pouvoir prédire l'effet de l'agencement temporel des répétitions. Les auteurs proposaient que dans le cas de tentatives de récupération sans feedback, l'agencement expansif est plus efficace car il permet de maintenir élevée la probabilité de succès de récupération tout au long de la phase d'apprentissage. En effet, le fait d'avoir de courts IIR au départ favorise la récupération correcte, qui est toujours favorisée ensuite grâce aux IIR de durée croissante.

En 1996, Cull et al. (Expériences 1 à 5) ont répliqué les résultats de Landauer et Bjork (1978), en observant que, pour différents types de matériel, une succession de trois tests sans feedback agencés de façon expansive (i.e., 1-5-9) entraînait, lors d'un test de mémoire ultérieur, des performances plus élevées que l'agencement uniforme (i.e., 5-5-5).

Cependant, des études ultérieures ont remis en question l'apparente généralité de ces observations.

Résultats contradictoires

Cull (2000, Expérience 1) a utilisé dans ses expériences trois façons différentes de présenter l'information au cours de la phase d'apprentissage. Le premier type de présentation est appelé *Etude seule* : les deux éléments constitutifs de l'item apparaissent toujours l'un avec l'autre. Par exemple, dans le cas où les items sont des paires de mot, l'intégralité de la paire apparaît lors de chaque présentation. Le second type de présentation possible est appelé *Test seul* : lors du premier essai, les deux éléments de la paire sont présentés ensemble, puis lors des essais ultérieurs, un seul élément constitutif de la paire est présenté, et le participant doit alors rappeler l'élément associé. Dans cette condition, aucun feedback n'est fourni suite à la réponse du participant. C'est ce type de présentation qui était employé dans l'Expérience 2 de Landauer et Bjork (1978) avec les éléments visage et prénom présentés à chaque essai, la réponse attendue étant le nom de famille. Enfin, le troisième type de présentation est dit *Test-étude* : l'élément indice apparaît seul pendant quelques secondes, le participant doit tenter de rappeler l'élément cible, puis celui-ci apparaît finalement pendant le reste du temps de présentation, et fait office de feedback. Une variante de ce type de présentation correspond à une situation dans laquelle le participant reçoit un feedback dès qu'il a donné sa réponse, et non pas après un intervalle fixé. Nous ne ferons pas de distinction entre ces deux variantes. (Le type *Test-étude* peut également être intégré dans un protocole d'apprentissage, comme dans Cepeda et al., 2008, 2009, dans lequel l'item est présenté jusqu'à obtenir une réponse correcte de la part du participant.) Nous adopterons dans la suite du texte la nomenclature utilisée par Cull (2000) pour décrire les différents types de présentation.

Dans l'Expérience 1 de Cull (2000), les sujets devaient apprendre des paires associant un mot rare et un mot commun. Les agencements expansifs (1-5-9) et uniforme (5-5-5) ont été comparés, mais contrairement aux résultats antérieurs, pour aucun type de présentation la différence entre les performances des agencements expansif et uniforme n'était significative.

Dans sa deuxième expérience, Cull (2000) a comparé des agencements uniforme et expansif avec des IIR plus longs, c'est-à-dire de l'ordre de quelques minutes et non plus en termes d'items intercalés. De nouveau, l'agencement n'avait pas d'effet sur la performance, et ce, quel que soit le type de présentation utilisé.

Carpenter et DeLosh (2005, Expérience 2) ont comparé un agencement uniforme (3-3-3) avec un agencement expansif (1-3-5), avec pour stimuli des paires associant un visage et un nom. Par ailleurs, ils ont comparé des essais de type *Etude seule* avec des essais de type *Test seul*. Contrairement aux études antérieures, l'agencement uniforme s'est révélé plus efficace pour la rétention que l'agencement expansif, et ce pour les deux types de présentation. Dans l'Expérience 3, ils ont comparé un agencement uniforme (3-3-3) avec un agencement expansif (3-5-7) tel que, cette fois, le premier IIR était de même durée et l'IIR moyen n'était pas le même dans les deux conditions. L'objectif était de tester l'hypothèse selon laquelle l'agencement expansif est plus efficace car il permet l'augmentation progressive de la difficulté de récupération au fil des répétitions (Bjork, 1988). Ainsi, dans le cas des agencements utilisés, le critère de difficulté progressive était respecté dans l'agencement expansif mais pas dans l'agencement uniforme, ce qui aurait dû conduire à une supériorité de l'agencement expansif. Or, les résultats n'ont pas révélé de différence entre les deux agencements, et ce dans aucun des deux types de présentation, remettant en question à la fois l'hypothèse de l'avantage général de l'agencement expansif et celle de la théorie de difficulté croissante de récupération.

Trois articles récents (Balota, Duchek, Sergent-Marshall, & Roediger, 2006 ; Karpicke & Roediger, 2007 ; Logan & Balota, 2008) se sont intéressés à la comparaison des agencements expansif et uniforme dans des études à court terme. Tous trois testaient l'hypothèse selon laquelle c'est le premier IIR, celui entre l'essai de présentation (P1) et le premier essai test (P2), qui détermine la rétention et la performance ultérieure de mémoire. Les trois études utilisaient des agencements similaires aux agencements déjà rencontrés, par exemple 3-3-3 vs. 1-3-5, et employaient la procédure *Test seul* et parfois *Test-étude*. Le matériel à apprendre était un ensemble de paires de mots ou de paires de vocabulaire. Balota et al. étudiaient un groupe de patients souffrant de la maladie d'Alzheimer, et Logan et Balota un groupe de sujets âgés normaux. Pour pouvoir étudier plus avant les processus en jeu pendant la phase d'apprentissage, les scores aux tests étaient enregistrés, ainsi que les temps de réponse dans les études de Karpicke et Roediger et de Logan et Balota. De plus, dans ces deux dernières études a été ajoutée, en plus du DR de quelques minutes classiquement employé, une condition dans laquelle le rappel final était plus tardif : 24h pour Logan et Balota et 48h pour Karpicke et Roediger. Les résultats ont montré les effets suivants. Premièrement, il y avait un avantage général et toujours observé de l'agencement expansif sur l'agencement uniforme pendant la phase d'apprentissage en termes de scores de récupération en condition *Test seul*. Ce n'était pas toujours le cas en condition *Test-étude*. Cependant, la différence de succès de récupération entre les agencements expansif et uniforme était de plus en plus faible au fil des essais de récupération. Autrement dit, lors de la première tentative de récupération, le succès pour l'agencement expansif était très important, et il diminuait d'essai en essai, rejoignant alors les niveaux de succès de l'agencement uniforme. Il y avait donc une convergence des performances. Ce phénomène peut s'expliquer par le fait que les IIR étaient de plus en plus longs dans la condition expansive, favorisant de plus en plus l'oubli des informations entre les répétitions. Deuxièmement, les temps de réponse diminuaient progressivement d'essai en essai dans les deux agencements, mais cette diminution était plus ample dans l'agencement uniforme, qui par ailleurs démarrait d'une valeur initiale plus haute. Comme les auteurs considéraient que les temps de réponse reflètent la difficulté de récupération, alors les essais seraient de moins en moins difficiles dans les deux agencements. Or, ce résultat s'oppose à la théorie énoncée par Bjork (1988) selon laquelle l'avantage de l'agencement expansif sur l'agencement uniforme serait dû au fait que la difficulté de récupération est croissante d'essai en essai dans la condition expansive, mais pas dans la condition uniforme. Troisièmement, les scores finaux lorsque le test se déroulait lors de la même session n'ont pas révélé de différence entre les agencements. Cependant, de façon intéressante, lorsque le test avait lieu après 24h ou 48h c'est l'agencement uniforme qui présentait un avantage significatif sur l'agencement expansif.

Ces résultats remettaient ainsi en question l'hypothèse selon laquelle la difficulté de récupération est croissante dans l'agencement expansif (Bjork, 1988) puisqu'au contraire les temps de récupération étaient de plus en plus courts pendant la phase d'apprentissage. Au contraire, il semble plutôt que le plus grand effort de récupération a lieu pendant le premier essai de récupération et que cette récupération nécessite plus d'effort dans la condition uniforme que dans la condition expansive. Cette hypothèse a aussi été confirmée par l'observation de Karpicke et Roediger (2007, Expérience 3) et de Carpenter et DeLosh (2005) selon laquelle, lorsqu'on compare des agencements qui commencent par le même intervalle, alors on n'observe pas de différence entre les agencements.

Ainsi, de façon globale, il s'avère qu'après un court DR, l'avantage de l'agencement expansif a été observé dans la moitié des études, l'autre moitié montrant soit une égalité soit un avantage de l'agencement uniforme. Au contraire, après un DR de 24h ou plus,

l'avantage de l'agencement uniforme a été observé dans toutes les études. Il semble que lorsque le test de rétention a lieu après un certain délai, c'est l'agencement de type uniforme qui serait le plus efficace. Les études ont également montré que pour favoriser la rétention à long terme, il est avantageux d'avoir un premier intervalle long avant le premier test. L'idée sous-jacente est qu'avec un important IIR entre la présentation initiale et le premier test, la difficulté de récupération est alors importante, et que si la récupération est réussie alors celle-ci est très utile pour la rétention ultérieure. Il semble par ailleurs que la première tentative de récupération est très importante tandis que les tentatives ultérieures le sont moins. Cela remet donc en cause l'idée généralement acceptée que l'agencement expansif est plus favorable que l'agencement uniforme. On peut considérer également que les agencements expansifs consistent en un intervalle initial très court, voire de type massé, qui est efficace pour la rétention à court terme, alors que les agencements uniformes consistent en un intervalle initial plus long, qui est efficace pour la rétention à long terme. Cette explication pourrait expliquer pourquoi l'agencement uniforme semble être de plus en plus efficace au fur et à mesure de l'allongement du DR.

Il faut cependant remarquer que les études ayant conduit à cette conclusion utilisaient des présentations de type *Test seul* ou *Test-étude*, qui incitent le sujet à récupérer explicitement les réponses pendant la phase d'apprentissage, et non celles de type *Etude seule*, qui est le type employé classiquement dans les études sur les effets d'espacement et d'intervalle. Il est possible que d'autres explications prévalent lorsque ce dernier type de présentations est employé. Globalement, on peut conclure que l'effet de l'agencement est différent selon le type de présentation utilisé.

Nous avons décidé de nous intéresser aux situations dans lesquelles l'information n'est pas testée de manière répétée mais est présentée de manière répétée. Nous pensons que certains mécanismes sont communs à ces deux situations mais que, cependant, des particularités sont associées à chaque situation. Il est important de noter que dans les quelques études ayant utilisé la procédure *Etude seule* (i.e., Landauer & Bjork, 1978, Expérience 2 ; Cull, 2000, Expériences 1 et 2 ; Carpenter & DeLosh, 2005), soit il n'y avait pas de différence observée entre les agencements uniforme et expansif, soit l'agencement uniforme était plus efficace. L'agencement expansif n'a jamais été significativement plus efficace que l'agencement uniforme dans le cas de présentations successives, c'est-à-dire en condition *Etude seule*.

1. 4. 2. À long terme

Nous avons considéré jusqu'à présent les études portant sur l'effet de l'agencement temporel sur une courte échelle de temps. Cependant, les informations présentes dans l'environnement et que nous devons retenir peuvent également se répéter sur plusieurs jours. D'autre part, ces informations qui se répètent (par exemple lorsque nous interagissons avec les autres, que nous lisons le journal, que nous apprenons quelque chose ou que nous révisons nos leçons) ne nous imposent pas de « test », elles se présentent simplement à nouveau, ce qui, expérimentalement, correspond à des essais de type *Etude seule*. La question de l'agencement des répétitions sur une longue échelle de temps, c'est-à-dire sur plusieurs jours, a été le sujet de recherches menées au début du siècle (e.g., Peterson, et al., 1935 ; Burt & Dobbie, 1925; Lyon, 1914)¹ mais fait l'objet de peu d'études récentes,

¹ De façon intéressante, voici ce qu'écrivait Lyon en 1914 : « [...] la loi la plus générale qui peut être énoncée [...] est que la méthode la plus économique est de distribuer les lectures sur une période relativement longue - les intervalles entre les lectures étant de proportion arithmétique. Par exemple, pour un individu qui mémorise un poème de vingt strophes, la meilleure rétention était obtenue en distribuant les lectures de la façon suivante : 2 heures, 8 heures, 1 jour, 2 jours, 4 jours, 8 jours, 16 jours, etc. [...] Les associations

probablement pour des raisons pratiques. Parmi ces études, quatre d'entre elles présentent un intérêt particulier pour notre problématique. Il s'agit, dans l'ordre chronologique, de Tsai (1927), Sones et Stroud (1940), Gay (1973), et Cull (2000).

Dans l'étude de Sones et Stroud (1940), les auteurs ont comparé l'efficacité de trois agencements temporels et, à la différence des études actuelles qui égalisent le DR, ont égalisé la durée totale entre la première occurrence et le test final (durée totale de 42 jours). Cela aboutit au fait que la durée d'apprentissage ainsi que la durée du DR différaient entre les agencements (la durée de la phase d'apprentissage était de 4, 16, ou 18 jours respectivement pour les agencements expansif, uniforme, et contractant). Le matériel utilisé était de plus particulier dans la mesure où il s'agissait de faire lire aux participants des passages de texte. Ainsi, le fait qu'ils aient observé un avantage de l'agencement expansif par rapport aux agencements contractant et uniforme est à prendre avec grande précaution.

Gay (1973, Expérience 2) a comparé l'effet de différents agencements de sessions d'apprentissage avec du matériel mathématique auprès d'élèves du collège. Le principe était le suivant. Les élèves devaient apprendre quatre règles mathématiques au total. Lors de la première session, l'apprentissage consistait à faire en sorte que l'élève parvienne à résoudre à la suite deux problèmes relatifs à la même règle, et ce pour les quatre règles. Les sessions de révision avaient lieu les jours 2 et 3 pour le premier groupe (i.e., agencement uniforme), les jours 2 et 8 pour le deuxième groupe (i.e., agencement expansif) et les jours 7 et 8 pour le troisième groupe (i.e., agencement contractant). Les sessions de révision comportaient un test sur les règles précédemment apprises, suivi de la présentation d'un nouvel exemple si l'élève échouait, puis le passage à une autre règle lorsque l'élève réussissait. Le test de rétention final avait lieu le jour 22 et consistait en un test similaire à ceux proposés lors des sessions d'apprentissage et de révision. Dans cette expérience, tout comme dans celle de Sones et Stroud (1940), l'intervalle entre la première session et le test de rétention final a été rendu égal entre les groupes. Ici, le DR n'était donc pas le même entre toutes les conditions. Cependant, pour les groupes expansifs et contractants, la dernière session de révision avait lieu le jour 8 ; donc le DR était le même dans les deux cas, et l'on pouvait donc les comparer. Les scores ont montré que le nombre de problèmes correctement résolus était plus important en condition expansive (sessions les jours 1, 2, 8) qu'en condition contractante (les jours 1, 7, 8), lui-même plus important qu'en condition uniforme (les jours 1, 2, 3). L'agencement expansif a donc permis une meilleure rétention finale que l'agencement contractant. Il faut cependant prendre en considération la particularité du matériel étudié et la tâche réalisée. En effet, il s'agissait d'appliquer une règle mathématique apprise à des exemples nouveaux, ce qui s'apparente davantage à de l'apprentissage procédural qu'à de la mémoire déclarative. Or, il n'est pas assuré que les phénomènes observés dans les tâches procédurales soient les mêmes que celles observées dans des tâches déclaratives.

Tsai (1927, Expériences 2 et 3) a été le premier à étudier l'effet de l'agencement temporel de sessions d'apprentissage sur plusieurs jours. Les stimuli utilisés étaient des mots de vocabulaire anglais - chinois présentés à des étudiants chinois. Les participants

qui ont été formées à un moment donné doivent être rappelées avant qu'un intervalle trop long se soit écoulé, lequel faisant que les associations originales perdent alors leurs « couleurs » et ne peuvent plus être rappelées dans les mêmes forme, temps, et ordre. En général, il a été trouvé que la méthode la plus économique pour conserver le matériel à l'abri de la disparition une fois mémorisé, était de revoir le matériel dès qu'il commençait à s'« effacer ». Cette citation reflète à la fois l'utilité de l'agencement expansif et la notion de *récupération en phase d'étude* comme nous le verrons dans le chapitre 2. Toutefois, les arguments étaient davantage ceux, intuitifs, de pédagogues, que basés sur des résultats expérimentaux.

ont été répartis en trois groupes et chaque groupe bénéficiait de sessions d'apprentissage correspondant à un agencement particulier.

Dans l'Expérience 2, trois agencements, distribués sur un total de 11 jours et consistant en 6 sessions d'apprentissage ont été comparés. L'agencement expansif consistait en les intervalles successifs suivants : 24h-24h-24h-72h-96h. Cet agencement peut également se représenter de la façon suivante : XXXX--X---X. Dans cette nomenclature, les 'X' représentent les jours où l'item était présenté, tandis que les tirets représentent les jours où l'item n'était pas présenté. Par souci de clarté, nous adopterons cette nomenclature pour représenter les agencements à long terme. L'agencement contractant était le suivant : X---X--XXXX, et l'agencement uniforme était le suivant : X-X-X-X-X-X. Ainsi, pour les trois agencements, l'espacement moyen entre deux répétitions successives était de 48h.

Le test de rappel indicé avait lieu le 14^{ème} jour et consistait à donner, en réponse aux mots anglais, le mot chinois associé. Le DR avait donc une durée de 3 jours. Un second test de même nature avait lieu le 18^{ème} jour pour les mêmes sujets (DR de 7 jours). Les performances de rappel se sont révélées plus élevées dans la condition expansive (61% et 45% respectivement pour le test à J14 et J18) que dans la condition contractante (52% et 41%), et les performances de rappel étaient plus élevées dans la condition contractante que dans la condition uniforme (48% et 35%). L'article ne précise pas la validité statistique de la comparaison de ces scores.

Dans l'Expérience 3, trois nouveaux agencements étaient comparés, dans un protocole comportant 8 sessions d'apprentissage distribuées sur un total de 15 jours. Les trois agencements avaient un IIR moyen d'une durée de 48h comme dans l'Expérience 1, et se schématisaient de la façon suivante:

- Agencement expansif : XXXX-X-X--X---X
- Agencement contractant : X---X--X-X-XXXX
- Agencement uniforme : X-X-X-X-X-X-X-X

Les tests avaient lieu le 18^{ème} jour (DR de 3 jours) et le 32^{ème} jour (DR de 17 jours). Les performances de rappel étaient plus élevées dans la condition expansive (74% et 54% respectivement pour le test à J18 et J32) que dans la condition contractante (64% et 48%), et elles étaient plus élevées dans la condition contractante que dans la condition uniforme (56% et 40%).

Ainsi, de façon générale, ces expériences ont montré que l'agencement de type expansif était plus efficace que les autres avec des répétitions consistant en de simples présentations sur plusieurs jours. Cependant, il est délicat de tirer des conclusions claires quant à l'interaction entre l'effet du DR et l'effet de l'agencement. En effet, bien que dans chacune de ces expériences deux durées de DR aient été testées, les deux tests portaient tous les deux sur l'ensemble du matériel chez les mêmes sujets. Les deux tests successifs n'étaient donc pas indépendants et il est possible que les réponses fournies par les participants lors du premier test aient permis une réactivation et un réapprentissage des items rappelés, alors que ce n'était pas le cas pour les items non rappelés.

Dans sa troisième expérience, Cull (2000) a comparé deux agencements temporels, pour lesquels l'ensemble des sessions étaient réparties sur 7 jours. Deux agencements temporels étaient comparés, et ce de façon inter-sujets : l'agencement expansif était de type XX-X--X, tandis que l'agencement uniforme était de type X-X-X-X. Les items appris étaient des paires de vocabulaire. La procédure d'apprentissage utilisée était similaire à

celle de Cepeda et al. (2008, 2009) puisque lors de la première session, on s'assurait que les sujets avaient répondu correctement à chaque item au moins une fois ; lorsqu'une réponse était incorrecte, la solution était présentée, et l'item testé de nouveau ultérieurement. Les performances finales mesurées le 10^{ème} jour (DR de 3 jours) n'ont pas révélé de différence entre les agencements. L'Expérience 4 répliquait l'Expérience 3 avec un DR de 8 jours et n'a pas non plus révélé de différence d'efficacité entre les agencements.

En somme, les rares études existantes ayant comparé différents agencements sur le long terme présentent des résultats relativement contradictoires, et de plus, les protocoles étaient différents dans chaque cas. Par ailleurs, la comparaison des études réalisées à court terme avec celles réalisées à long terme est délicate dans la mesure où les études à court terme employaient majoritairement une procédure d'apprentissage à base de tests alors que les études à long terme employaient majoritairement une procédure d'apprentissage à base de présentations simples. On ne sait donc pas si les mêmes effets de l'agencement à long terme seraient observés dans le cas d'essais de type *Test-étude* / *Test seul* ou *Etude-seul*.

L'objectif du présent travail était de mieux comprendre les effets mnésiques de l'agencement temporel sur une longue échelle de temps, en conduisant des études s'inspirant de celles réalisées par Gay (1973), Cull (2000) et surtout Tsai (1927). Les deux chapitres suivants seront consacrés à la description des causes potentielles de ces effets d'agencement et, plus généralement, des effets de pratique distribuée.

Chapitre 2 : Les théories psychologiques de l'effet de pratique distribuée

One venerable topic in experimental-cognitive psychology is how and why learning occurs . Roediger et Karpicke (2006)

Introduction

Nous avons présenté dans le chapitre 1 un certain nombre d'études ayant montré que le délai temporel avec lequel une information est répétée influence la rétention de cette information. Ces observations nous amènent à nous interroger sur le ou les mécanismes par lesquels la durée de l'intervalle inter-répétition (IIR) peut influencer l'apprentissage et la rétention. En quoi le passage du temps entre deux (ou davantage) occurrences d'un même item peut-il avoir un effet qui s'observera ultérieurement lors de la récupération ? Nous nous intéresserons principalement aux situations où seules deux occurrences de l'information ont lieu, pour des raisons de simplicité et parce que peu de théories ont été avancées pour expliquer l'effet de trois répétitions ou plus. Nous appellerons P1 et P2 respectivement la première et la seconde occurrence d'une information.

Avant d'exposer les différentes théories en présence, qui sont toujours actuellement débattues, plusieurs points doivent être évoqués. Premièrement, les auteurs n'ont pas toujours distingué l'effet d'espacement et l'effet d'intervalle dans l'élaboration de leurs théories, ce qui ne rend pas toujours claire la limite entre les théories s'appliquant à l'un et celles s'appliquant à l'autre. Deuxièmement, les théories ont été majoritairement élaborées dans le cadre de l'effet de la pratique distribuée à court terme, et il n'est pas évident que les théories soient également valables sur le long terme. Troisièmement, il n'est pas non plus

évident, ni toujours discuté dans la littérature, que les théories soient valables à la fois dans les cas où l'apprentissage se produit par le biais de présentations répétées d'un item (essais de type *Etude seule*, voir la section 1.4.1) et dans les cas où l'apprentissage se produit par le biais de tests successifs sur le même item (essais de type *Test seul* et *Test-étude*). Nous nous concentrerons essentiellement sur les théories relatives aux essais de type *Etude seule* dans la mesure où ce type de présentations sera utilisé dans nos expériences.

La logique générale de l'explication de l'effet de pratique distribuée est la suivante. On peut raisonnablement penser que les processus en jeu lors de P1, c'est-à-dire lorsque l'on voit l'information la première fois (du moins la première fois dans le contexte expérimental particulier) sont les mêmes quel que soit ce qui se passe après. La différence entre les différentes conditions est donc due :

- soit à des phénomènes qui se produisent lors de l'intervalle entre P1 et P2,
- soit au fait que, lors de P2, des phénomènes diffèrent entre les différentes conditions,
- soit encore une combinaison des deux phénomènes précédents.

De plus, les phénomènes en question peuvent avoir des conséquences non seulement pendant l'apprentissage, mais également sur la façon dont ce qui a été appris va persister dans le temps après P2, c'est-à-dire lors du délai de rétention (DR).

Nous allons dans ce chapitre présenter les théories développées dans la littérature psychologique pour expliquer les résultats observés, tandis que le chapitre 3 sera consacré aux théories existantes dans le champ plus vaste de la neuroscience cognitive. Au fur et à mesure des nombreuses études réalisées sur la thématique de la pratique distribuée, de nombreuses théories psychologiques ont vu le jour. Les explications les plus intuitives qui ont été proposées sont les suivantes :

Premièrement, deux hypothèses faisant reposer l'origine de l'effet sur l'intervalle P1-P2 ont été proposées : celle de la répétition mentale (*rehearsal*) et celle de la consolidation. L'hypothèse de la répétition mentale stipule que le fait d'avoir un long IIR entre P1 et P2 pourrait permettre au sujet d'avoir plus de temps pendant cet intervalle pour 'réviser' les items à apprendre, par rapport à un court IIR. En effet, on sait que lorsqu'un sujet a pour tâche d'apprendre une liste d'items, il peut lui arriver de se répéter mentalement les items précédents pendant la présentation de certains items. Cette hypothèse a été rapidement écartée comme cause principale des effets de pratique distribuée car la répétition mentale est une activité volontairement mise en œuvre par le sujet lorsqu'il fait un effort pour apprendre des informations. Or, l'effet d'espacement a été observé même en situation d'apprentissage incident (e.g., Shaughnessy, 1976 ; voir cependant Greene, 1989). L'hypothèse de la consolidation stipule quant à elle que le passage du temps entre P1 et P2 permet à la trace mnésique élaborée à P1 de se consolider. Si l'épisode P2 survient trop tôt, cette consolidation est alors interrompue, ce qui est défavorable à la rétention finale. Au contraire, si l'épisode P2 survient après un IIR d'une certaine durée, la consolidation de P1 a pu atteindre un niveau plus avancé et la rétention finale sera alors plus élevée. Cette hypothèse de la consolidation a également été rapidement écartée, car elle fait reposer l'effet de la pratique distribuée sur P1 alors que de nombreuses études ont montré que le locus de l'effet d'espacement est P2 (e.g., Hintzman, Block, & Summers, 1973). Nous développerons toutefois, dans la section 3.3, les conceptions actuelles sur la consolidation dans le champ des neurosciences, qui sont toujours discutées dans le cadre de l'effet d'espacement.

Deuxièmement, deux hypothèses faisant reposer l'origine de l'effet sur P2 ont été proposées. Ces hypothèses, dites du traitement déficitaire, stipulent que les sujets allouent

moins d'attention au stimulus répété lors de P2 lorsque l'IIR est court que lorsqu'il est plus long. L'hypothèse inattentionnelle de Hintzman (1974) propose que des mécanismes volontaires sont à l'œuvre dans ce processus. Le sujet, dans une condition d'intervalle court (par exemple 5 secondes), porterait moins attention au traitement de P2 que dans la condition d'intervalle long (par exemple 2 minutes), car la répétition lui procurerait alors un « sentiment de savoir ». Cette hypothèse a été écartée car, même dans des conditions où le sujet obtient une récompense proportionnelle à son rappel final, la condition massée reste moins favorable au rappel final que la condition distribuée (e.g., Hintzman, et al., 1975). Ainsi ce mécanisme d'attribution volontaire d'attention ne peut pas expliquer à lui seul les effets d'espacement et d'intervalle. On ne peut toutefois pas exclure que de tels effets attentionnels se produisent et ont un rôle explicatif partiel dans certains effets obtenus. L'hypothèse de l'amorçage (*priming*) considère au contraire que des mécanismes involontaires sont à l'œuvre dans ce processus. Le sujet, dans une condition d'IIR court, allouerait moins de ressources au traitement de P2 que dans la condition d'IIR long, car, ayant toujours l'item en mémoire à court terme, il ne réaliserait les mêmes traitements sur l'information que dans une condition d'intervalle long. Cette hypothèse a reçu certains appuis empiriques, et nous développerons ce point dans la section 2.4 et dans la section 3.1 sur l'imagerie fonctionnelle.

Enfin, des théories moins intuitives ont également été proposées. Les deux principales théories en vigueur actuellement sont celles de la variabilité de l'encodage et de la récupération en phase d'étude. D'autres hypothèses moins populaires mais néanmoins intéressantes sont celles de l'accessibilité - reconstruction et celle de la métacognition. Etant donné que notre problématique porte sur les effets de la répétition sur le long terme, nous orienterons dans la mesure du possible la présentation des théories en ce sens. Le lecteur désireux d'en savoir plus sur les théories expliquant davantage les effets à court terme peuvent se référer aux revues de Hintzman (1974), Cepeda et al. (2006), et Crowder (1976, chapitre 9).

2. 1. Théorie de la variabilité de l'encodage

2. 1. 1. Généralités et historique

La théorie de la variabilité de l'encodage est l'une des théories les plus souvent mentionnées dans l'explication des effets d'espacement et d'intervalle. Elle est toutefois également évoquée pour expliquer d'autres effets psychologiques, comme par exemple les effets de récence et de contiguïté (e.g., Howard & Kahana, 2002). Elle est intitulée dans la littérature « encoding variability theory », mais certains auteurs l'intitulent également « contextual-variability theory ». L'objectif de cette section est de décrire cette théorie ainsi que de mettre en avant ses atouts et ses limites dans l'explication des effets de pratique distribuée. Nous profitons de cette introduction pour situer le cadre conceptuel de la mémoire dans lequel nous nous positionnons. D'une façon globale, nous considérons que la mémoire déclarative est constituée de traces, créées par le processus d'encodage, qui sont réactivées au moment de la récupération. L'encodage est le processus par lequel une information perçue est transformée en trace mnésique dans l'esprit et le cerveau.

D'une façon générale, l'hypothèse de la variabilité de l'encodage postule qu'une même information peut être encodée en mémoire de plusieurs façons différentes. En outre, plus les différents encodages d'une même information sont différents entre eux, plus seront disponibles de « routes » de récupération de cette information au moment où celle-ci devra être récupérée. Autrement dit, plus une information aboutit à une (ou

des) trace(s) mnésique(s) diversifiée(s), meilleur sera le rappel ultérieur. Un des facteurs les plus importants susceptibles de faire varier la façon d'encoder l'information est le contexte présent au moment de l'encodage. En effet, la trace mnésique créée lors de l'encodage d'une information inclut également certains éléments du contexte courant. Le second principe de la théorie est que, lorsque l'IIR augmente, les contextes d'encodage sont de plus en plus différents entre P1 et P2. Il découle de ces postulats que de longs IIR sont bénéfiques à la récupération finale. Nous constaterons également que certaines formalisations de la théorie mettent en avant l'idée que la performance de récupération finale est déterminée par la concordance entre les éléments contextuels présents lors du test et les éléments contextuels encodés dans les traces mnésiques. Nous reviendrons en détail sur ces points.

Certains principes clés de la théorie de la variabilité de l'encodage ont été énoncés la première fois par Estes en 1955, qui a avancé l'idée novatrice de *fluctuation aléatoire* des stimuli lors des intervalles entre les essais successifs. Selon l'auteur, un stimulus exactement identique pouvait être traité par le sujet de façon différente lors de différents essais, et ces fluctuations de traitement suivaient des lois probabilistes. Cependant, le modèle proposé par Estes visait à expliquer les phénomènes de conditionnement et d'extinction et non directement les effets observés en mémoire déclarative qui nous intéressent ici.

Le terme de « variabilité de l'encodage » apparut en 1968 dans un article de Martin. Celui-ci postulait, dans le cadre d'un apprentissage déclaratif -en l'occurrence un apprentissage de paires de mots- qu'un même item peut être encodé de différentes façons. En particulier, un même item peut se voir attribuer différentes significations, ou différentes interprétations par le sujet, lors de ses différentes apparitions. Le fait d'interpréter de façon similaire ou différente un même item qui est présenté deux fois va influencer sa récupération ultérieure.

Il est important de noter dès à présent que dès ces premières énonciations de la théorie, la reconnaissance des stimuli lorsqu'ils étaient répétés (c'est-à-dire la *récupération en phase d'étude*) avait une grande importance. En effet, dans des articles précédents (e.g., Martin, 1967), Martin avait observé que la performance finale était diminuée si les items répétés n'étaient pas reconnus lors de leur répétition. Nous reviendrons sur ce point dans la section 2.2.

2. 1. 2. Le modèle de Bower (1972)

C'est Bower en 1972 qui a véritablement développé la notion de variabilité de l'encodage, en reprenant les formulations de Estes (1959) et de Martin (1968). Bower plaçait au centre de sa théorie la notion d'encodage, en définissant l'encodage comme un processus opérant sur le *stimulus réel* et fournissant en sortie le *stimulus fonctionnel*, cette transformation se faisant par le biais de divers *opérateurs d'encodage*. Selon Bower, plusieurs stimuli fonctionnels sont possibles pour un stimulus réel donné, et le choix de l'encodage réellement opéré sur un stimulus est un événement probabiliste. De plus, l'auteur postulait que, lors d'un essai, un seul et unique encodage se produit sur le stimulus réel. Il énonçait également qu'un stimulus est reconnu comme une répétition seulement si il a été encodé de la même façon lors des deux présentations.

Le choix des opérateurs d'encodage mis en oeuvre sur un stimulus est déterminé entre autres par un ensemble de facteurs contextuels. Selon Bower, le contexte comprend les stimulations externes prévalentes lors de la présentation des stimuli (comme la posture, la température, la pièce, les bruits de fond, etc) ainsi que les stimuli internes physiologiques

(comme la gorge sèche, les battements cardiaques, les gargouillis d'estomac, la nausée, l'ennui, etc). Encore plus important, selon Bower, est ce que le sujet est en train de penser, le « *monologue interne produit par le flux libre de la conscience* » qui décrit la situation expérimentale. Cette description par le sujet contient sa conception, instant par instant, de la structure de la tâche expérimentale, les instructions, la nature du matériel rencontré, les stratégies d'encodage utilisées et ce qu'il pense être utile de faire, et ce qu'il pense du but de l'expérience. Il survient tout un ensemble d'associations libres qui comportent des pensées sur « *la stupidité (sic) de l'expérimentateur, sur sa propre stupidité, ce qu'il a mangé au déjeuner, ce qu'il fera ce soir* », etc... L'auteur suppose que ces facteurs influencent l'état de la machinerie d'encodage de telle sorte que les certains types d'encodage deviennent plus ou moins probables par rapport à d'autres. Autrement dit, des changements du contexte auront pour conséquence des changements dans l'encodage. Ces changements sont de deux types, systématique (intentionnel) et aléatoire (involontaire).

L'expérimentateur peut produire des changements systématiques dans l'encodage en présentant du matériel non destiné à être appris explicitement, comme des éléments contextuels, en même temps que les items cibles. Le sujet a alors pour consigne de penser à la relation entre le matériel et le contexte. On peut utiliser par exemple des mots polysémiques, tel que *avocat*, et induire un sens différent à chaque présentation. On peut aussi utiliser des éléments qui vont attirer l'attention des sujets, comme par exemple écrire une lettre en rouge. On peut aussi faire varier le contexte en modifiant le moment pendant lequel le sujet passe l'expérience : après manger, le matin ou la nuit, dans une salle différente, debout ou assis, la couleur du fond d'écran, etc.

L'auteur pense cependant que ce sont les changements aléatoires qui sont responsables des effets de pratique distribuée. Bower postule qu'il se produit un phénomène de dérivation (*drift*) graduelle du contexte au fil du temps dans la mesure où un ensemble d'événements se produisent pendant le passage du temps expérimental. Ainsi, lorsque le temps passe suite à un instant t , le contexte courant est de plus en plus différent du contexte au moment t . Ce changement contextuel induit une modification des opérateurs d'encodage. Selon l'auteur cependant, ce changement contextuel est rarement total, les « états d'esprits » revenant de façon cyclique.

Dans le cas où une répétition d'un stimulus a lieu, la seconde occurrence va faire augmenter la performance de mémoire lors d'un test ultérieur, car elle fournit une occasion pour de nouveaux éléments d'être encodés et ajoutés à la trace mnésique. L'avantage apporté par ce second épisode dépend de l'amplitude des changements d'encodage réalisés pendant l'intervalle entre les présentations, puisque plus il y a d'encodages réalisés, plus la récupération ultérieure sera facilitée.

2. 1. 3. Le modèle de Glenberg (1979)

La théorie de Bower (1972) n'a cependant pas été élaborée pour expliquer directement les effets d'espacement ou d'intervalle. Glenberg (1979) a fait une synthèse générale des théories de la variabilité de l'encodage (Bower, 1972 ; Estes, 1955 ; Martin, 1968 ; Glenberg, 1976) en l'affinant afin qu'elle rende compte des effets de pratique distribuée. Pour résumer, sa théorie énonce qu' « *un stimulus subit des changements d'encodage de présentation en présentation, et l'amplitude du changement est positivement corrélée avec l'espacement des présentations. Ces changements dans l'encodage ajoutent de l'information à la représentation épisodique des présentations du stimulus et sous-tendent les améliorations de performance.* » Cette énonciation est similaire aux formulations antérieures. Le changement majeur par rapport aux théories précédentes est que l'auteur

n'évoque pas l'idée qu'on encoderait différemment les items répétés, c'est-à-dire que les processus d'encodage seraient de nature différente au fil des répétitions. Il propose plutôt que ce sont les éléments contextuels stockés dans la trace aux côtés de la représentation du stimulus qui varient, mais que, dans la trace, l'élément focal (c'est-à-dire représentant le stimulus lui-même) est encodé de la même façon.

La particularité de cette théorie dans l'explication des effets de pratique distribuée est d'insister sur trois points essentiels. Premièrement, une répétition est potentiellement efficace dans la mesure où la seconde présentation permet le stockage d'informations distinctes de celles stockées lors de la première présentation. Deuxièmement, la réalisation de ce potentiel est contrôlée par les conditions présentes au moment de la tâche de récupération, c'est-à-dire l'environnement présent lors de cette récupération. Troisièmement, la théorie énoncée par Glenberg postule que les effets d'espacement peuvent être dus à la variation de l'encodage de plusieurs types de composants de l'information. Trois types de composants sont considérés comme importants dans les tâches de mémoire verbale : les composants contextuels, descriptifs, et structurels, que nous définirons plus loin. Nous décrirons dans un premier temps les mécanismes en jeu lors des processus d'encodage et de stockage selon cette théorie, puis ceux mis en jeu lors des processus de récupération, ce qui permettra enfin d'expliquer les effets de la pratique distribuée couramment observés.

Les processus d'encodage et de stockage

Selon le modèle de Glenberg (1979), lors de la présentation d'un stimulus verbal, les processus perceptuels encodent les caractéristiques physiques du stimulus. Ces caractéristiques sont utilisées pour accéder aux composants sémantiques associées au stimulus. Une fois ces composants récupérés, une trace à composants multiples est créée en mémoire pour représenter l'occurrence du stimulus dans l'expérience. Cette trace contient des composants propositionnels issus de la mémoire sémantique, ainsi que d'autres composants de la trace, spécifiques à l'épisode, et qui sont déterminés conjointement par le stimulus traité, la nature de la tâche, les stratégies du sujet et le contexte dans lequel le stimulus a été présenté. Ce modèle se place dans le courant général des modèles à traces multiples développées dans les années 1970 (e.g., Hintzman, 1988 ; Hintzman & Block, 1971).

Les composants de trois niveaux, contextuels, structurels, ou descriptifs, diffèrent quant à la probabilité qu'ils soient inclus dans les traces représentant les différents items ; un composant inclus dans de nombreuses traces est dit général. Les composants contextuels sont généraux car les mêmes composants contextuels sont inclus dans toutes les traces encodées dans un contexte spécifique.

Les *composants contextuels* représentent le contexte dans lequel l'événement est encodé et incluent des informations telles que les caractéristiques de l'environnement physique, le temps, et les état affectifs et cognitifs du sujet. Les événements perçus sont inclus dans le flot de conscience du sujet, comme le mentionnait Bower (1972). Par exemple, la trace qui représente la présentation d'un mot dans une tâche de rappel libre contient le fait que ce mot soit présenté, mais également qu'il l'a été dans une pièce donnée, dans une certaine liste, en réalisant une opération cognitive précise. Les composants contextuels ont un statut particulier dans le sens où les éléments contextuels sont encodés *automatiquement* pendant la tâche et parce qu'on considère qu'ils peuvent influencer l'encodage des autres composants contenus dans la trace. Le fait que tous les composants

soient finalement soumis aux variations de contexte fait que la variation contextuelle est primordiale dans cette théorie.

Une répétition est potentiellement efficace dans la mesure où des informations différentes sont stockées lors des deux présentations d'un événement répété. Le contexte lui-même est constitué de divers éléments qui ne varient pas tous au même rythme au fil du temps, et ont donc des taux de fluctuation variables. Il existerait les composants dits locaux, qui fluctuent rapidement et continuellement, comme par exemple les items présentés. D'autres composants fluctuent plus lentement, comme par exemple l'augmentation de la sensation de faim au fil de l'expérience. D'autres composants enfin, les caractéristiques dites globales du contexte, sont invariants ou fluctuent très lentement, et sont donc très généraux. Selon l'auteur, on peut expérimentalement séparer les effets dus aux différents taux de fluctuation des composants contextuels dans leur influence sur la performance mnésique finale.

L'amplitude de la fluctuation du contexte et donc la quantité de changement dans les composants contextuels stockés dans la trace est directement liée à l'intervalle entre les présentations d'un même item. La corrélation positive entre le nombre de composants différents stockés dans la trace et l'IIR est appelé « stockage différentiel » (*differential storage*).

Les *composants structurels* représentent la structure que les sujets imposent sur les événements individuels, c'est-à-dire les associations, groupements et catégorisations réalisés sur les stimuli. Ces composants ne sont pas encodés automatiquement mais dépendent des processus de contrôle engagés par le sujet. Ces composants spécifiques sont déterminés par le contexte local, défini par exemple par les items présentés juste avant ou juste après un item cible. Lors de la répétition d'un item, le sujet peut, par exemple, être en train de traiter un ensemble d'items différents de ceux de la première présentation, ce qui permet le stockage de composants différents dans la trace du stimulus. Puisque des changements contextuels induisent des changements dans les composants structurels, les composants structurels seront eux-mêmes caractérisés par un stockage différentiel lorsque l'IIR augmente.

Lors de la répétition d'un stimulus, la variabilité dans les composants structurels encodés est fonction à la fois des processus de contrôle et du contexte local. On peut inciter le sujet à encoder différents aspects du stimulus lors des deux occurrences en changeant la tâche d'orientation réalisée lors des deux occurrences. Plus simplement, les changements dans le contexte local peuvent faire varier les composants structurels inclus dans la trace.

Les *composants descriptifs*, enfin, concernent les éléments strictement relatifs aux stimuli. Par exemple, dans l'apprentissage de mots, ces composants incluent des informations sur l'orthographe, l'articulation, la signification, etc. Ces composants dépendent du traitement dans lequel le sujet est engagé et du contexte, par exemple de la profondeur de traitement. Bien que les composants descriptifs soient les plus uniques inclus dans la trace, c'est-à-dire les plus spécifiques à la trace, il est possible de manipuler leur encodage ou de créer des stimuli qui incluent tous les mêmes composants. Par exemple, forcer le sujet à prononcer à voix haute des stimuli qui riment entraîne l'inclusion de composants descriptifs similaires dans toutes les traces. Ces composants deviendront alors généraux.

Selon l'auteur, la trace d'un stimulus répété après un IIR long inclura un plus grand nombre de composants contextuels, structurels et descriptifs que la trace d'un stimulus répété après un IIR plus court. Cependant, la théorie ne prédit pas toujours une augmentation de la performance mnésique ultérieure proportionnelle à l'amplitude

du stockage différentiel (c'est-à-dire l'amplitude avec laquelle des composants différents et variés sont stockés dans la trace lors des répétitions). En réalité, la corrélation entre stockage différentiel et performance à la tâche mnésique peut varier en amplitude et même en direction en fonction de l'environnement de récupération. Nous allons décrire ces processus dans la partie suivante.

Les processus de récupération

Se souvenir d'un stimulus spécifique nécessite de récupérer la trace épisodique représentant l'occurrence de ce stimulus dans l'expérience. L'accès aux traces est rendu possible, entre autres, par des indices de récupération, qui sont les éléments présents dans l'environnement au moment où la récupération est réalisée. Pour permettre effectivement de récupérer une ou des trace(s) épisodique(s) stockée(s), ces indices doivent contenir des composants similaires aux composants inclus dans les traces, et, par conséquent, des traces elles-mêmes. Ainsi, par similarité, les composants de l'indice déclencheront une activation temporaire des composants identiques stockés dans les traces. Selon la théorie de Glenberg (1979), le degré d'activation d'un composant dans une trace individuelle est inversement lié au nombre de traces dans lesquelles ce composant est inclus, c'est-à-dire sa généralité. Ainsi, plus un composant est général, c'est-à-dire présent dans de nombreuses traces, moins il va déclencher une activation forte des traces qui le contiennent. À l'opposé, un composant de l'environnement très spécifique à une trace unique va entraîner l'activation intense de ce composant dans cette trace. D'une façon générale, le degré avec lequel une trace est activée est proportionnel à la somme des activations de ses composants individuels. Par conséquent : 1) l'activation de la trace et sa récupération sont fonction du nombre de composants partagés par l'indice et la trace ; 2) l'activation de la trace est faible lorsque les composants de la trace ou de l'indice sont généraux ; 3) l'activation de la trace augmente avec le nombre de composants inclus dans la trace.

Les composants spécifiques présents dans les indices dépendent de la nature du test de mémoire, c'est-à-dire de l'information disponible au moment du test. Typiquement, cette information inclut les consignes, le contexte au moment du test et les indices explicites de récupération fournis. Les indices et leurs effets diffèrent également selon le type de tâche demandée au sujet.

Dans le cas d'une tâche de *reconnaissance*, on fournit au sujet des copies physiques des stimuli originaux ainsi que des distracteurs et il a pour tâche de discriminer les items anciens des items nouveaux. Le sujet commence la tâche en encodant les stimuli un par un, et produit donc des indices comportant à la fois les composants descriptifs des stimuli et les composants contextuels instanciés à ce moment-là. Ces indices activent alors des composants identiques contenus dans les traces en mémoire. Si la reconnaissance se produit après un court DR, alors les composants contextuels de l'indice vont activer les traces contenant des composants contextuels similaires, et donc des traces récentes. En effet, le contexte fluctue de façon progressive et régulière, et par conséquent celui-ci n'aura pas beaucoup fluctué depuis les occurrences récentes. Après des DR plus longs, les composants contextuels jouent un rôle décroissant dans l'accès aux traces pour deux raisons. Premièrement, après un long DR, les composants contextuels de l'indice ne correspondent pas aux composants contextuels des traces mnésiques, donc les composants de l'indice n'activent aucun composant des traces. En effet, le contexte aura eu le temps de fluctuer de façon très importante. Deuxièmement, les composants contextuels partagés par l'indice et par la trace sont alors inclus dans de nombreuses traces (i.e., ils sont très généraux), et ainsi chaque trace sera seulement faiblement activée. En réalité, en

tâche de reconnaissance, la source prédominante d'activation de traces sera l'ensemble des composants descriptifs. L'indice, c'est-à-dire le stimulus présenté, contient un grand nombre de composants descriptifs uniques. Généralement, si l'indice est physiquement identique à un mot présenté antérieurement, la trace de l'item sera fortement activée et entraînera la reconnaissance de cet item. D'autre part, si l'encodage des composants descriptifs lors de la tâche de reconnaissance est radicalement différent de l'encodage lors de l'apprentissage, alors l'indice ne va pas permettre de récupérer la trace.

Les mêmes processus sont en jeu dans d'autres types de tâches mais avec des composants d'autres niveaux. Dans le cas d'un *rappel libre*, les seuls indices disponibles sont ceux de l'environnement présent lors du test. En effet, dans une tâche de rappel libre, le sujet a pour tâche de récupérer les informations qu'il a encodées auparavant, sans bénéficier d'aucun indice explicite, contrairement à la tâche de reconnaissance. Après un DR court, les composants contextuels locaux vont activer les traces qui comportent des composants similaires, c'est-à-dire ceux traités en fin d'apprentissage. Ce ne sera pas le cas après un long DR puisque ce seront alors les composants contextuels globaux qui déclencheront les activations de toutes les traces. Comme ces composants sont inclus dans les traces de tous les items appris, cependant, l'indice est relativement peu efficace. Il faut également mentionner que les traces récupérées avec les indices contextuels peuvent elles-mêmes devenir des sources d'indices de récupération.

L'effet de pratique distribuée

Le modèle ainsi défini, nous pouvons maintenant comprendre comment il permet d'expliquer les effets de pratique distribuée, ainsi que l'influence du délai de rétention, dans le cas d'une tâche finale de reconnaissance ou de rappel libre.

Comme l'énonce le modèle, augmenter l'IIR va typiquement mener à un stockage différentiel des composants sur un ou plusieurs niveaux. Or, en général, le stockage différentiel va faciliter la récupération ultérieure. Plus la trace mnésique contient un nombre important de composants différents, plus elle aura de chances de partager des composants avec l'indice lors du test de récupération, et donc être activée et récupérée.

Le modèle parvient à expliquer pourquoi, dans certaines situations, l'effet d'intervalle est monotone croissant et dans certaines autres, il prend une forme en U inversé (e.g., Glenberg, 1976). En effet, à tout niveau, le fait de manipuler la relation entre l'indice de récupération et l'information stockée dans la trace peut produire une performance de récupération en fonction de l'IIR qui peut être monotone décroissante, en U inversé, ou monotone croissante. La Figure 5 décrit ces prédictions en considérant un seul niveau de composant (contextuel, structurel, ou descriptif). Dans la figure du haut sont représentés de haut en bas les moments durant lesquels un événement peut être présenté et répété, de t1 à t8. En face de chaque moment sont représentés les composants encodés à ce moment-là. Les composants sont symbolisés par des lettres de l'alphabet allant de g à q, qui symbolisent les composants présents dans l'environnement et encodés dans la trace. Avec le passage du temps, ces composants environnementaux évoluent progressivement, ce qui est symbolisé par l'avancée progressive des lettres dans l'alphabet. Au moment t1 on est en présence des composants g, h, i, et j, au moment t2 on évolue vers les composants h, i, j, et k, et ainsi de suite, jusqu'au moment t8 qui contient les composants n, o, p, et q. Il s'agit, bien entendu, d'une schématisation, les changements contextuels dans une expérience n'étaient certainement pas aussi simples que dans cette représentation.

Temps		Composants	
...		...	
t1		ghij	
t2		hijk	
t3		ijkl	
t4		jklm	
t5		klmn	
t6		lmno	
t7		mnop	
t8		nopq	
...		...	

Intervalle inter-répétition		Structure de la répétition			
		0	1	2	6
Moments de présentation		t7 + t8	t6 + t8	t5 + t8	t1 + t8
Composants dans la trace		mnopq	lmnopq	klmnopq	ghijnopq
Nombre de composants activés	Indice n°1 (mnop)	4	4	4	3
	Indice n°2 (lnqs)	2	3	3	2
	Indice n°3 (random)	4-5 / n	4-6 / n	4-7 / n	4-8 / n

Figure 5 : Reproduction du schéma illustrant le modèle de Glenberg (1979).

Dans la figure du haut sont présentés les moments auxquels un item peut être présenté et les composants (d'un seul niveau) stockés lors de la présentation. Dans la figure du bas sont présentés les composants inclus dans les traces formées lors des répétitions après quatre intervalles différents, la structure de trois indices de récupération, et le nombre de composants inclus dans les traces activées par les indices.

Considérons un item présenté à deux reprises avec des IIR différents, et les composants encodés dans la trace. Ces situations sont représentées dans la figure du bas, avec P2 se produisant toujours à t8. Par exemple, un item présenté au moment t5 puis répété au moment t8 produit une trace contenant des composants allant de k à q (troisième colonne). De la même façon, un item présenté au moment t7 puis répété au moment t8 produit une trace contenant des composants allant de m à q (première colonne). On remarque que plus l'IIR est long, plus le nombre d'éléments contextuels encodés est important, ce qui illustre la notion de stockage différentiel.

La figure représente également, en bas de la figure du bas, l'activation de la trace fournie par trois indices différents parmi tous les indices possibles au moment de la

récupération (Indices n°1, 2, et 3). A l'intersection de la ligne (représentant l'indice lors du test) et de la colonne (représentant l'IIR), est représenté le nombre de composants en commun entre la trace et l'indice. L'activation de la trace, et par conséquent la performance finale, est proportionnelle à ce nombre de composants en commun. Les indices de récupération diffèrent par les composants qu'ils contiennent et peuvent être plus ou moins biaisés envers des composants stockés dans la trace à t8. En effet, le *biais* de l'indice se définit comme la tendance à inclure dans l'indice des composants stockés lors d'une occurrence de l'item (P1 ou P2), et d'exclure ceux stockés lors des autres occurrences. À l'inverse, un indice non biaisé contient des composants tels que les composants encodés dans la trace lors des deux présentations ont la même probabilité d'être inclus dans l'indice.

Le biais de l'indice peut être contrôlé par une manipulation expérimentale directe, ou par le DR de façon indirecte. Quand la récupération se produit après un court DR, le contexte présent au moment de la récupération est similaire au contexte au moment de la seconde occurrence. Les composants contextuels inclus dans l'indice de récupération seront par conséquent biaisés envers ceux encodés lors de la seconde présentation. Cette situation correspond à l'indice n°1, représenté dans la Figure 5. On observe que pour un tel indice, ce sont les IIR les plus courts qui induisent les meilleures performances de récupération : l'activation est la plus forte (niveau 4) pour les IIR plus courts, tandis qu'elle est la plus faible (niveau 3), pour l'IIR de longueur 6, c'est-à-dire un item présenté à t1 et t8.

L'indice n°2 est un exemple de biais envers les composants présentés lors de la seconde occurrence, mais le biais est moins important que pour le premier indice. L'activation de la trace, basée sur le second indice, est une fonction en U inversé de l'IIR. En effet, les IIR associés à la plus grande activation sont les IIR de niveau 1 et 2.

L'indice n°3 n'est biaisé envers aucune présentation. Les nombres présentés dans la figure indiquent le nombre attendu de composants activés en considérant que quatre composants différents (sur un total de n possible) sont choisis aléatoirement pour l'inclusion dans l'indice. Seulement dans ce dernier cas la performance est une fonction monotone de l'IIR. En effet, plus l'IIR est long, plus le nombre de composants différents encodés dans la trace est important. Dans la plupart des expériences, les indices de récupération ne sont pas biaisés, ce qui correspond à un indice du type 3 et à l'habituel effet d'intervalle monotone.

Pour résumer, cette théorie permet de rendre compte de la forme de la fonction de performance en fonction de l'IIR et du DR, par l'intermédiaire d'un jeu de chevauchement des composants dans la trace et dans les indices au moment du test. C'est la théorisation de l'hypothèse de la variabilité de l'encodage la plus aboutie et la plus précise disponible dans la littérature. Un ensemble d'expériences a présenté des éléments allant dans son sens. C'est l'objet de la section suivante.

2. 1. 4. Arguments expérimentaux en faveur de la théorie

Dans son article de 1979, Glenberg proposait trois expériences visant à tester les prédictions de la théorie de la variabilité de l'encodage en manipulant les différents niveaux de composants. Nous présenterons les deux premières expériences. Dans la première, des paires de mots liés entre eux étaient présentées, par exemple *lame-couteau*, le mot cible étant le second mot (*couteau*). Dans la condition dite de *constance* d'encodage, la paire était répétée à l'identique (*lame-couteau*). Dans la condition dite de *variation de l'encodage*, le mot cible était répété au sein d'une paire contenant un autre mot indice (*cuillère-couteau*). En plus de cette manipulation, certaines paires étaient répétées en condition massée et d'autres en condition distribuée. Le facteur espacement servait à induire un stockage différentiel sur les composants contextuels, tandis que le facteur type d'encodage

servait à induire un stockage différentiel sur les composants descriptifs. Les prédictions sur l'interaction entre le type d'encodage et l'espacement étaient différentes selon la tâche de récupération employée.

Dans une tâche de *rappel indicé*, ce sont les composants descriptifs qui sont responsables de l'accès à la trace car ils sont fournis dans l'indice lors du test. On ne devrait donc pas observer l'effet d'espacement en condition de *variation de l'encodage* car on a alors induit un maximum de variabilité pour l'encodage des composants descriptifs, ainsi le rappel est élevé et non influencé par l'espacement. Ce sont les résultats qui ont été observés. En *rappel libre*, ce sont les composants contextuels présents lors du test qui servent d'indices permettant l'accès à la trace. Comme le stockage différentiel des composants contextuels est indépendant de la manipulation du type d'encodage, l'effet d'espacement devrait être observé dans les deux conditions d'encodage. C'est également ce résultat qui a été observé.

Dans l'Expérience 2, plusieurs listes de mots étaient présentées puis une tâche de rappel libre avait lieu. L'effet d'intervalle était testé en répétant certains mots au sein d'une liste avec des IIR différents. L'auteur manipulait le stockage différentiel des composants contextuels globaux en réalisant des répétitions des mots entre les listes, en considérant que les autres types de composants, c'est-à-dire les composants structurels et descriptifs, maintenaient un stockage différentiel constant. Le DR était manipulé : le rappel libre avait lieu soit immédiatement après la présentation de la liste, soit immédiatement après la présentation de toutes les listes, soit deux heures après la présentation de toutes les listes. La manipulation du DR servait à biaiser ou non les composants contextuels globaux instanciés au moment du rappel libre. Les résultats ont montré que pour les items répétés dans des listes différentes, la performance de rappel en fonction de l'IIR inter-liste était monotone décroissante lorsque le rappel libre était immédiat, et monotone croissante lorsque le rappel libre était tardif. Ces deux expériences ont donc montré que la manipulation des composants de différents niveaux, ainsi que de la durée du DR, peuvent influencer la forme de l'effet d'intervalle, et ce de façon cohérente avec les prédictions du modèle.

D'une façon générale, et en mettant de côté l'existence de plusieurs niveaux de composants, un des arguments expérimentaux en faveur de cette théorie devrait être l'observation qu'une induction de variabilité de l'encodage élimine ou atténue l'effet d'espacement, comme ce qui a été observé dans les Expériences 1 et 2 de Glenberg (1979).

Les premières études ayant fait varier le contexte d'encodage entre deux présentations d'un item ont conduit à des conclusions mitigées. Certaines ont testé indirectement l'hypothèse de la variabilité de l'encodage, c'est le cas de Hintzman, et al. (1973), et de Hintzman, Summers, Eki, & Moore (1975, Expérience 1). Hintzman et al. (1973) ont fait varier la modalité sensorielle des présentations, ou l'ont gardée identique, et n'ont observé d'influence de cette manipulation ni sur la fréquence d'occurrence estimée par les sujets, ni sur l'effet d'espacement. Hintzman et al. (1975, Expérience 1) ont modifié le contexte lors de P2 en y ajoutant un signal sonore, mais cela n'a pas modifié l'effet d'espacement observé. Ces deux observations contredisent donc les prédictions de la théorie de la variabilité de l'encodage.

Parmi les études ayant modifié le contexte d'encodage directement pour tester l'hypothèse de la variabilité de l'encodage, citons Dempster (1987), Maskarinec et Thompson (1976), et Postman et Knecht (1983). Dempster a comparé la rétention de vocabulaire dans trois conditions : a) seule la définition du mot était fournie, b) une phrase de contexte était présentée suite à la définition, c) trois phrases de contexte étaient présentées suite à la définition. L'effet d'espacement a été observé, et l'introduction de

phrases contextuelles n'a pas montré d'effet significatif et n'a pas non plus influencé l'effet d'espacement. Maskarinec et Thompson ont testé deux façons de modifier le contexte pour avoir des conditions d'encodage identiques ou différentes au fil des répétitions. Dans la première expérience, les auteurs ont modifié les items adjacents aux items cibles ; dans la deuxième expérience, en condition d'apprentissage incident, ils ont fait varier la tâche d'orientation réalisée sur les items. Dans aucun cas la variabilité de l'encodage n'a interagi avec l'effet d'espacement.

Postman et Knecht (1983) avaient pour objectif de vérifier si la variation de l'encodage était dans tous les cas plus favorable au rappel par rapport à une condition de constance d'encodage. Dans l'Expérience 1, ils faisaient lire aux sujets des phrases contenant chacune un mot cible, qui était l'objet du rappel final. Chaque mot cible était présenté trois fois, soit dans la même phrase répétée, ce qui constitue la condition de *constance* d'encodage, soit dans trois phrases différentes, ce qui constitue la condition de *variabilité* de l'encodage. Dans ce dernier cas, les trois phrases différentes induisaient la même signification du mot cible. Ensuite, une tâche de rappel libre ou de rappel indicé avaient lieu. Le rappel indicé fournissait comme indice une ou plusieurs phrases sans le mot cible. En condition *d'indice unique*, le rappel indicé fournissait une seule phrase contexte, tandis qu'en condition *d'indices multiples*, le rappel indicé fournissait les trois contextes différents. Il a été montré que pour une condition donnée d'encodage, la meilleure performance était obtenue lorsque le nombre d'indices fournis correspondait au nombre d'indices pendant l'encodage. La condition de *variabilité* n'était donc pas toujours avantageuse.

Dans l'Expérience 2, Postman et Knecht (1983) ont répliqué l'expérience précédente en prenant en compte uniquement les conditions où le nombre d'indices lors de l'encodage et lors du test correspondaient. Ils comparaient donc la situation où une phrase contexte était présentée à l'étude et lors du test, avec la condition où trois phrases contexte étaient présentées à l'étude et lors du test. Ils ont également ajouté une condition dans laquelle le test se produisait 24h plus tard. En rappel libre, il n'y avait pas de différence de performance entre les deux conditions, et le rappel diminuait selon la même amplitude entre le test immédiat et le test différé. En rappel indicé, la condition avec une phrase à l'encodage et lors du test était plus efficace, et les deux conditions déclinaient selon la même amplitude entre le test immédiat et le test différé. Ainsi, l'avantage de la variabilité de l'encodage n'était pas observé.

Ainsi, bien que la théorie de la récupération en phase d'étude ait été formalisée de façon détaillée et présente de nombreux attraits, tous les arguments expérimentaux ne vont pas dans son sens. Nous allons maintenant aborder un deuxième type d'explication, la récupération en phase d'étude, et nous verrons ensuite que, loin d'être antagonistes, ces deux hypothèses peuvent se combiner.

2. 2. Théorie de la récupération en phase d'étude

Le terme de récupération en phase d'étude (*study-phase retrieval*) apparaît pour la première fois dans l'article de Hintzman, Summers, et Block (1975). Pourtant le concept principal sous-tendant la théorie avait déjà été décrit dans des études antérieures, par exemple dans Hintzman et Block (1973). La théorie de la récupération en phase d'étude consiste en la proposition suivante : « *un effet typique d'une seconde présentation d'un item est de récupérer la trace de la première présentation, incluant les associations contextuelles. Cela est possible uniquement si cette première trace est toujours accessible en mémoire* » (Hintzman & Block). Tout comme l'hypothèse de la variabilité de l'encodage,

l'hypothèse de la récupération en phase d'étude n'est pas limitée à l'explication des effets d'espacement et d'intervalle, mais peut aussi expliquer d'autres phénomènes, en particulier les effets observés lors d'un rappel de la position sérielle et lors d'un jugement de récence. Par exemple, dans l'article de Hintzman et Block, les sujets devaient estimer la durée de l'espacement qui séparait deux occurrences d'un item pendant la phase d'apprentissage. La théorie de la récupération en phase d'étude avait l'avantage de pouvoir expliquer pourquoi les jugements d'espacement pour un même item sont très précis, et comment l'information sur l'espacement est stockée en mémoire. L'hypothèse est que, lors de P2, on récupère la trace encodée lors de P1 ainsi que l'estimation de la durée depuis P1, et on stocke cette information. Au contraire, les jugements d'espacement pour deux items différents sont généralement beaucoup moins précis, ce qui pourrait s'expliquer par le fait que la présentation d'un item donné ne déclenche pas la récupération d'un autre item précédemment présenté et, par conséquent, l'information sur la récence de cet autre item n'est pas stockée en mémoire.

Un autre argument en faveur de l'existence de la récupération en phase d'étude se trouve dans Tzeng et Cotton (1980). Les auteurs présentaient une liste de mots aux sujets qui devaient les apprendre sans connaître la nature du test final. Ce dernier était un test de discrimination de récence, c'est-à-dire que les sujets devaient indiquer, parmi deux mots présentés, lequel avait été présenté le plus tardivement dans la liste. Le taux de bonnes réponses était plus élevé lorsque les deux mots de la paire appartenaient à la même catégorie sémantique (e.g., *fraise* et *pomme*) que lorsqu'ils étaient d'une catégorie sémantique différente (e.g., *chien* et *voiture*). Ces résultats peuvent s'interpréter de la façon suivante : au sein de la liste, un mot d'une certaine catégorie sémantique, disons par exemple « *pomme* » devrait automatiquement rappeler au sujet qu'il a vu auparavant un autre mot de la même catégorie, disons « *fraise* ». Cette remémoration au moment de la présentation du mot « *fraise* » est un événement lui-même stocké en mémoire. Ainsi, lorsque la paire « *pomme* - *fraise* » apparaît au moment du test final, le sujet devrait se souvenir que le second mot lui avait rappelé le premier et il devrait pouvoir choisir avec succès le mot le plus récent. À l'inverse, deux mots n'appartenant pas à la même catégorie sémantique, disons « *chien* » et « *voiture* », ne devraient pas bénéficier de cette récupération (ou de façon moins importante) car il n'y a pas de raison particulière pour que, lors de la présentation du mot « *voiture* », le sujet se souvienne qu'il a vu auparavant le mot « *chien* ». Des effets similaires ont été rapportés par Winograd et Soloway (1985), Hintzman, Summers et Block (1975) et Hintzman, Block et Summers (1973).

Concernant plus spécifiquement l'effet de pratique distribuée, nous avons déjà mentionné que, très tôt dans la littérature, certains auteurs (e.g., Martin, 1968 ; Melton, 1967) avaient remarqué qu'il était utile au rappel final que, au moment de P2, le sujet reconnaisse l'item comme étant une répétition. Par exemple, Melton faisait varier l'IIR entre les différentes occurrences d'un mot, et de plus, faisait réaliser aux sujets une tâche de reconnaissance lors de la phase d'apprentissage : pour chaque mot présenté le sujet devait annoncer s'il pensait que le mot était une répétition ou un nouveau mot. Lors du rappel libre ultérieur, les mots qui avaient été reconnus comme étant répétés étaient effectivement mieux rappelés que les mots non reconnus. Cependant, cet aspect de reconnaissance pendant l'étude n'avait jamais été réellement approfondi ou explicité jusqu'aux articles de Thios et D'Agostino (1976) et Johnston et Uhl (1976), que nous décrirons plus tard.

Il est important de noter que le terme de *récupération en phase d'étude* tel qu'on le retrouve dans les articles récents (e.g., Verkoijen, Rikers, & Schmidt, 2004 ; Logan et Balota, 2008) peut référer à deux conceptions différentes. La première conception est

employée dans des situations d'apprentissage où une récupération active du matériel est exigée de la part du sujet, c'est-à-dire qu'un ou plusieurs tests sont employés comme essais d'apprentissage (i.e., conditions *Test seul* ou *Test-étude*, telles que nous les avons définies dans le chapitre 1) comme dans Thios et D'Agostino (1976). La seconde conception, davantage liée au phénomène de réminiscence ou remémoration (*reminding*) décrit dans Tzeng et Cotton (1980), et dans Hintzman et Block (1973) par exemple, est employée dans les situations où une simple répétition (i.e., condition *Etude seule*) déclenche chez le sujet une récupération automatique de souvenirs, qui se manifeste par un sentiment de reconnaissance. Nous décrirons d'abord les hypothèses s'appliquant aux conditions de récupération active, puis nous décrirons ensuite les hypothèses s'appliquant à la récupération automatique, qui nous concernent plus directement dans le présent travail.

2. 2. 1. Première acception : récupération active

Cette première conception de la théorie de la récupération en phase d'étude correspond à une situation expérimentale où les essais d'apprentissage sont des tests de récupération, suivis ou non d'un feedback (respectivement conditions *Test seul* ou *Test étude*). Selon cette conception, une récupération réussie du matériel pendant un test ayant lieu pendant la phase d'apprentissage est utile à la récupération finale. Thios et D'Agostino (1976) ont voulu tester directement cette hypothèse, qu'ils considéraient opposée à celle de la variabilité de l'encodage. Selon les auteurs, une fonction importante du second essai est de déclencher la récupération active de la représentation stockée pendant le premier essai, et ce sont les phénomènes associés à cette récupération qui pourraient être la cause des effets d'espacement et d'intervalle observés dans le rappel ultérieur. Leurs hypothèses étaient que les effets d'espacement et d'intervalle observés dans des conditions qui encouragent la récupération active en phase d'étude devraient être substantiellement réduits lorsque la récupération est empêchée. Les Expériences 1 et 2 se déroulaient de la façon suivante. Des phrases étaient présentées et le sujet devait les apprendre (e.g., *The conductor boarded the express train*), puis il y avait un intervalle de durée variable pendant lequel le sujet réalisait une tâche distractive. Le second essai d'apprentissage consistait pour le sujet à fournir la phrase vue auparavant mais en la transformant pour qu'elle soit à la voix passive (e.g., *The express train was boarded by the conductor*). Ce deuxième essai appartenait à l'une ou l'autre des deux conditions suivantes. Dans la condition de *récupération*, seul l'objet de la phrase était donné comme indice, et le sujet devait se souvenir du reste de la phrase pour la transformer et la donner à la voix passive (e.g., *Express train*). Dans la condition de *non récupération*, l'objet de la phrase était également donné mais était accompagné de l'intégralité de la phrase déjà vue, ainsi le sujet n'avait pas à récupérer la phrase en mémoire mais pouvait faire la transformation directement à l'aide de la phrase fournie (e.g., *Express train - The conductor boarded the express train*). Enfin, immédiatement après la fin de la phase d'apprentissage, un rappel libre des phrases avait lieu. Les trois résultats suivants ont été observés : 1) En condition de *non récupération*, l'effet d'intervalle n'était pas observé : tous les intervalles permettaient le même rappel final. 2) En condition de *récupération*, la probabilité de récupération en phase d'étude, c'est à dire la capacité à restituer la phrase à la voix passive lors de P2, diminuait au fur et mesure que l'intervalle augmentait. Cela est cohérent avec la notion d'oubli lié au passage du temps. 3) En condition de *récupération*, la probabilité de rappel final correct de la phrase augmentait avec l'intervalle entre les deux essais, mais seulement lorsque la récupération avait pu être correctement réalisée par le sujet. Il s'agit de l'effet d'intervalle classiquement observé. 4) En condition de *récupération*, mais lorsque le sujet ne parvenait pas à donner la phrase à la voix passive, ce qui correspond à un échec à la tâche de récupération, l'effet d'intervalle était moins franc (cependant

les conclusions sont limitées par le petit nombre d'observations dans ce cas). Ainsi, ces résultats montrent que la récupération de l'information lors des répétitions a une influence essentielle sur le rappel final et sur l'effet d'intervalle.

Les auteurs ont proposé l'explication suivante. Partant du postulat que la quantité de traitement (ou effort de récupération) de la seconde occurrence est positivement corrélée à la performance de rappel final, alors cette quantité de traitement doit être la plus élevée possible. Or, plus l'intervalle P1-P2 est long, plus la récupération en phase d'étude est difficile, étant donné que l'oubli du matériel se produit progressivement pendant cet intervalle. Ainsi, plus l'IIR est long, plus le traitement réalisé au moment de P2 est important et, par conséquent, meilleur sera le rappel final. (Ceci n'est cependant valable que dans la mesure où l'information est effectivement récupérée en mémoire à P2). Or, cet effort de récupération augmente avec l'IIR. Bjork (1975, cité dans Roediger & Karpicke, 2006) a postulé que la profondeur de récupération pourrait opérer de façon similaire à la profondeur de traitement lors de l'encodage, c'est-à-dire qu'une récupération nécessitant plus d'effort serait un événement potentiellement plus utile à la rétention finale qu'une récupération dite superficielle. Cependant il est nécessaire que, pour jouer ce rôle, la récupération soit réussie, ou alors qu'un feedback soit donné au sujet en cas d'échec. Cette hypothèse d'effort de récupération présente des points communs avec celle de la reconstruction (voir la section 2.4).

Une autre interprétation, proposée par d'autres auteurs, est la suivante. Plus l'intervalle P1-P2 est long, plus la récupération exigée lors de P2 est similaire à la récupération exigée lors du test final, et donc plus elle favorise ce dernier car elle constituerait alors un « entraînement » pour le rappel final. Cet argument est à rapprocher de l'hypothèse du *transfer appropriate processing* qui énonce que la performance à une tâche de mémoire est d'autant plus élevée que les processus requis pour la réaliser correspondent aux opérations d'encodage engagés pendant l'apprentissage (Morris, Bransford, & Franks, 1977). En effet, un test réalisé lors d'un épisode P2 peu de temps après l'épisode P1 pourrait nécessiter des processus différents de ceux requis lors d'un test de récupération finale (e.g., une récupération d'information à partir de la mémoire à court terme). À l'opposé, plus l'intervalle P1-P2 augmente, plus les processus de récupération engagés à P2 sont proches de ceux requis lors du test final, qui a lieu après un DR d'une certaine durée.

Ces hypothèses sont similaires à celles avancées par les auteurs des études récentes sur l'effet de l'agencement temporel des essais à court terme (Karpicke & Roediger, 2007 ; Logan & Balota, 2008) et qui mettent en évidence, comme nous l'avons vu dans la section 1.4.1, que la durée séparant la présentation initiale et le premier test serait un facteur essentiel à la rétention finale.

Les expériences de Thios et d'Agostino (1976) avaient pour mérite de constituer un test direct de la théorie de la récupération en phase d'étude. Cependant, les processus mis en jeu lorsque, à l'apprentissage, on exige du sujet une récupération active ne sont pas représentatifs des processus mis en jeu lorsqu'un item est présenté à nouveau. Le titre de l'article de Glenberg et Smith (1981) illustre bien cette différence : « *Espacer les répétitions et résoudre des problèmes ne sont pas la même chose* ». En d'autres termes, la récupération ici est dite « active ». Plus classiquement, la théorie de la récupération en phase d'étude fait appel à la notion de récupération automatique, liée à la reconnaissance. Nous nous focaliserons davantage sur cette seconde conception.

2. 1. 2. Seconde acception : récupération automatique déclenchée par la répétition

C'est cette formulation de la théorie de la récupération en phase d'étude qui est le plus souvent mentionnée dans la littérature et c'est cette conception que nous évoquerons lorsque nous citerons la théorie de la récupération en phase d'étude par la suite. Dans de nombreux articles (e.g., Cepeda et al., 2006) l'hypothèse de la récupération en phase d'étude est présentée comme s'opposant à celle de la variabilité de l'encodage. En effet, la théorie de la variabilité de l'encodage devrait logiquement prédire que si l'occurrence de l'item n'est pas reconnue comme une répétition au moment de P2, alors la variabilité de l'encodage devrait être la plus importante, ce qui devrait favoriser le rappel libre ultérieur (e.g., Johnston & Uhl, 1976). Or, toutes les données empiriques ont contredit cette prédiction et ont même montré l'inverse, c'est-à-dire qu'une reconnaissance de l'item au moment de P2 était favorable à la récupération ultérieure. Nous allons examiner quelques-unes de ces données.

Johnston et Uhl (1976, Expérience 2) présentaient une liste de mots et, pour chaque mot présenté, le sujet devait réaliser une tâche de reconnaissance c'est-à-dire estimer si le mot avait déjà été présenté auparavant ou non. Certains mots étaient répétés, et ce avec des IIR variables, et d'autres n'étaient pas répétés. Ensuite une tâche de rappel libre avait lieu. De façon concordante avec les observations de Thios et d'Agostino (1976) et de Melton (1967), les items reconnus à P2 avaient beaucoup plus de chances d'être rappelés que les items non reconnus. D'autre part, alors que la probabilité de reconnaître un mot comme répété lors de P2 diminuait avec l'IIR, la probabilité de finalement rappeler un mot préalablement reconnu augmentait avec l'IIR. Autrement dit, l'intervalle entre P1 et P2 bénéficiait plus aux mots reconnus à P2 qu'aux mots non reconnus. Par conséquent, soit la théorie de la variabilité de l'encodage est fausse, soit il faut lui ajouter la condition de reconnaissance à P2. C'est ce qu'ont fait les modèles dits « mixtes », c'est-à-dire qui combinent variabilité de l'encodage et récupération en phase d'étude, que nous décrirons dans la section 2.3.

D'autres arguments en faveur de la théorie de la récupération en phase d'étude existent. Premièrement, lorsque les items ne sont pas reconnus à P2 comme étant une répétition, alors l'effet d'espacement n'est pas observé (e.g., Braun & Rubin, 1998). Deuxièmement, la théorie de la récupération en phase d'étude permet d'expliquer assez simplement pourquoi la courbe de performance en fonction de l'IIR prend parfois une forme en U inversé (e.g., Glenberg, 1976). Selon la théorie, lorsqu'une certaine durée de l'intervalle P1-P2 est atteinte, la récupération en phase d'étude est de moins en moins probable, et par conséquent les performances finales chutent. C'est le parti pris défendu par Verkoijen et al. (2005) et Toppino et Bloom (2002). Troisièmement, cette hypothèse permet d'expliquer pourquoi les effets d'espacement et d'intervalle ne concernent pas des paires d'items différents. En effet, Ross et Landauer (1978) ont constaté que beaucoup de théories postulent que des traces mnésiques distinctes sont créées pour des événements distincts (e.g., Hintzman & Block, 1971), même si ces événements concernent un même item qui se répète. Ainsi, pour que le rappel final de l'item soit couronné de succès, il faut que le sujet récupère alors soit la trace relative à P1, soit la trace relative à P2, soit les deux. Formulée de façon mathématique, la probabilité de rappel final est la suivante :

$$\text{Proba (Rappel)} = \text{Proba (Rappel P1)} + \text{Proba (Rappel P2)} - \text{Proba (Rappel P1 et P2)}.$$

Il se trouve que dans le cas d'un stimulus répété, par exemple le mot *maison* présenté deux fois, alors cette probabilité de rappel est différente selon l'intervalle qui sépare P1 et P2. Si la répétition d'un item n'est pas une situation spéciale relativement aux traces mnésiques créées en mémoire (i.e., une nouvelle trace est créée lors de la répétition), alors on devrait observer les mêmes effets pour des événements P1 et P2 qui concernent des

items différents, par exemple le mot *banane* puis le mot *chat* présentés dans la même liste. Or, cela n'a jamais été observé. Par conséquent, les processus en jeu lorsqu'un item est répété sont différents de ceux en jeu lorsque deux items différents sont présentés : quelque chose se produit lors de la répétition d'un item qui ne se produit pas lors de la présentation d'un item différent. Manifestement, il s'agit de la récupération en phase d'étude.

Ainsi, un ensemble d'éléments expérimentaux soulignent l'importance des phénomènes de récupération en phase d'étude dans l'explication des effets de pratique distribuée. Nous allons maintenant développer des hypothèses récentes qui développent ces hypothèses, en particulier celle de la remémoration récursive.

2. 1. 3. Remémoration récursive

En 2004, Hintzman proposa la notion de remémoration récursive (*recursive reminding*) pour expliquer pourquoi les réponses des sujets dans les tâches de jugement de fréquence et de jugement de confiance en reconnaissance ne peuvent pas s'expliquer par un facteur unique. Hintzman défendait l'idée qu'il faut considérer comment les processus mentaux engagés par la seconde étude d'un item diffèrent de ceux engagés lors de la première présentation. Un candidat évident est justement la remémoration (*reminding*). La remémoration est une *recollection* qui se produit durant un essai d'étude, c'est-à-dire une réminiscence spontanée d'événements reliés au stimulus, particulièrement des événements antérieurs ayant eu lieu au sein de l'expérience. Cette recollection contiendrait l'item, le traitement réalisé sur l'item et les éléments contextuels associés à l'occurrence de l'item ; cette remémoration serait, de plus, associée à une expérience subjective chez le sujet. L'auteur postulait que la seconde présentation d'un item n'est pas seulement encodée en tant que telle, mais que, comme elle rappelle aussi au sujet la première présentation, alors l'ensemble de ces événements cognitifs (i.e., recollection de P1 + nouveau traitement de l'item) est encodé dans la nouvelle trace. Si une troisième présentation a lieu, elle rappelle au sujet la deuxième présentation, qui contenait elle-même le souvenir de la recollection de P1, et tous ces événements cognitifs sont encodés dans la trace (i.e., recollection de P2 qui contient la recollection de P1 + nouveau traitement de l'item). Ce processus peut se poursuivre pour chaque nouvelle présentation. Les recollections seraient donc enchâssées les unes dans les autres, c'est-à-dire que le sujet ne se souviendrait pas d'événements distincts correspondant à P1, P2, P3 etc ; plutôt, l'item évoquerait la réminiscence de sa dernière occurrence, qui elle-même avait déclenché la réminiscence de l'occurrence précédente, qui elle-même avait déclenché la réminiscence de l'occurrence précédente, etc.

Ainsi, en 2010, Hintzman postulait que la théorie de la remémoration récursive pouvait être une bonne alternative aux modèles explicatifs classiques de l'effet de répétition, qui sont les modèles basés sur les notions de trace cumulative d'une part, et de traces indépendantes d'autre part. Il a observé que la répétition peut faire paraître un item plus récent qu'il ne l'est réellement, mais aussi moins récent qu'il ne l'est réellement, et que cela dépend de l'espacement des répétitions. D'autre part, les sujets peuvent juger de la récence de deux présentations d'un item avec un certain degré d'indépendance, particulièrement quand ils se souviennent correctement que l'item a été présenté deux fois.

Benjamin et Tullis (2010 ; voir aussi Benjamin & Ross, *in press*) aboutissaient à une conclusion similaire. Dans le cadre de l'effet de pratique distribuée, ils émettaient des réserves au sujet des théories de la variabilité de l'encodage et proposaient à la place un nouveau modèle. Ce modèle consistait à apporter aux modèles mathématiques existants tels que SAM (Raaijmakers, 2003) ou MINERVA 2 (Hintzman, 1988) des améliorations incluant la notion de remémoration récursive, afin de mieux rendre compte

des effets de répétition sur la rétention ultérieure. Les auteurs énonçaient trois principes qui supplémentent les principes de base de la récupération en phase d'étude :

- Les items présentés à P1 sont sujets à l'oubli avec le temps, selon une fonction de type puissance.
- Les items présentés à P2 varient dans leur capacité à déclencher spontanément la remémoration de l'épisode ayant eu lieu à P1. Les répétitions à l'identique sont censées être associées à de hauts potentiels de remémoration, la présentation d'items liés entre eux à des potentiels moyens, et la présentation d'items non liés entre eux à de très faibles potentiels.
- L'action de récupération potentialise la mémoire, et le degré de potentialisation est positivement relié à la difficulté de récupération. C'est-à-dire que la remémoration suite à un haut degré d'oubli favoriserait davantage la mémoire qu'une remémoration suite à un faible oubli. Ainsi, le degré de difficulté de la remémoration entre en ligne de compte également, ce qui rejoint les conceptions des théories de l'accessibilité et de la reconstruction développées dans la section 2.4.

La notion de remémoration récursive n'est pas, selon les auteurs, encore assez aboutie pour être considérée comme un modèle. Il s'agit cependant d'une piste prometteuse à explorer dans le futur. La nouveauté de cette perspective, par rapport à la théorie de la récupération en phase d'étude -et bien que basée sur celle-ci- est que le traitement réalisé sur un item répété est différent selon la recollection qu'elle déclenche chez le sujet. Par exemple, si la deuxième occurrence du mot *pomme* fait se rappeler au sujet l'épisode où il a vu *pomme*, alors le traitement qu'il va réaliser (i.e., stratégie d'apprentissage, éléments évoqués, métacognition, etc...) sera différent du cas où la répétition ne lui rappelle rien.

En guise de synthèse, un des intérêts des hypothèses de la récupération en phase d'étude et de la remémoration récursive, est que la récupération de la trace mnésique a pour conséquence de faire le lien, en mémoire, entre les deux (ou plus) occurrences d'un même item. Cependant, à elle seule, elle explique difficilement les effets d'intervalle ou d'espacement ; elle peut simplement expliquer la chute des performances pour de longs IIR. Il semble implicite, dans la littérature, que cette récupération est corrélée au sentiment de reconnaissance de l'item, mais aucune discussion sur cette idée n'apparaît dans les articles, sauf chez Hintzman (2004, 2010) qui parle d'une expérience subjective chez le sujet.

Nous allons maintenant développer les modèles théoriques qui combinent la théorie de la variabilité de l'encodage et celle de la récupération en phase d'étude, que nous appelons modèles mixtes.

2. 3. Modèles mixtes combinant variabilité de l'encodage et récupération en phase d'étude

Quelques auteurs ont proposé des explications de l'effet de pratique distribuée qui combinent les hypothèses de variabilité de l'encodage et de récupération en phase d'étude. Ce courant n'a cependant pas été très suivi par les auteurs actuels puisque dans la majorité des articles récents (e.g., Cepeda et al., 2006), ces deux théories sont mises en opposition. Nous allons développer dans cette section les études apportant des arguments expérimentaux en faveur de ces modèles appelés ici « mixtes ».

Nous avons vu précédemment (section 2.1.4) que les expériences dans lesquelles on faisait varier les contextes d'encodage à P1 et à P2 échouaient dans bon nombre de cas à améliorer le rappel final (e.g., Postman & Knecht, 1983). Les études de Bellezza et collègues

(Bellezza, Winkler, & Andrasik, 1975 ; Bellezza & Young, 1989 ; Young & Bellezza, 1982) ont tenté d'apporter des explications à ces échecs sans toutefois abandonner l'hypothèse de la variabilité de l'encodage. Young et Bellezza (1982) ont montré, tout comme Postman et Knecht (1983), que certaines situations dans lesquelles l'encodage était constant entre deux occurrences d'un même item menaient à de meilleures performances finales que lorsque l'encodage était variable. En effet, dans 3 expériences sur 4 réalisées, la condition de constance d'encodage était plus favorable au rappel final que la condition de variabilité d'encodage. Les auteurs concluaient de ces résultats que, dans certains cas, il est plus utile de maintenir à P2 la « structure organisationnelle » créée lors de P1 que d'augmenter le nombre de composants associés à l'item. Selon les auteurs, la performance de rappel final serait améliorée seulement si la variabilité de l'encodage entraîne l'addition d'information au même locus fonctionnel, c'est-à-dire à la même trace mnésique ; ce ne serait pas le cas lorsqu'elle induit la création de traces fonctionnelles séparées. Pour réaliser cette addition d'information, il est nécessaire, au moment de la répétition, de récupérer la trace mnésique correspondant à la première occurrence de l'item, ce qui correspond à la récupération en phase d'étude.

Or, dans les expériences présentées dans l'article, il est possible que cette condition de création d'une trace unique n'ait pas été remplie. Il est possible que, dans les conditions d'encodage variable, la seconde présentation ait entraîné la formation d'une nouvelle trace mnésique, tandis que, dans les conditions d'encodage constant la seconde présentation ait entraîné la récupération de la trace mnésique existante. Dans leur article de 1989, les auteurs ont proposé *l'hypothèse du chunking* (Bellezza & Young, 1989). Cette hypothèse postule que lorsque la seconde présentation d'un item a lieu, le sujet peut reconnaître cet item comme étant une répétition. Dans ce cas-là, la trace mnésique de la première présentation est récupérée et l'information de la deuxième occurrence est introduite dans cette trace déjà existante ; il y a *chunking* des deux contextes au sein d'une même trace. Au contraire, si le sujet ne reconnaît pas l'item comme étant une répétition, alors une nouvelle trace mnésique est créée indépendamment de la première. La formation d'une trace mnésique unique et élaborée entraîne un rappel optimal. Ainsi, selon cette théorie, le rappel optimal se produirait lorsque les deux conditions suivantes sont remplies : 1) la trace mnésique formée à P1 est récupérée durant P2 ; et 2) le contexte lors de P2 est assez différent de celui de P1 pour qu'une nouvelle information soit additionnée à la trace unique. Selon cette hypothèse, l'amplitude de variabilité de l'encodage est cruciale pour déterminer la performance de rappel ultérieur. Elle ne doit être ni trop faible, ni trop importante. Si la variabilité de l'encodage est trop importante (comme par exemple lorsqu'on exige du sujet une tâche d'orientation différente lors des deux occurrences), elle va empêcher, au moment de P2, la reconnaissance de l'item et la récupération de la trace. Si la variabilité de l'encodage est trop faible (comme par exemple lorsque l'item est répété à l'identique peu de temps après la première occurrence), elle ne garantit pas que la trace récupérée bénéficiera d'une élaboration. Greene et Stillwell (1995) ont répliqué les observations de Bellezza et collègues dans une tâche de jugement de fréquence.

Grâce à ces conceptions, il devient possible d'expliquer pourquoi certaines expériences ont montré un avantage de la constance d'encodage et d'autres un avantage de la variabilité d'encodage. Les deux types de résultats dépendaient du fait que P2 soit reconnu ou non comme une répétition de P1. En effet, la probabilité de ne pas récupérer la trace de P1 au moment de P2 est plus importante en condition de variabilité d'encodage. Ainsi, dans les expériences, la constance d'encodage était supérieure à la variabilité d'encodage dans le cas où la condition de variabilité d'encodage empêchait la récupération en phase d'étude. A

contrario, la variabilité d'encodage était supérieure à la constance d'encodage dans le cas où la condition de variabilité d'encodage permettait la récupération en phase d'étude.

Verkoijen et al. (2004) ont également conclu à la pertinence d'un modèle combinant variabilité de l'encodage et récupération en phase d'étude. Dans leurs deux expériences, les participants devaient apprendre une liste contenant des mots présentés en condition massée ou distribuée. Ces mots étaient répétés soit dans le même contexte soit dans un contexte différent, le contexte étant opérationnalisé par l'arrière-plan des items. Le composant variabilité de l'encodage du modèle prédit qu'une variation du contexte devrait avoir un effet bénéfique sur le rappel libre des répétitions massées. A contrario, le composant récupération en phase d'étude du modèle prédit qu'une variation du contexte devrait avoir un effet délétère sur le rappel libre des répétitions espacées. En effet, dans cette condition, la récupération en phase d'étude n'est pas assurée. Les résultats des deux expériences étaient consistants avec ces prédictions car il a été observé une interaction croisée entre l'espacement et la congruence du contexte lors des deux occurrences : les répétitions massées étaient mieux rappelées quand présentées dans différents contextes, alors que les présentations espacées étaient mieux rappelées quand répétées dans un même contexte. Ces résultats suggèrent donc qu'une explication adéquate de l'effet d'espacement en rappel indicé devrait incorporer les mécanismes de variabilité de l'encodage et de récupération en phase d'étude. Ainsi, il semble que l'effet facilitateur dû à l'espacement sur le rappel libre s'observera uniquement si le souvenir de la première occurrence de l'item est récupéré lors de la seconde occurrence.

Le modèle de Raaijmakers (2003) ainsi que celui de Mozer et al. (2009) implémentent des modèles de type mixte. Ces modèles seront développés dans la section 3.2.

Nous avons vu que les théories de la variabilité de l'encodage et de la récupération en phase d'étude ont toutes deux des avantages ainsi que des limites ; ainsi, au lieu de les opposer, combiner ces deux hypothèses pourrait permettre de mieux expliquer les effets de pratique distribuée. Nous allons maintenant nous intéresser à des hypothèses moins couramment évoquées dans la littérature.

2. 4. Traitement déficitaire, amorçage, accessibilité et reconstruction

Un troisième type d'hypothèses postule qu'un *traitement déficitaire* des occurrences répétées en condition d'IIR court est responsable de l'effet de pratique distribuée. Ces hypothèses ont été formulées pour expliquer les effets d'espacement et d'intervalle à court terme, et sont probablement moins pertinentes pour expliquer les effets de pratique distribuée à long terme. Il est toutefois utile de rappeler brièvement le principe de ces hypothèses dans la mesure où, comme nous le verrons, des expériences récentes en imagerie fonctionnelle font appel à des concepts proches de ces théories.

Ces hypothèses ont pour point commun l'idée que la seconde occurrence d'un item est traitée de façon moins optimale en condition massée qu'en condition distribuée, ce qui a pour conséquence de produire un encodage de moindre qualité. Ainsi, la proximité temporelle des deux occurrences empêcherait le sujet de réaliser un second encodage optimal pour la mémorisation. Il a été montré que ces effets se produisent de façon automatique et indépendamment de la volonté du sujet (e.g., Shaughnessy, 1976). Bien que des hypothèses faisant appel à des processus volontaires de la part du sujet (e.g., *rehearsal* et *inattention*) ont aussi été proposées, les arguments expérimentaux ont globalement contredit ce caractère volontaire (e.g., Toppino, 1991 ; voir cependant Toppino, Fearnow-

Kenney, Kiepert, & Teremula, 2009). Plusieurs hypothèses ont été proposées pour décrire les mécanismes en présence.

En 1993, Challis a défendu l'idée que *l'amorçage sémantique* pouvait être la cause du traitement déficitaire : lors de la seconde occurrence de l'item en condition massée, celui-ci serait l'objet d'un amorçage de type sémantique, qui empêcherait le sujet de traiter correctement l'item. Le fait de présenter la première fois le mot *chapeau* constituerait un événement amorce, qui faciliterait le traitement de la seconde occurrence du mot *chapeau*. Cette facilitation constituerait un désavantage pour le traitement de la seconde occurrence dans la mesure où il « court-circuiterait » en quelque sorte les processus cognitifs engagés normalement lors de la présentation de l'item. Comme l'amorçage sémantique n'a qu'un effet temporaire, alors en condition distribuée, puisque l'IIR est plus long, un traitement normal de la seconde occurrence pourrait se produire.

En 1998, Russo, Parkin, Taylor, et Wilks ont étendu cette hypothèse en introduisant comme facteur le type de traitement engagé par le sujet lors de la phase d'apprentissage. Lorsqu'un traitement sémantique des informations est requis (par exemple, juger le caractère plaisant, ou apprendre volontairement les mots), alors un amorçage sémantique se produirait lors de P2, ce qui aurait une influence délétère sur la condition massée. C'est l'hypothèse initiale de Challis. Cependant, lorsqu'un traitement non sémantique des informations est requis (par exemple, juger la rime, détecter la présence de lettres dans un mot), alors un amorçage non sémantique (i.e., perceptif) se produirait lors de P2, ce qui aurait également un effet délétère sur la condition massée. De façon cohérente avec cette hypothèse, il a été observé que des modifications physiques d'un item d'une occurrence à l'autre, comme par exemple un changement de police d'écriture pour les mots (e.g., Mammarella, Avons, & Russo, 2004 ; Russo, Mammarella, & Avons, 2002) ou un changement d'orientation pour les visages (e.g., Mammarella, Russo, & Avons, 2002) ont pu éliminer l'effet d'espacement. Autrement dit, la condition massée permettait une récupération ultérieure aussi élevée que la condition distribuée. Or, dans ce cas, un même item était répété mais l'amorçage perceptif était empêché du fait de la modification des caractéristiques visuelles de l'item. Ces observations montrent donc qu'en situation de répétition d'un item à l'identique, un phénomène se produit, qui ne se produit pas lorsque la répétition prend une forme différente.

Ces hypothèses peuvent être rapprochées de celle avancée par Jacoby (1978). Selon l'auteur, lors de P2, les processus d'encodage ayant été réalisés pendant P1 sont plus ou moins récupérés en mémoire. Si ces processus ne peuvent pas être récupérés tels quels, alors il est nécessaire de les reconstruire, c'est-à-dire de les réaliser à nouveau, et cette reconstruction serait bénéfique à la mémoire. L'auteur utilisait une tâche dans laquelle des paires de mots étaient présentées en intégralité lors de P1 (e.g., *foot-shoe*), puis, lors de P2, le mot cible était incomplet et le sujet devait le retrouver (e.g., *foot-s--e*). Dans un troisième temps, un rappel indicé était demandé aux sujets (e.g., *foot- ?*). Si l'intervalle entre P1 et P2 est court, alors il est probable que l'on puisse facilement récupérer en mémoire le mot cible, puisque la paire a été présentée récemment. En revanche, si P2 se produit longtemps après P1, il va être nécessaire pour le sujet de « résoudre » le problème du mot incomplet, et donc de *reconstruire* certains des processus engagés lors de P1. C'est pourquoi, en général, l'espacement serait bénéfique pour la mémoire. En 1982, Cuddy et Jacoby mirent en avant la notion d'accessibilité : plus la trace de P1 est difficilement accessible en mémoire lors de P2, ce qui se produit suite à un long IIR, plus les traitements engagés lors de P2 seront importants, ce qui est favorable à la rétention.

Cependant, Glenberg et Smith (1981) ont avancé des arguments contraires à cette hypothèse de *reconstruction*. En particulier, bien que l'hypothèse semble cohérente lorsque des essais de *Test-étude* ou *Test-seul* sont employés, ils considéraient que lorsque des essais de type *Etude seule* sont employés, l'effet d'espacement ne peut pas être expliqué seulement par une telle hypothèse. Selon les auteurs, les résultats expérimentaux ne montraient pas que, dans des conditions classiques d'apprentissage de liste (i.e., avec des essais de type *Etude seule*), les sujets s'engagent davantage dans des activités constructives lorsque l'IIR est long que lorsqu'il est court. Premièrement, l'effet d'intervalle n'est pas toujours monotone, ce qui devrait pourtant être observé si l'hypothèse de reconstruction était exacte (e.g., Glenberg, 1976). Deuxièmement, si l'item est présenté dans une modalité différente lors de P1 et de P2, alors le phénomène de reconstruction devrait nécessairement avoir lieu, et ce, quel que soit l'IIR. Donc l'effet d'espacement devrait disparaître en condition de modalité différente, ce qui n'est pas le cas (e.g., Hintzman, Block, & Summers, 1973). Troisièmement, si une tâche d'orientation différente est réalisée lors de P1 et de P2, alors le phénomène de reconstruction devait nécessairement être réalisé lors de P2, et ce, quel que soit l'IIR. Donc l'effet d'espacement devrait disparaître en condition de tâche d'orientation différente. Or, les expériences de Glenberg et Smith ainsi que des expériences antérieures (e.g., Maskarinec & Thompson, 1976) n'ont pas montré cette disparition de l'effet d'espacement annoncée, ce qui s'oppose à la théorie de la reconstruction.

Ainsi, cette théorie de la reconstruction semble être pertinente surtout dans le cas où les répétitions sont des tests et non pas de simples présentations. D'autre part, elle concerne davantage la différence entre les conditions massée et distribuée que l'effet d'intervalle. De plus, cette théorie émet des prédictions opposées à celles de la théorie de la récupération en phase d'étude : plus l'IIR est grand, plus la reconstruction est importante et donc plus la rétention devrait être importante.

Une autre hypothèse, très proche de la précédente, intitulée hypothèse de *l'accessibilité*, a été développée en 2005 par Pavlik et Anderson dans un modèle mathématique. Ce modèle est décrit dans la section 3.2.2. Pour résumer, le modèle implémente le principe suivant : le taux de déclin de la trace créée lors d'une répétition d'un item est lié au niveau d'activation de la trace laissée par l'occurrence précédente. Ainsi, si la trace de l'item est fortement active en mémoire au moment de la répétition, ce qui se produit après un IIR court, alors le taux de déclin de l'activation de la nouvelle trace sera élevé, et par conséquent, l'oubli rapide. À l'inverse, si la trace de l'item est faiblement active au moment de l'occurrence, ce qui se produit après un IIR long, alors le taux de déclin de l'activation de la nouvelle trace sera faible, et par conséquent, l'oubli faible. Comme nous le verrons dans le chapitre consacré à la modélisation, les simulations réalisées avec ce modèle étaient relativement convaincantes.

Les hypothèses de l'accessibilité et de la reconstruction sont difficiles à tester sans modifier le paradigme classique des études sur les effets de pratique distribuée, paradigme qui met en jeu de simples présentations, c'est-à-dire des essais de type *Etude seule*, lors de l'apprentissage. En effet, il est difficile d'accéder aux processus de reconstruction à l'œuvre chez le sujet sans inclure des tests pendant l'apprentissage. C'est pourquoi les études utilisant l'imagerie cérébrale pourraient être utiles car elles pourraient permettre d'obtenir des informations sur ce qui se produit chez le sujet sans modifier la tâche réalisée par ce dernier. Or, il apparaît que la notion de suppression de répétition (voir la section 3.1) pourrait être mise en lien avec la notion d'amorçage et la théorie de l'accessibilité-reconstruction. Quand l'IIR est court, les zones cérébrales activées lors de P2 ne sont pas autant activées

que lors de P1, alors que lorsque l'IIR est long, cette sous-activation n'est pas observée. Ces niveaux d'activation cérébrale pourraient justement refléter les opérations d'encodage dont parlait Jacoby (1978). Cependant, cette hypothèse soulève des questions d'une portée plus large que la problématique de l'effet d'espacement. On peut s'interroger sur la définition exacte de la notion de construction de processus, et sur celle de récupération de processus. Plus largement, ces questions sont liées à la connaissance des mécanismes en jeu lors de la reconnaissance. Ces questions vont donc bien au-delà du cadre du présent travail.

2. 5. Explications métacognitives

Pour conclure sur les théories psychologiques de l'effet de pratique distribuée, nous trouvons pertinent de présenter les hypothèses liées au comportement volontaire d'apprentissage réalisé par le sujet. En effet, jusqu'à présent, nous n'avons considéré que des théories qui font abstraction du sujet en situation d'apprentissage, de ses choix, de ses motivations, et de ses stratégies. Il est cependant évident que les choix du sujet -notamment dans des tâches où le celui-ci doit apprendre du matériel dans le but de s'en souvenir plus tard- en termes d'allocation de ressources cognitives, de stratégies d'apprentissage, d'effort engagé ou non à la tâche, et tout ce qui constitue ce qu'on appelle la métacognition, peuvent entrer en ligne de compte dans les effets mnésiques. Ces phénomènes pourraient s'ajouter à ceux qui échappent à la volonté et à la conscience du sujet, comme la variabilité de l'encodage, la récupération en phase d'étude, etc. Ces considérations sont particulièrement pertinentes dans des situations relativement écologiques telles qu'un apprentissage sur plusieurs jours, qui nécessitent un certain engagement de la part du sujet.

En 2005, Bahrack et Hall ont été parmi les premiers à mettre en avant cette notion de métacognition comme facteur explicatif des effets de pratique distribuée ; ils se sont concentrés en particulier sur les situations dans lesquelles l'apprentissage a lieu sur le long terme, c'est-à-dire sur au moins plusieurs jours. Les auteurs insistent sur le fait qu'il faut prendre en considération les processus cognitifs qui se produisent au moment de l'apprentissage car ils joueraient un rôle causal dans le niveau de rétention ultérieur. Par exemple, les expériences de Bahrack (1979) et Bahrack et al. (1993 ; voir la section 1.3.2) ont montré que de très longs intervalles entre les sessions d'apprentissage sont associés à de faibles performances pendant l'apprentissage, contrairement à des sessions rapprochées. En revanche, de très longs intervalles entre les sessions d'apprentissage sont associés à une meilleure rétention ultérieure de l'information, par rapport à des intervalles plus courts. (Dans les expériences de Bahrack et collaborateurs, les phases d'apprentissage étaient constituées d'essais de type *Test-étude*, ce qui permet de rendre compte des performances du sujet pendant l'apprentissage.) Pour expliquer ces résultats, les auteurs faisaient l'hypothèse suivante : les différentes stratégies engagées par le sujet pour apprendre peuvent avoir une durée d'efficacité variable. Mais on ne peut se rendre compte que la stratégie que l'on utilise n'est pas efficace qu'en faisant l'expérience d'un échec de récupération. Si tous les IIR sont courts le sujet ne fera pas l'expérience de l'échec (i.e., il réussira les tests), et ne changera donc pas de stratégie. À l'inverse, si les IIR sont longs, le sujet fera plus probablement l'expérience que ses stratégies antérieures n'étaient pas efficaces, et sera amené à les perfectionner.

Les expériences réalisées dans l'étude de Bahrack et Hall (2005) ont montré que les sujets étaient plus enclins à changer de stratégie d'apprentissage lorsque les IIR étaient longs que lorsque les IIR étaient courts et que cela était bénéfique pour la rétention ultérieure. En outre, les temps d'étude, décidés par les sujets, étaient également plus

longs pour les items auxquels ils n'avaient pas pu répondre correctement juste auparavant. Ainsi les sujets adaptaient leurs stratégies en fonction des succès ou échecs aux tests. Il est raisonnable de penser que, d'une façon générale, les sujets ajustent leurs stratégies d'apprentissage même lorsque des essais de type *Etude seule* sont employés pendant la phase d'apprentissage. Par exemple, lors de la répétition d'un item, il est possible que les sujets tiennent ce type de discours intérieur : « Je reconnais ce pseudomot, mais je n'aurais pas pu me souvenir de sa traduction » et perfectionnent alors leur apprentissage. De la même façon, si l'IIR est trop long, il est probable que les sujets ne reconnaissent pas l'item répété et qu'ils remettent en jeu la même stratégie d'apprentissage qui s'était révélée inefficace.

Ces considérations mettent donc en avant l'idée qu'il ne faut pas négliger les comportements délibérés du sujet dans l'explication des effets de pratique distribuée.

2. 6. Conclusion

Cette revue des théories psychologiques relatives à l'effet de pratique distribuée révèle la grande complexité des phénomènes en jeu lors des répétitions d'information. De très nombreux facteurs sont, semble-t-il, à prendre en considération, facteurs qui sont complexes et font eux-mêmes l'objet de recherches spécifiques. Citons notamment les effets contextuels, les processus de récupération en mémoire, la notion de trace mnésique, les stratégies mises en œuvre par le sujet, etc... De plus, toutes les théories ne sont pas mutuellement exclusives. En particulier, les hypothèses de variabilité de l'encodage et de récupération en phase d'étude pourraient être combinées dans des modèles mixtes ; par ailleurs, les effets dus aux stratégies élaborées par le sujet lors de l'apprentissage pourraient se surajouter aux effets dus à des processus involontaires. Dans le présent travail, nous nous pencherons en particulier sur la théorie de la récupération en phase d'étude (Expérience 2) et sur celle de la variabilité de l'encodage (Expérience 4).

En outre, d'autres disciplines des sciences cognitives se sont penchées sur les effets de pratique distribuée, et ont apporté leurs propres conceptions et méthodes d'étude à ces questions. Le chapitre 3 se propose de présenter une revue de ces travaux.

Chapitre 3 : Neurosciences cognitive de l'effet de pratique distribuée

In contrast with the strategy-based hypotheses, some researchers have noted the uncommon generality and persistence of the spacing effect and have proposed that the phenomenon reflects the operation of automatic memory processes that may be so fundamental that they are hard-wired into the memory system.

Toppino (1991)

Nous avons présenté dans les deux chapitres précédents une revue des effets de pratique distribuée dans la discipline de la psychologie cognitive. Cet effet a également été abordé par d'autres approches des sciences cognitives. Dans ce chapitre, nous nous consacrerons tout d'abord à décrire les études réalisées en imagerie cérébrale sur ce thème. Nous considérerons ensuite les tentatives de modélisation mathématique qui ont été réalisées pour rendre compte des effets de la répétition et du temps, avant de mettre en avant des

considérations sur les notions de consolidation et de taux d'oubli, et en prenant enfin une perspective évolutionniste.

3. 1. Apport de l'imagerie cérébrale pour l'étude de l'effet de pratique distribuée

Les discussions sur les causes de l'effet de pratique distribuée peuvent bénéficier de nouveaux types d'arguments expérimentaux avec l'utilisation des techniques d'imagerie fonctionnelle. Les techniques d'imagerie cérébrale fonctionnelle, en particulier l'IRM fonctionnelle, permettent d'observer l'activité cérébrale d'un individu en train de réaliser une tâche cognitive. À notre connaissance, aucune étude en imagerie cérébrale ne s'est intéressée à l'effet de la pratique distribuée en comparant plusieurs IIR supérieurs à 24 h, ni sur l'effet de l'agencement temporel, à court ou à long terme. Cependant, étant donné le caractère novateur de ces techniques pour répondre à la problématique plus générale de l'effet de pratique distribuée, et les arguments qu'elles apportent en termes de théorie explicative, nous présenterons dans ce chapitre les résultats de cinq études récentes qui ont porté sur les bases neurales de l'effet d'espacement et de l'effet d'intervalle à court terme. L'étude de Takashima et al. (2007) a également porté sur cette problématique mais, à cause de problèmes méthodologiques (i.e., pas d'égalisation du DR) rendant les conclusions difficilement interprétables, nous choisissons de ne pas la développer ici.

Avant de décrire ces expériences, il est nécessaire de développer quelques résultats issus des études en imagerie fonctionnelle. Tout d'abord, il est possible de mesurer les niveaux d'activation au moment de l'encodage dans certaines régions cérébrales connues comme ayant un rôle potentiel dans l'encodage de nouveaux items, telles que le cortex préfrontal et le lobe médian temporal (en particulier la région para-hippocampique). Il s'est avéré que ces niveaux d'activation étaient plus importants pour les items qui étaient reconnus ultérieurement que pour les items qui n'étaient pas reconnus (e.g., Brewer, Zhao, Desmond, Glover, & Gabrieli, 1998 ; Wagner, Dale, Rosen, & Buckner, 1998). D'autre part, il a été montré que les manipulations de profondeur d'encodage (Otten, Henson, & Rugg, 2001) et de familiarité du matériel étudié (Chee, Westphal, Goh, Graham, & Song, 2003) augmentent l'activité neuronale de ces régions durant l'encodage, et améliorent généralement aussi la rétention ultérieure du matériel.

Il a également été observé en IRM fonctionnelle et en TEP chez l'humain, ainsi que par l'enregistrement de neurones isolés chez l'animal ou les patients implantés, le phénomène dit de suppression de répétition (*repetition suppression*). Il s'agit de l'observation d'une diminution de la réponse hémodynamique ou neuronale dans certaines régions au moment de la répétition d'un stimulus. En effet, on observe un certain niveau d'activation lors de la première présentation du stimulus (P1) puis, lors des répétitions ultérieures du même stimulus (P2, P3, etc), ce niveau d'activation est plus faible. Les régions cérébrales présentant ce phénomène dépendent du type de stimuli traités et de la tâche d'orientation réalisée. La suppression de répétition est alors considérée comme le corrélat neurophysiologique de l'amorçage à court terme (e.g., Henson, 2003). Il a de plus été observé que l'intervalle temporel entre les répétitions (IIR) d'un stimulus influence l'amplitude de la suppression de répétition. Pour les items familiers, lorsque l'IIR augmente, alors l'activité observée lors de P2 augmente. Autrement dit, après un intervalle court, on observe une faible activation lors de P2, qui reprend peu à peu son niveau normal au fur et à mesure de l'augmentation de l'IIR (Henson, Rylands, Ross, Vuilleumier, & Rugg, 2004 ;

Henson, Shallice, & Dolan, 2000). Ainsi, la suppression de répétition est de plus en plus faible lorsque l'IIR augmente.

La suppression de répétition et sa modulation par l'IIR a également été observée au niveau de l'hippocampe et des cortex périrhinaux, entorhinaux et inférotemporaux. Il a été suggéré que cette diminution d'activité serait la base de la reconnaissance (Brown & Xiang, 1998) : on reconnaît qu'un item est une répétition lorsque le niveau d'activité dans ces régions est plus faible qu'en temps normal. Par conséquent, on peut se demander si le phénomène de suppression de répétition pourrait être à l'origine des effets d'espacement et d'intervalle classiquement observés. Selon cette hypothèse, la diminution d'activation observée lorsque l'IIR est court pourrait être à l'origine de la faible rétention ultérieure de l'item.

Il est intéressant de remarquer que l'hypothèse de suppression de répétition peut être considérée comme une variante neurophysiologique de l'hypothèse du traitement déficitaire : en condition massée ou d'intervalle court, le traitement réalisé sur l'information à P2 est moindre que celui réalisé en condition distribuée, ce qui est délétère pour les processus d'encodage et donc pour la récupération ultérieure.

3. 1. 1. Effet d'espacement en imagerie cérébrale

Xue et al. (2010a et 2010b) et Callan et Schweighofer (2010) avaient pour objectif d'étudier les bases neurales de l'effet d'espacement, et pour ce faire, ils ont utilisé des paradigmes relativement similaires. Les stimuli étaient présentés quatre fois, et ces quatre présentations étaient soit massées, c'est-à-dire présentées successivement avec un intervalle inter-essais d'une seconde, soit espacées, c'est-à-dire que les présentations successives étaient séparées par un intervalle plus important comportant d'autres items (entre 6 et 20 items intercalés chez Xue et al. ; entre 5 et 10 minutes dans Callan et Schweighofer). Les deux études différaient par les items utilisés, et par conséquent sur la zone d'intérêt anatomique analysée, puisque le phénomène de suppression de répétition impacte les régions perceptives qui sont spécifiques aux items traités. Xue et al. (2010a) faisaient apprendre aux sujets un ensemble de visages, Xue et al. (2010b) un ensemble de caractères coréens, et Callan et Schweighofer un ensemble de paires mot-pseudomot. La problématique étudiée dans les études était cependant quelque peu différente : Xue et al. se demandaient si l'effet d'espacement était dû à la suppression de répétition neurale réduite dans la condition distribuée, tandis que Callan et Schweighofer ont essayé de départager la théorie du traitement déficitaire de la théorie de la variabilité de l'encodage.

Xue et al. (2010a) voulaient tester l'hypothèse selon laquelle l'effet d'espacement est dû à un déficit de traitement des occurrences répétées d'un item en condition massée, ce déficit étant lui-même associé à un amorçage perceptuel à court terme. Cette hypothèse a été avancée en 1998 par Russo et al., comme nous l'avons vu dans le chapitre 2. L'hypothèse des auteurs était que l'apprentissage espacé devrait entraîner une meilleure reconnaissance finale par rapport à un apprentissage massé en réduisant la suppression de répétition pendant l'apprentissage.

Les auteurs ont utilisé des visages non familiers qu'ils ont répétés 4 fois, soit de façon massée, soit de façon distribuée (entre 6 et 20 items intercalés, moyenne de 12 items). Un test de reconnaissance des visages était réalisé une heure après l'apprentissage. La phase d'apprentissage était réalisée dans un IRM, ce qui permettait d'observer les activations cérébrales associées aux répétitions et donc d'observer la suppression de répétition. La zone cérébrale d'intérêt était dans cette étude le cortex fusiforme, puisqu'une activité de cette région est classiquement observée lors de la perception de visages.

Les résultats ont révélé, d'une part, l'effet d'espacement comportemental classique, c'est-à-dire que l'apprentissage espacé permettait une meilleure reconnaissance finale des visages que l'apprentissage massé. D'autre part, au plan cérébral, les auteurs ont montré que le cortex fusiforme bilatéral était davantage activé pour les items qui étaient reconnus ultérieurement que pour les items non reconnus ultérieurement. Cette même zone présentait également le phénomène de suppression de répétition c'est-à-dire que son activité diminuait au fil des répétitions d'un même item, de telle façon que les items reconnus ultérieurement correspondaient à une situation de plus faible suppression de répétition. Enfin, l'activité du cortex fusiforme bilatéral était affectée par le type d'apprentissage, tel que l'apprentissage espacé était responsable d'une suppression de répétition de moins grande ampleur. Ainsi, la reconnaissance finale était associée à une plus faible suppression de répétition durant l'apprentissage, or l'espacement réduisait la suppression de répétition et améliorait la reconnaissance finale. Cette étude apportait donc des arguments en faveur du rôle causal de la suppression de répétition dans l'effet d'espacement.

Xue et al. (2010b) ont tiré les mêmes conclusions en utilisant comme matériel des caractères coréens nouveaux pour les sujets. Ainsi l'espacement réduisait d'une part la suppression de répétition neurale dans la région du cortex fusiforme gauche et, d'autre part, améliorait la reconnaissance finale.

Callan et Schweighofer (2010) ont fait apprendre aux sujets des paires mot-pseudomot à quatre reprises. Ils ont comparé les prédictions faites par deux hypothèses : celle du traitement déficitaire, plus particulièrement celle de la suppression de répétition, avec celle de la variabilité de l'encodage. Leurs régions d'intérêt étaient les zones cérébrales connues pour présenter une activité lors de l'encodage de matériel verbal, notamment l'operculum frontal gauche. Selon la théorie du traitement déficitaire, à l'encodage, les traitements devraient être moins importants lors des répétitions en condition massée dans un ensemble de zones dont l'operculum frontal. Selon la théorie de la variabilité de l'encodage, l'activation des régions hippocampiques et parahippocampiques devrait être plus importante lors des répétitions en condition distribuée, reflétant une plus grande activité de traitement sémantique des items. La tâche de rappel indicé était réalisée 10 minutes puis un jour après la phase d'encodage.

Les résultats ont montré, en plus de l'effet d'espacement comportemental classique, que l'activité dans la région de l'operculum frontal gauche était plus importante en condition espacée, et qu'elle diminuait de répétition en répétition dans la condition massée uniquement. Ces résultats vont dans le sens de la théorie du traitement déficitaire. A l'inverse, il n'a été observé aucune différence de pattern d'activité dans l'hippocampe au fil les épisodes successifs entre les conditions massée et espacée, ce qui n'est pas concordant avec la théorie de la variabilité de l'encodage.

Ainsi, les trois études ayant porté sur l'effet d'espacement ont révélé des résultats concordants avec l'hypothèse du traitement déficitaire : la répétition massée se caractérise par une diminution d'activation des régions cérébrales d'intérêt au fil des répétitions, contrairement à la répétition distribuée. Peut-on tirer les mêmes conclusions lorsqu'on compare des IIR plus longs ?

3. 1. 2. Effet d'intervalle en imagerie cérébrale

Wagner et al. (2000) ont cherché à mettre en rapport les effets d'amorçage comportemental et neural lors des répétitions avec la performance mnésique ultérieure. L'amorçage neural est l'observation d'une activité réduite dans une région cérébrale à P2 par rapport à P1. Les

auteurs ont mesuré l'activité cérébrale, par le biais de l'IRMf, lors de l'encodage incident d'une liste de mots. Une même liste était encodée à deux reprises, soit avec un IIR de 2 minutes, ce qui constitue un intervalle court, soit après un intervalle de 25h, ce qui constitue un IIR long avec un épisode de sommeil intercalé. Une tâche de reconnaissance des mots avait lieu 48h après la seconde présentation de la liste. Comme l'encodage incident consistait à catégoriser chaque mot selon qu'il était abstrait ou concret, on pouvait analyser les performances à cette tâche lors de P2 et évaluer l'amorçage comportemental. Or, en effet, les mots répétés suite à l'intervalle court ont été catégorisés plus rapidement que les mots répétés suite à l'intervalle long. Autrement dit, l'amorçage comportemental était bien observé et était moins important en condition d'intervalle long que court.

Concernant l'amorçage neural lors de P2, les régions cérébrales montrant une diminution d'activité lorsque les items catégorisés avaient déjà été présentés étaient le cortex préfrontal inférieur gauche et le cortex fusiforme gauche, et ce, quel que soit l'IIR. Cet amorçage neural était plus important lorsque l'IIR était court que lorsqu'il était long. Ainsi, l'amorçage comportemental et l'amorçage neural étaient concordants : les items répétés suite à un IIR court étaient à la fois catégorisés plus rapidement (i.e., amorçage comportemental) et présentaient une plus grande suppression de répétition (i.e., amorçage neural) que les items répétés suite à un IIR long.

Quel était le lien entre amorçage et reconnaissance finale ? Premièrement, la reconnaissance était plus précise lorsque l'IIR était long que court. Les auteurs ont ensuite déterminé dans la condition d'IIR long, 1) la corrélation entre l'amorçage comportemental et les performances de reconnaissance, et 2) la corrélation entre l'amorçage neural dans le cortex préfrontal inférieur gauche et postérieur et les performances de reconnaissance. Ces corrélations se sont révélées significativement négatives. Autrement dit, plus l'amorçage comportemental et/ou neural était important, plus la reconnaissance ultérieure était faible.

Les auteurs interprétaient ces résultats de la façon suivante : l'amorçage suscité lors d'une répétition suite à un IIR court est associé à un traitement cognitif restreint du stimulus répété. Or, ce traitement restreint est délétère pour la rétention à long terme. Ainsi, cette étude apporte des arguments, tout comme les études précédentes, en faveur d'une interprétation de l'effet d'intervalle en termes de traitement déficitaire.

Cependant, les résultats de Wagner et al. (2000) sont à considérer avec précaution car ils n'ont pas été répliqués par Stark, Gordon, et Stark (2008). Les auteurs ont exactement répliqué, mais au plan comportemental uniquement, l'expérience de Wagner et al. Ils ont également observé que l'IIR long était associé à la fois à une reconnaissance finale plus importante et à un effet d'amorçage plus faible, et que l'IIR court était associé à la fois à une reconnaissance finale plus faible et un effet d'amorçage plus fort. Cependant, ils n'ont pas répliqué la corrélation négative observée par Wagner et al. au niveau des individus, c'est-à-dire que les individus ayant présenté le plus grand niveau d'amorçage auraient dû présenter également un plus faible niveau de reconnaissance. Les observations n'ont pas non plus été répliquées au niveau des items, c'est-à-dire que les items ayant subi le plus grand effet d'amorçage auraient dû présenter aussi un plus faible niveau de reconnaissance. Ils concluent que les deux processus mnésiques que sont l'amorçage - processus mnésique implicite - et la reconnaissance - processus mnésique explicite - agissent en parallèle sans interférer l'un sur l'autre.

Par conséquent, ces deux dernières études présentent des données contradictoires et ne permettent pas de confirmer ou d'infirmer l'idée que l'effet d'amorçage est responsable de la diminution des performances finales de reconnaissance ou de rappel lorsqu'on compare

des IIR de 2 minutes et de 25 heures. De nouvelles études sont donc nécessaires pour en savoir plus à ce sujet.

3. 1. 3. Conclusion

L'exploration des bases neurales de l'effet de pratique distribuée n'en est qu'à ses débuts. Nous avons vu qu'une des théories les plus prometteuses actuellement est celle de la suppression de répétition, du moins lorsque les IIR sont courts. Les études à venir auront pour objectif de répondre à plusieurs questions. Premièrement, lorsque les IIR sont supérieurs à 24h, la suppression de répétition est-elle un facteur pertinent pour expliquer les effets de pratique distribuée ? On peut s'interroger sur ce point, car il semble que l'effet de suppression de répétition est observé lorsque les répétitions se produisent suite à des intervalles de l'ordre de quelques minutes, mais pas lorsque les intervalles sont plus longs (e.g., Brozinsky, Yonelinas, Kroll, & Ranganath, 2005). Deuxièmement, pour des protocoles à long terme, il serait utile de pouvoir observer les corrélats neuroanatomiques de la consolidation mnésique, laquelle devrait être plus avancée à P2 suite à un IIR long que suite à un IIR court. Une piste d'étude pourrait être liée à la théorie de la récupération en phase d'étude : serait-il possible de détecter des corrélats de la récupération lors des répétitions, c'est-à-dire un signal qui serait associé à la récupération en mémoire de la trace de la (ou des) présentation(s) antérieures ? Cela permettrait d'obtenir des informations sur les processus de récupération des traces sans avoir à faire réaliser au sujet un test sur l'item. De nombreuses questions sont donc ouvertes.

3. 2. Modélisation mathématique

L'effet de pratique distribuée a fait l'objet de plusieurs tentatives de modélisation mathématique. Loin de vouloir décrire dans le détail ces modèles, nous décrivons dans cette section en quoi, selon nous, les efforts de modélisation ont permis d'apporter de nouveaux éléments pour répondre au problème de l'origine des effets observés.

En général, les modèles mathématiques de l'effet de pratique distribuée complètent des modèles de mémoire existants pour les faire s'adapter à ces effets (e.g., les modèles ACT-R de Pavlik & Anderson, 2005 ; et SAM de Raaijmakers, 2003) ou combinent les caractéristiques de plusieurs modèles en un seul (e.g., le modèle MCM de Mozer et al., 2009). Le modèle de Pavlik et Anderson, par exemple, essaie de modéliser les effets de pratique distribuée en intégrant auparavant les lois, mieux connues, déterminant la courbe d'oubli et la courbe d'apprentissage.

En quoi est-il utile de réaliser des modèles mathématiques de phénomènes psychologiques ? En premier lieu, la modélisation permet d'aboutir à la description simple des phénomènes observés, sans nécessiter de faire appel à des théories psychologiques. Par exemple, une description verbale n'est pas aisée pour décrire l'interaction entre l'IIR et le DR observée par exemple dans Cepeda et al. (2008). Une description par le biais d'équations permet de simplifier notre appréhension de la globalité d'un phénomène, en remplaçant plusieurs lignes d'explication verbale.

De la même façon, les modèles mathématiques permettent d'appréhender la complexité des phénomènes et de certaines théories. Par exemple, la théorie de la variabilité de l'encodage décrite par Glenberg (1979) est particulièrement complexe au regard des notions de chevauchement de composants contextuels, de fluctuation aléatoire, etc. Il est certain que le développement de ce type de théories est rapidement limité si on s'en tient à des descriptions verbales qui favorisent le flou dans la description de la théorie

et donc dans les prédictions du modèle. Benjamin et Tullis (2010) déclarent à ce sujet que « *la majorité des approches théoriques n'ont pas été implémentées computationnellement et sont par conséquent sujettes à des différences d'interprétation* ».

En second lieu, les modèles permettent de faire des simulations qui font gagner du temps par rapport à des expériences empiriques longues et, parfois, peuvent même remplacer des expériences qui, en pratique, ne peuvent pas être réalisées. Par exemple, Lindsey et al. (2009) expliquent qu'il est impossible de tester expérimentalement toutes les combinaisons d'agencement temporel de plus de deux épisodes d'apprentissage, sur toutes les durées possibles, avec différents matériels. Grâce aux modèles mathématiques, on peut simuler une situation complexe et la valider ensuite empiriquement en ne testant que certains cas parmi tous les cas possibles.

Enfin, les formulations mathématiques et les simulations font émerger de nouveaux questionnements et de nouvelles pistes d'exploration dans l'étude d'un effet. En particulier, plusieurs modèles (Raaijmakers, 2003 ; Pavlik & Anderson, 2005 ; Mozer et al., 2009) placent au cœur de leur formulation la notion de niveau d'activation de la trace en mémoire, et son évolution avec le temps. Cette notion est assez rarement mentionnée dans la littérature sur les effets de pratique distribuée (sauf par exemple dans Anderson & Schooler, 1991) et, quand elle est mentionnée, elle l'est de façon floue et peu explicitée. Or, il nous semble que la notion d'activation ou de force d'un souvenir est un concept très important qui peut être utile pour décrire et expliquer les effets de répétition.

Nous présenterons trois modèles qui ont eu une certaine influence dans la littérature des effets de pratique distribuée. Il s'agit des travaux de Raaijmakers (2003), Pavlik et Anderson (2005), et Mozer et al. (2009).

3. 2. 1. Le modèle de Raaijmakers (2003)

Le modèle de Raaijmakers avait pour objectif de modéliser la théorie de la variabilité de l'encodage dans l'explication des effets d'espacement et d'intervalle, et en particulier le modèle développé par Glenberg en 1979. Pour ce faire, l'auteur a perfectionné le modèle déjà existant intitulé SAM développé par Raaijmakers et Shiffrin (1980). Dans ce modèle, il est postulé que durant le stockage, l'information est représentée sous forme de traces mnésiques qui contiennent des informations sur l'item, sur les associations en cours, et sur le contexte. Le rappel est quant à lui basé sur les indices de récupération, tels que les mots fournis par l'expérimentateur ou les indices contextuels. Le fait de récupérer une trace dépend de la *force d'association* entre les indices de récupération et les informations contenues dans la trace. Le modèle implémente également une fluctuation contextuelle continue au fil du temps, selon laquelle le contexte courant devient de plus en plus dissimilaire au contexte initial. Ainsi, la force d'association entre les indices contextuels et les informations contenues dans la trace dépend à la fois de l'intervalle entre les deux occurrences du stimulus (IIR) et du délai de rétention (DR). De plus, le modèle implémente le processus de récupération en phase d'étude, qui n'était pas explicite chez Glenberg (1979) ; l'auteur considère comme important que l'item soit récupéré ou reconnu lorsqu'il se répète, pour que les éléments contextuels additionnels soient stockés dans la trace existante et soient utiles à la récupération finale. Si l'item n'est pas récupéré ou pas reconnu, alors une nouvelle trace est formée. Par conséquent, ce modèle simule un modèle mixte de la pratique distribuée.

Le modèle a été confronté aux données empiriques de trois expériences réalisées à court terme. Les simulations ont fait apparaître l'interaction entre l'IIR et le DR, ainsi que la relation en U inversé de la performance en fonction de l'IIR (e.g., Glenberg, 1976).

3. 2. 2. Le modèle de Pavlik et Anderson (2005)

Le modèle de Pavlik et Anderson avait pour objectif de simuler la théorie de l'accessibilité (voir section 2.3). Il s'agissait de perfectionner le modèle ACT-R existant (Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004) qui ne parvenait pas à rendre compte de l'effet d'intervalle et de l'interaction entre l'IIR et le DR. Le modèle considérait comme centrale la force -ou niveau d'activation- d'un item mnésique, notée m . En effet, la probabilité de rappel de cet item au moment du test était fonction de son niveau d'activation m . Le niveau d'activation m d'un item correspondait à la somme de tous les renforcements de cet item, renforcements produits par les répétitions de cet item au fil du temps; chacun de ces renforcements s'ajoutait au niveau m courant. Or, chaque renforcement de l'activation possédait lui-même son propre taux de déclin, noté d , qui décrivait une fonction puissance. Dans l'ancienne version du modèle, le taux de déclin d était constant. Dans la nouvelle version du modèle, le taux de déclin d était différent selon le niveau d'activation m de l'item au moment de l'occurrence. Ainsi, si l'item se trouve dans un état fortement actif au moment de l'occurrence (m élevé), alors le taux de déclin d de l'activation pour cette occurrence sera également élevé. À l'inverse, si l'item est faiblement activé au moment de l'occurrence (m faible), alors le taux de déclin d de l'activation pour cette occurrence sera également faible. Il en découle qu'il est plus avantageux qu'une occurrence de l'item se produise lorsque celui-ci possède un faible niveau m , ce qui est plus probable après un long IIR. On retrouve ainsi l'effet d'intervalle.

Le modèle parvenait à simuler de façon convaincante plusieurs expériences antérieures sur les effets d'espacement et d'intervalle, dont quatre sur le court terme, et, ce qui est notable, une sur le long terme (i.e., l'expérience de Bahrick, 1979, décrite dans la section 1.3.2). L'intérêt particulier de ce modèle est qu'il mettait l'accent sur la notion de déclin de la trace, qui est une notion proche de celle du taux oubli, dont nous discuterons dans la section 3.3.

3. 2. 3. Le modèle de Mozer et al. (2009)

Le modèle MCM (*Multiscale Context Model*) de Mozer et al. avait pour objectif de caractériser l'état et l'évolution de la force de la trace mnésique d'items appris lors d'une ou plusieurs sessions. La particularité de ce modèle était de pouvoir faire des prédictions sur la performance de rappel d'un item en fonction du (ou des) IIR et du DR, mais également de la nature de l'item à apprendre.

Le modèle se basait sur deux fonctions essentielles de l'apprentissage : la fonction d'espacement et la fonction d'oubli. Ces deux fonctions sont décrites dans la Figure 6. La fonction d'espacement décrit une courbe en U inversé de la performance finale (axe des ordonnées) en fonction de l'IIR (axe des abscisses) : pour un IIR très court, ce qui correspond à une répétition massée, la performance est faible. Lorsque l'IIR augmente, la performance augmente également jusqu'à un IIR optimal, puis se met à décliner jusqu'à atteindre une asymptote correspondant à une performance moyenne (e.g., Glenberg, 1976 ; Cepeda et al., 2008). La fonction d'oubli décrit une performance qui décroît selon une fonction puissance (e.g., Wixted & Carpenter, 2007 ; Wixted & Ebbesen, 1991) en fonction de la durée du DR.

Concrètement, le modèle MCM est un modèle complexe qui combine les propriétés de deux modèles antérieurs, le modèle SAM de Raaijmakers (2003), présenté dans cette section, et le modèle MTS de Staddon, Chelaru, et Higa (2002). Ce qui a été inspiré du modèle SAM est, d'une part, la notion de fluctuation contextuelle progressive, inspirée par

la théorie de la variabilité de l'encodage (e.g., Glenberg, 1979), et, d'autre part, l'importance de récupérer la trace de la (ou des) occurrence(s) antérieure(s) au moment de la répétition, inspiré par la théorie de la récupération en phase d'étude (e.g., Thios & D'Agostino, 1976). Ainsi, ce modèle constitue également une simulation d'un modèle « mixte ». Ce qui a été inspiré du modèle MTS est la notion d'utilité prédictive que nous allons développer.

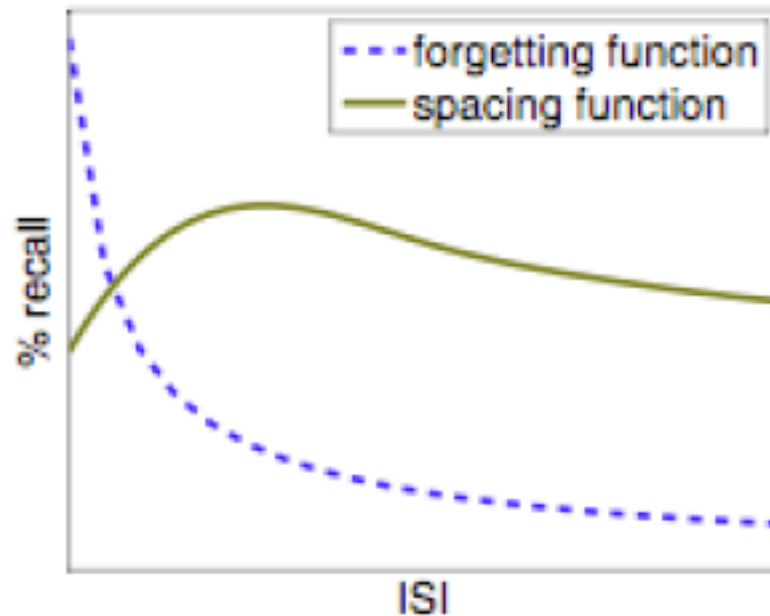


Figure 6 : Reproduction des fonctions d'oubli et d'espacement décrites dans le modèle de Mozer et al. (2009). La fonction d'espacement (spacing function) représente le rappel lors du test suite à deux sessions d'apprentissage séparées par un IIR donné (ISI). La fonction d'oubli (forgetting function) représente le rappel en fonction du DR (en abscisses).

Les théories de l'utilité prédictive (*predictive-utility theories*) ont pour postulat que la mémoire possède une capacité limitée et/ou est imparfaite et permet des intrusions. Pour pouvoir atteindre des performances optimales, certains souvenirs doivent être effacés s'ils ne sont pas nécessaires dans le futur. Anderson et Milson (1989), à partir d'une analyse rationnelle, ont pu estimer la probabilité future d'avoir besoin d'un souvenir. Quand un item est étudié de nombreuses fois avec un IIR donné, l'analyse rationnelle suggère que la probabilité de besoin diminue rapidement suite au dernier épisode d'étude, une fois qu'un DR plus grand que l'IIR s'est écoulé. Par conséquent, augmenter l'IIR doit mener à une trace mnésique plus persistante. (Nous développerons davantage ces notions dans la section 3.3). La notion d'utilité prédictive est implémentée dans le modèle MTS de Staddon et al. (2002). Sans décrire dans le détail les équations du modèle, celui-ci énonce que quand un item est présenté de façon répétée avec des IIR courts, la trace peut être représentée par des intégrateurs ayant des constantes de temps courtes et, par conséquent, la trace va décliner rapidement. Augmenter l'IIR fait donc en sorte que les intégrateurs ont des taux de déclin plus faibles.

Le modèle MCM implémentait ces notions d'une façon complexe. De façon simplifiée, il considérait que lorsqu'un item est étudié, une trace mnésique est formée qui inclut le contexte psychologique ; ce dernier est supposé fluctuer de façon aléatoire avec le temps.

La probabilité de rappel ultérieur dépend en partie de la similarité entre les représentations du contexte lors de l'étude et lors du test. La caractéristique de ce modèle qui le distingue des autres (en particulier de celui de Raaijmakers, 2003) est que le contexte est supposé varier sur de multiples échelles temporelles, ce qui lui permet de modéliser des situations dans lesquelles l'apprentissage a lieu sur de multiples sessions.

L'objectif des auteurs, dans une perspective d'application, était de pouvoir prédire l'espacement optimal en fonction de plusieurs éléments comme le type de matériel et la nature des essais d'apprentissage. Leur approche consistait à récupérer empiriquement les paramètres de la fonction d'oubli grâce à un test réalisé juste avant la session de répétition, puis à partir de ces données, de faire des prédictions sur la fonction d'espacement, ce qui permettait ensuite de faire des recommandations en termes d'intervalles optimaux.

Les auteurs ont réalisé une simulation à partir des données des expériences de Cepeda et al. (2008 ; 2009). Les paramètres de la fonction d'oubli étaient calculés à partir des résultats du rappel indicé au début de la seconde session d'apprentissage, et donc après des IIR variables. La courbe d'espacement était quant à elle déterminée à partir des résultats du rappel final. À partir des équations du modèle et des paramètres déterminés, les courbes d'espacement pouvaient être prédites et comparées aux courbes observées. Or, les courbes prédites se sont révélées très similaires aux courbes observées. Le modèle s'est révélé capable d'anticiper le sommet des courbes d'espacement et la forme globale de ces courbes en fonction du DR. De plus, avec les paramètres déterminés grâce aux données de Cepeda et al. (2008 ; 2009) et sans nouvelle modification, ils ont obtenu une simulation proche de la réalité pour les données recueillies dans la méta-analyse de Cepeda et al. (2006).

L'existence de plusieurs modèles aboutis rendant compte des effets de pratique distribuée appelle leur comparaison afin de déterminer lequel est le plus pertinent.

3. 2. 4. L'étude de Lindsey et al. (2009)

L'étude de Lindsey et al. avait pour objectif de comparer la validité respective des modèles de Mozer et al. (2009) et de Pavlik et Anderson (2005), et pour ce faire, ils se sont attachés aux situations dans lesquelles plus de deux occurrences d'un item se produisent. Il s'agissait d'observer quelles étaient les prédictions des deux modèles concernant l'effet de l'agencement des répétitions, c'est-à-dire quel agencement (uniforme, expansif, ou contractant) était le plus bénéfique pour la rétention finale. Il se trouve que l'effet de l'agencement temporel n'avait jamais été simulé auparavant. Les auteurs ont réalisé des milliers de simulations de chacun des modèles pour la situation suivante : un item était présenté lors de trois occurrences -ce qui correspond à deux IIR successifs de durées variables- puis testé après un certain délai de rétention (DR). La durée totale entre P1 et P3 n'était pas fixée. Les DR étaient compris entre 10 secondes et 300 jours, et les auteurs faisaient varier aléatoirement les paramètres libres des modèles. Pour un DR donné, le modèle donnait en réponse les durées du premier et du second IIR qui produisaient la meilleure performance de rappel final. De façon intéressante, les deux modèles faisaient des prédictions robustes mais contradictoires concernant l'agencement optimal. Les résultats sont présentés dans la Figure 7, avec les prédictions du modèle de Pavlik et Anderson en haut et celles du modèle de Mozer et al. en bas. Les graphes présentent, en fonction du DR en abscisses, la proportion des simulations qui concluaient que l'agencement uniforme (en bleu), expansif (en vert), ou contractant (en rouge) était l'agencement optimal. Le modèle de Pavlik et Anderson prédisait toujours qu'un agencement contractant était le plus efficace. Le modèle de Mozer et al. faisait des prédictions différentes qui, d'une part, contredisaient celles du modèle de Pavlik et Anderson, et, d'autre part, variaient

selon la durée du DR. Au fur et à mesure de l'allongement du DR, la proportion des prédictions qui donnaient l'avantage à l'agencement expansif augmentait tandis que la proportion des prédictions qui donnaient l'avantage à l'agencement uniforme diminuait. Les prédictions en faveur de l'agencement contractant restaient relativement faibles. Ainsi, selon les prédictions du modèle, l'agencement expansif était supposé devenir de plus en plus avantageux avec l'augmentation du DR. Or, les rares données empiriques existantes sur l'effet de l'agencement (e.g., Landauer & Bjork, 1978 ; Tsai, 1927) ont observé que l'agencement contractant était le moins favorable des trois agencements ; ces résultats vont donc à l'encontre de ceux du modèle de Pavlik et Anderson, et dans le sens du modèle de Mozer et al. (2009). Par ailleurs, les simulations ont prédit que lorsque le DR augmente, les durées des IIR optimaux augmentent également, ce qui est cohérent avec les données empiriques existantes (e.g., Cepeda et al., 2008).

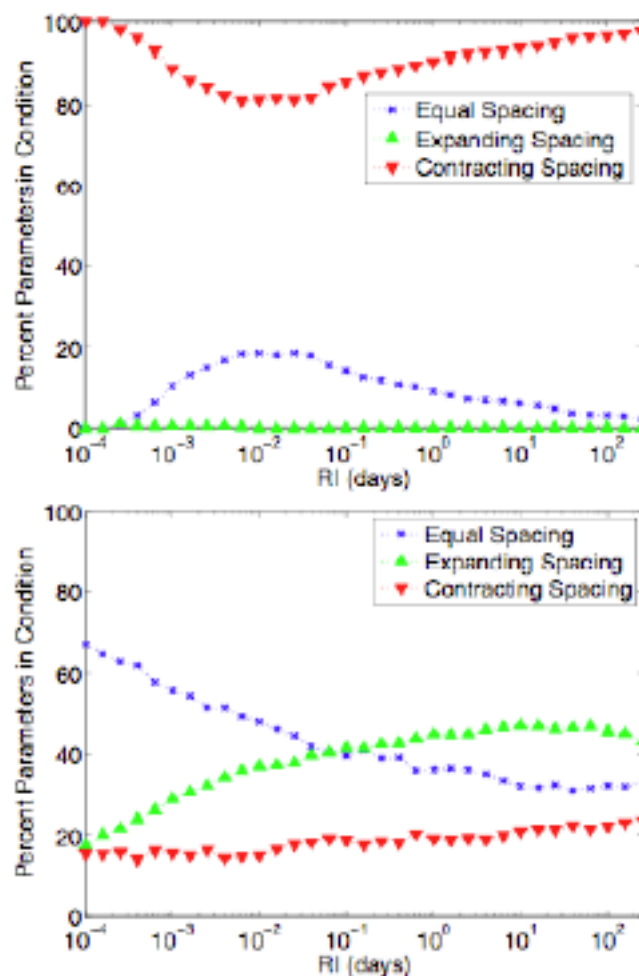


Figure 7 : Reproduction des résultats des simulations réalisées par Lindsey et al. (2009) du modèle de Pavlik et Anderson (2005, en haut), et de Mozer et al. (2009, en bas). En abscisses est représenté le délai de rétention (RI) en jours. En ordonnées est représenté le pourcentage relatif des simulations révélant l'avantage d'un agencement particulier (uniforme en bleu, expansif en vert, contractant en rouge).

Les éléments présentés dans cette section montrent bien que la modélisation mathématique se révèle un outil intéressant pour étudier les effets de la répétition d'une

information sur la mémorisation, dans la mesure où on peut apporter des arguments en faveur d'une théorie en confrontant les prédictions a priori du modèle aux résultats empiriques observés a posteriori. En l'occurrence, le modèle de Mozer et al. (2009) semble le plus abouti car ses prédictions sur les résultats de Cepeda et al. (2008 ; 2009) -qui sont les plus écologiques en terme de durée des intervalles- étaient très fiables. Il est très intéressant pour notre problématique de l'agencement temporel des répétitions à long terme d'analyser les prédictions faites par le modèle de Mozer et al. Aujourd'hui, aucune étude à notre connaissance, ni à celle des auteurs, n'a montré d'interaction entre les effets de l'agencement temporel à long terme et la durée du DR ; il est intéressant d'observer que ce modèle prédit un avantage croissant de l'agencement expansif sur les autres lorsque le DR augmente.

3. 3. Consolidation et taux d'oubli

Nous avons vu au début du chapitre 2 qu'une des hypothèses visant à expliquer les effets de pratique distribuée s'intitulait l'hypothèse de la *consolidation*. Cette hypothèse postulait que, au moment de P2, la trace nouvellement formée héritait du degré de consolidation de la trace créée lors de P1. Selon cette hypothèse, plus l'intervalle entre les occurrences P1 et P2 est long, plus l'état de consolidation de la trace mnésique récupérée lors de P2 est avancé. Par conséquent, le rappel final est favorisé. Cette hypothèse n'a pas été beaucoup développée en psychologie dans la mesure où les études ont montré que le locus de l'effet de pratique distribuée se situe au moment de P2 et non au moment de P1 (e.g., Hintzman et al., 1973). La citation suivante reflète bien le parti pris des psychologues vis-à-vis de la consolidation dans le cadre des recherches sur l'effet de pratique distribuée dans les années 1970 : « *Je trouve rassurant que personne dans le symposium n'ait fait appel à la consolidation comme une explication de l'effet observé, mais que tous aient cherché, à la place, à trouver des hypothèses sur les activités de traitement du sujet* » (Melton, 1970). A contrario, on observe actuellement un intérêt particulier pour la notion de consolidation et pour celle de l'oubli dans le champ des neurosciences, et ce, en particulier, dans le cadre de la mémoire déclarative. Dans cette section, nous nous intéresserons aux phénomènes de consolidation, d'oubli, de reconsolidation, et du sommeil, dans leur lien potentiel avec l'effet de la répétition sur la mémoire, et terminerons par des considérations évolutionnistes sur cet effets.

3. 3. 1. Consolidation

La consolidation et les mécanismes qui la sous-tendent sont l'objet actuellement de nombreuses publications, lesquelles portent majoritairement sur des modèles animaux. Il se trouve que, globalement, assez peu de liens sont créés entre les conceptions psychologiques traditionnelles de la consolidation mnésique et les conceptions plus récentes issues des neurosciences, ce qui est regrettable (voir cependant Hardt, Orn Einarsson, & Nader, 2010). Concernant l'effet de pratique distribuée qui nous concerne, les deux traditions s'intéressent à des niveaux d'explication différents. La psychologie tente d'expliquer les effets observés en termes de processus cognitifs liés au traitement des informations par le sujet, tandis que les neurosciences s'intéressent à des phénomènes plus élémentaires se déroulant par exemple au niveau de la cellule. Il s'avère le terme de consolidation ne réfère pas exactement aux mêmes conceptions dans les deux traditions. Dans le champ de la psychologie cognitive, la consolidation réfère au fait qu'une trace mnésique, avec le temps, devient de plus en plus stable et de moins en moins sensible aux interférences. Dans le champ des neurosciences, la consolidation réfère à deux processus,

celui de la consolidation synaptique (elle-même composée d'une phase précoce et d'une phase tardive) et celui de la consolidation systémique (pour une revue, voir Frankland & Bontempi, 2005). La consolidation synaptique se déroule en général sur plusieurs heures et fait suite à une activité intense et répétée des synapses. Elle correspond à des modifications au sein des neurones qui aboutissent à un renforcement des connexions synaptiques sous-tendant les traces mnésiques. Elle se définit par la nécessité d'une synthèse protéique. La consolidation systémique, quant à elle, se déroule sur des jours, des mois, voire des années, et se définit par le transfert des régions de stockage des souvenirs de l'hippocampe vers le néocortex (Alvarez & Squire, 1994 ; McClelland, McNaughton, & O'Reilly, 1995).

Ainsi, des processus physiologiques sous-tendent les phénomènes mnésiques et, a fortiori, sous-tendent également les effets de pratique distribuée. Des études récentes ont montré l'influence de la distribution temporelle des stimulations sur certains processus physiologiques. Premièrement, un large ensemble de phénomènes mnésiques (e.g., habituation, sensibilisation, conditionnement...) sont influencés par la façon dont les occurrences d'un stimulus sont distribuées dans le temps, et ce dans des organismes tels que l'aplysie, la drosophile, la mouche, le rat, etc (Kandel, 2001). Deuxièmement, il a été montré que les processus physiologiques sous-tendant la potentialisation à long terme (PLT) sont influencés par la distribution des répétitions. Par exemple, Scharf et al. (2002) ont montré *in vitro* que la PLT dans l'hippocampe est plus importante suite à des stimulations électriques distribuées que massées. Troisièmement, des modifications cellulaires associées à l'apprentissage sont également influencées par la distribution des stimulations : Sisti, Glass, et Shors (2007) ont montré chez le rat qu'un apprentissage distribué d'une tâche de localisation spatiale favorisait davantage la survie des neurones nouvellement formés dans l'hippocampe, par rapport à un apprentissage massé. Wu, Deisseroth, et Tsien (2001) ont montré le développement de nouvelles dendrites dans l'hippocampe suite à des stimulations électriques distribuées mais pas massées. Enfin, des modifications génétiques peuvent faire en sorte que les individus porteurs d'une mutation particulière présentent un apprentissage massé aussi efficace que l'apprentissage distribué (Genoux et al., 2002 ; Pagani, Oishi, Gelb, & Zhong, 2009). Les voies cellulaires responsables des effets de pratique distribuée (e.g., MAPK, CREB) commencent à être mieux comprises.

Un phénomène essentiel lié à la consolidation est celui de la réactivation. En effet, selon Frankland et Bontempi (2005), « *la réactivation mnésique est le mécanisme au cœur de la théorie de la consolidation* ». Il s'avère que cette réactivation peut être spontanée (par exemple pendant le sommeil) mais également provoquée (c'est-à-dire pendant l'éveil). Or, la répétition d'un stimulus est un événement réactivateur par excellence. La question est alors de déterminer, pour que cette consolidation physiologique des souvenirs soit la plus efficace possible, l'intervalle optimal à l'issue duquel il faut réactiver les souvenirs. Cette question rejoint largement la notion de récupération en phase d'étude développée dans la section 2.2.

Une autre observation intéressante relative à la notion de réactivation est l'hypothèse que le cortex préfrontal inhiberait l'hippocampe lors de la récupération des souvenirs anciens. Des études d'imagerie chez les animaux (Maviel, Durkin, Menzaghi, & Bontempi, 2004) ont montré que l'activité hippocampique est fortement inhibée quand des souvenirs spatiaux et contextuels anciens sont récupérés. Il semblerait que le cortex préfrontal soit à l'origine de cette inactivation. Le fait que l'inhibition se produise ou non pourrait dépendre du fait que le stimulus perçu correspond à un souvenir existant ou non. Si la récupération de ce souvenir est réussie, la fonction hippocampique sera rapidement inhibée afin de ne

pas ré-encoder les souvenirs anciens. Si la récupération du souvenir est impossible (e.g., si l'information a été oubliée), alors il n'y aura pas d'inhibition et l'hippocampe sera ré-engagé. Cette hypothèse est intéressante relativement à la théorie de la récupération en phase d'étude : si les stimuli présentés sont connus, alors l'hippocampe subirait une inactivation et ne les encoderait pas de nouveau. Dans ce cas, des processus différents de ceux ayant eu lieu lors l'encodage initial sont probablement mis en œuvre.

Toutefois, peu d'informations sont disponibles à propos de l'influence de la pratique distribuée à long terme sur la consolidation mnésique (voir cependant Litman et Davachi, 2008 décrit plus loin), et encore moins à propos de l'effet de l'agencement temporel des répétitions à long terme.

3. 3. 2. Taux d'oubli

Un autre concept important lié à celui de la consolidation est le concept d'oubli. En effet, la consolidation est responsable du fait que les souvenirs se renforcent et deviennent, avec le temps, moins sujets à l'oubli et à l'interférence. Ainsi, un souvenir consolidé va présenter un taux d'oubli plus faible qu'un souvenir non consolidé, et le taux d'oubli permet de mesurer indirectement le niveau de consolidation d'un souvenir (e.g., Wixted, 2004).

Par exemple, en 2008, l'étude de Litman et Davachi visait à rapprocher les approches psychologique et physiologique dans l'explication de l'effet de pratique distribuée. L'hypothèse des auteurs était qu'un apprentissage avec un long IIR était avantageux par rapport à un apprentissage avec un court IIR non pas par la comparaison de la rétention à un moment donné, mais plutôt par la comparaison des taux d'oubli dans les deux types d'apprentissage. Les participants devaient apprendre un ensemble de paires de mots concrets présentés deux fois, soit avec un IIR de 5 minutes, soit avec un IIR de 24 h. Pour étudier le taux d'oubli, la rétention était mesurée à deux reprises, d'abord immédiatement après la seconde phase d'encodage, puis 24 h après. Les scores de rappel dans les deux conditions ne différaient pas après le délai de rétention de 5 min ; au contraire, après le délai de rétention de 24 h, les items fortement espacés étaient mieux rappelés que les items faiblement espacés. Ces données suggèrent donc que la réactivation des apprentissages après 24 h de consolidation rend les traces mnésiques plus résistantes à l'oubli qu'une réactivation plus immédiate. Les résultats sont présentés dans la Figure 8.

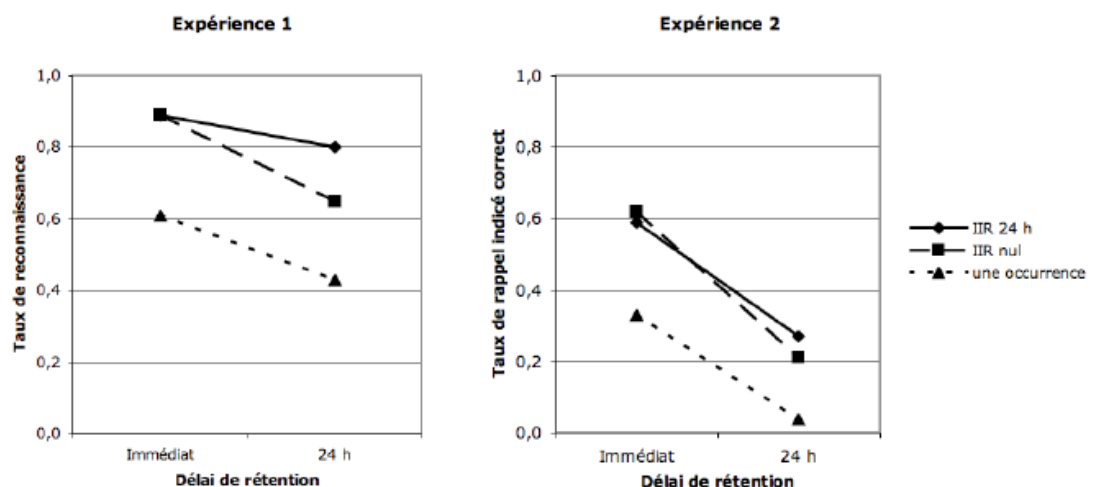


Figure 8 : Reproduction des résultats de l'étude de Litman et Davachi (2008, ; Expérience 1 à gauche et Expérience 2 à droite). Taux de bonnes réponses en fonction de l'intervalle inter-répétition (couleur de la courbe) et du délai de rétention (axe des abscisses).

En psychologie cognitive, quelques études ont porté sur l'évolution de la rétention avec le temps en fonction de l'IIR (e. g., Robinson 1921 ; Rohrer & Taylor, 2006). En 1921, Robinson a réalisé une étude dans le but d'étudier l'effet de pratique distribuée. Par exemple, dans une première expérience, les sujets devaient apprendre une liste de 10 nombres, la liste étant présentée 12 fois. Ces présentations étaient soit de type massé, c'est-à-dire que la liste était présentée 12 fois de suite avec un intervalle inter-liste de 6 secondes, soit de type distribué, c'est-à-dire que la liste était présentée 6 fois de suite un jour, puis 6 fois de suite le lendemain. Le DR était de 5 minutes, 20 minutes, ou 24h. De façon intéressante, un même sujet réalisait les six conditions expérimentales et ce à deux reprises. Les courbes de rétention sont présentées dans la Figure 9. Ces courbes mettent bien en évidence que même si la condition « massée » est parfois plus efficace lorsque le DR est très court, le taux d'oubli est globalement beaucoup plus important en condition massée qu'en condition distribuée. Dans l'Expérience 2 on observait même un croisement des courbes.

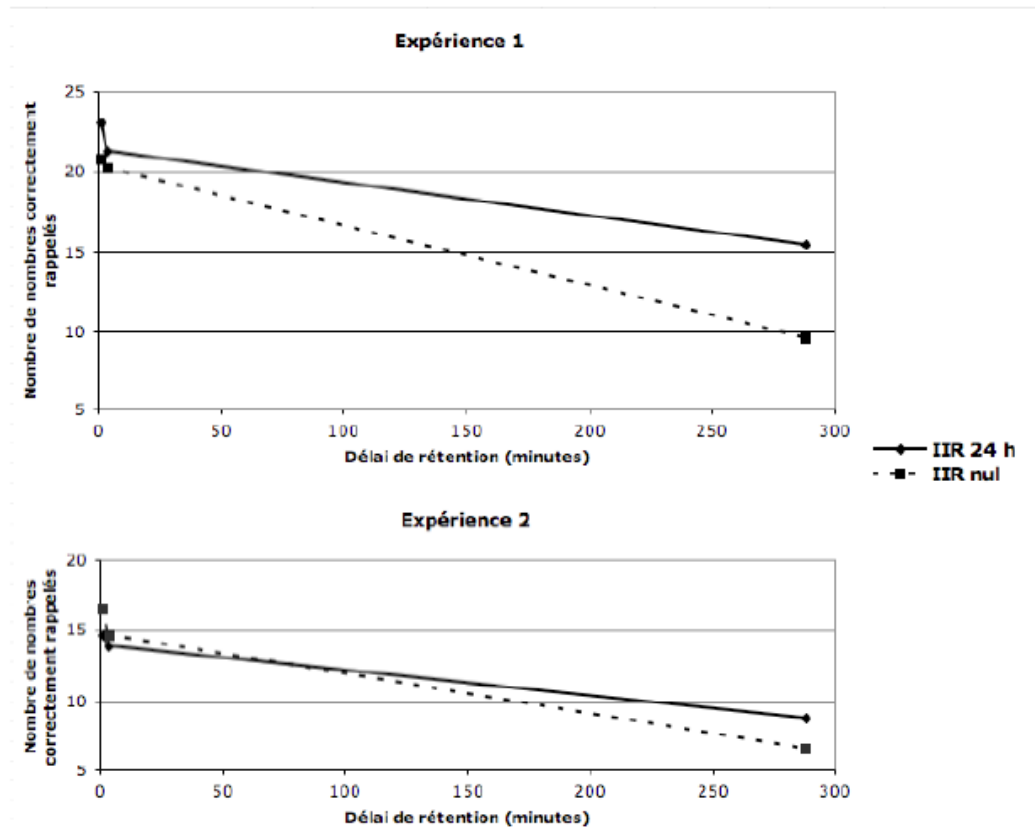


Figure 9 : Reproduction des résultats de l'étude de Robinson (1921 ; Expérience 1 en haut et Expérience 2 en bas). Performance de rappel en fonction de l'intervalle inter-répétition (couleur de la courbe) et du délai de rétention (axe des abscisses).

Dans la même veine, Rohrer et Taylor (2006, Expérience 1) ont fait apprendre à des étudiants des procédures pour résoudre des problèmes mathématiques. Dans l'Expérience 1, les étudiants étaient confrontés à 10 problèmes, soit lors d'une seule session expérimentale (condition pseudo-massée), soit ces 10 problèmes étaient répartis lors de deux sessions séparées par une semaine (condition distribuée). Il a été observé que la condition distribuée n'était pas supérieure à la condition pseudo-massée lorsque les connaissances étaient testées après une semaine, alors qu'elle lui était nettement supérieure lorsque le test avait lieu après quatre semaines. Ces résultats sont présentés dans la Figure 10. Par ailleurs, le même type de résultats a été observé dans les études sur une courte échelle de temps (e.g., Pavlik et Anderson, 2005).

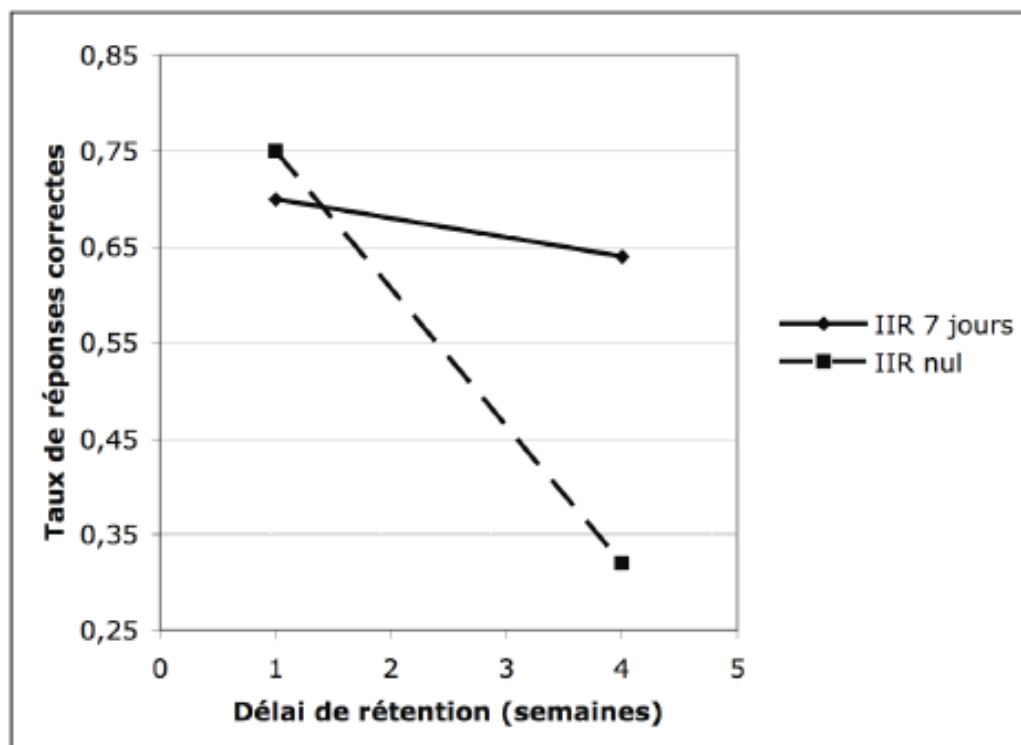


Figure 10 : Reproduction des résultats de l'étude de Rohrer et Taylor (2006, Expérience 1). Taux de rappel en fonction de l'intervalle inter-répétition (couleur de la courbe) et du délai de rétention (axe des abscisses).

Relativement aux conceptions neuroscientifiques de la consolidation, ces résultats peuvent indiquer que, au niveau cellulaire, les mécanismes de renforcement synaptique qui supportent la consolidation cellulaire pourraient être influencés de façon différente selon l'espacement des réactivations mnésiques, et induire des conséquences à long terme différentes (Litman & Davachi, 2008). En termes de consolidation systémique, ces résultats suggèrent l'explication suivante. Lorsque l'IIR est court, les traces mnésiques élaborées lors de P1, probablement sous-tendues par des mécanismes hippocampiques, pourraient ne pas être complètement stabilisées au moment de P2. Au contraire, ces traces ne pourraient commencer à se consolider, c'est-à-dire à « migrer » vers le néocortex, qu'après un intervalle de 24h. Cette consolidation serait par ailleurs largement mise en œuvre au cours du sommeil, comme nous le verrons dans la section 3.3.4. Il serait donc plus utile de réactiver un souvenir consolidé (i.e., IIR long) que labile (i.e., IIR court).

Au plan psychologique, des modèles rendant compte de ces observations ont été proposés, comme par exemple le modèle de Wickelgren (1973). Il s'agit d'un modèle mathématique du stockage en mémoire à long terme dont le présupposé de base est le suivant : la trace mnésique posséderait deux propriétés qui seraient critiques pour la caractérisation de son stockage, la *force* et la *résistance*. La récupération d'un item (lors d'une tâche de rappel ou de reconnaissance) dépendrait de la force de la trace mais pas de sa résistance. Cependant, la rapidité avec laquelle la trace décline dépendrait à la fois de sa force et de sa résistance. Ainsi, selon le modèle, le taux de déclin d'une trace suite à plusieurs essais d'apprentissage serait plus faible qu'après un unique essai d'apprentissage ; de plus, ce déclin serait d'autant plus lent que l'intervalle entre le premier et le dernier essai d'apprentissage est long. Toujours selon la théorie, la résistance de

la trace commencerait à augmenter après le premier essai d'apprentissage, et tous les incréments ultérieurs dus aux répétitions s'ajouteraient à la résistance déterminée par le premier apprentissage. Ces conceptions sont proches de celles développées dans le modèle de Pavlik et Anderson (2005, voir la section 3.2.2).

En guise de conclusion, l'effet de pratique distribuée, de par son caractère ubiquitaire, est un terrain sur lequel les traditions psychologiques et neuroscientifiques peuvent se rencontrer. La psychologie pourrait notamment s'inspirer de l'intérêt porté par les neuroscientifiques sur la notion de taux d'oubli.

3. 3. 3. Reconsolidation

En parallèle des processus de consolidation, d'autres types de processus peuvent potentiellement contribuer à l'avantage des IIR longs sur les IIR courts. En particulier, on peut s'intéresser à l'ensemble des processus à l'œuvre lors de la réactivation des représentations mnésiques durant P2. Il a été suggéré que le fait de réactiver une trace mnésique après une période de consolidation initie un processus de reconsolidation qui fournit une opportunité supplémentaire de consolidation (Alberini, 2005 ; Nader, Schafe, & LeDoux, 2000 ; Sara, 2000). De même que la consolidation, la reconsolidation fait l'objet de nombreuses publications actuellement dans le champ des neurosciences. La reconsolidation se définit comme le processus par lequel un souvenir déjà consolidé, lorsqu'il est réactivé, entre à nouveau dans un état instable pendant lequel il redevient sensible aux interférences, puis subit finalement une période de reconsolidation qui le fait revenir à un état stable. On appelle « *reminder* » l'événement qui déclenche la réactivation du souvenir et la phase de labilisation (ou fragilisation) des souvenirs. Il a été montré en effet que, chez l'animal, lorsqu'on réactivait un souvenir après qu'il ait été consolidé et qu'on injectait en même temps une substance bloquant la synthèse protéique, alors le souvenir disparaissait (e.g., Lattal & Abel, 2004). On empêcherait ainsi le processus de reconsolidation. La même injection sans que le souvenir soit réactivé auparavant ne modifiait pas le souvenir.

Des études s'intéressant à la reconsolidation des souvenirs épisodiques chez l'humain ont été menées ces dernières années (Forcato, Argibay, Pedreira, & Maldonado, 2009 ; Forcato et al., 2007 ; Forcato, Rodriguez, Pedreira, & Maldonado, 2010 ; Hupbach, Hardt, Gomez, & Nadel, 2008 ; Hupbach et al., 2007 ; Hupbach, Gomez, & Nadel, 2009) et ont montré que des souvenirs épisodiques pouvaient être perturbés si on les réactivait juste avant d'apprendre un nouveau matériel similaire. Par exemple, dans l'étude de Hupbach et al. (2007, Expérience 3), des étudiants devaient retenir des objets présentés dans un panier le jour 1. Le jour 3, certains d'entre eux recevaient un reminder tandis que les autres n'en recevaient pas. Le reminder consistait à mettre les participants en présence du panier en les incitant à se souvenir de ce qu'ils avaient fait le jour 1 sans toutefois mentionner les objets. Ensuite, tous les participants apprenaient un deuxième ensemble d'objets, différent de ceux du jour 1. Les souvenirs des objets présentés le jour 1 étaient testés le jour 5. Les sujets ayant reçu le reminder ont introduit des items appris le jour 3 dans le rappel des items du jour 1, contrairement aux sujets n'ayant pas reçu le reminder. (Toutefois, la présence du reminder n'influait pas le taux de bonnes réponses des items appris le jour 1). Cette expérience suggère donc que, le jour 3, le reminder a rendu instables les souvenirs du jour 1, et ainsi les traces des items du jour 3 ont pu s'intégrer aux traces des items du jour 1. Quand il n'y avait pas de reminder, les traces des items appris les jours 1 et 3 coexistaient simplement en mémoire. De façon intéressante, l'Expérience 2 a montré que lorsque le test final avait lieu le jour 3, juste après l'apprentissage de la deuxième liste, cet effet d'interférence n'était

pas observé. Ainsi le processus de reconsolidation, en l'occurrence l'intégration des traces du second épisode aux traces du premier épisode, semble nécessiter du temps pour se mettre en place et ne s'observe pas immédiatement. D'autre part, Forcato et al. (2007) ont montré que le désavantage dû à la reconsolidation ne pouvait se produire que si les items interférents apparaissaient dans une certaine fenêtre temporelle après l'apparition du reminder. Ils ont en effet montré qu'un intervalle de 5 minutes ou de 6 heures entre le reminder et la présentation des items interférents perturbait le souvenir des items initiaux alors qu'un intervalle de 10h ne le perturbait plus. Il existerait donc une fenêtre temporelle de plusieurs heures pendant laquelle le souvenir réactivé serait labile, mais suite à cette fenêtre temporelle le souvenir ne serait plus sensible à la perturbation et pourrait subir le processus de reconsolidation.

Ces observations laissent penser que l'apprentissage avec un long IIR est plus avantageux qu'avec un court IIR car un phénomène de reconsolidation (non perturbée) se produit au moment de P2 uniquement dans le cas du long IIR. En 2009, Hupbach et al. se sont interrogés sur l'origine des interférences en répliquant le même type d'expérience qu'en 2007. Il s'est avéré que les participants ayant été mis en présence du reminder attribuaient par erreur des objets vus le jour 3 à la liste des objets vus le jour 1, mais pas l'inverse. Les auteurs, ainsi que les résultats d'autres études, suggèrent que le reminder sert à réactiver le souvenir initial afin de le mettre à jour en y ajoutant des informations nouvelles plus pertinentes. Cette hypothèse a le mérite de montrer en quoi un processus qui rend instable des souvenirs auparavant stabilisés a une utilité pour l'individu.

Ainsi, d'une façon générale, les études sur la reconsolidation suggèrent que la réactivation suite à une période de consolidation mène à des processus de réorganisation des traces mnésiques qui aboutiraient à l'amélioration mnésique, ce qui ne serait pas le cas lorsque la réactivation survient trop précocement.

3. 3. 4. Rôle du sommeil dans la consolidation

Certains auteurs estiment que les réactivations spontanées pendant le sommeil permettent la consolidation des souvenirs et le transfert des représentations mnésiques de l'hippocampe vers les régions corticales. Le sommeil peut donc potentiellement avoir un rôle dans les effets de pratique distribuée lorsque l'apprentissage se déroule sur plusieurs jours. Il a en effet été montré qu'un intervalle de sommeil qui suit un épisode d'apprentissage favorise la rétention des informations par rapport à une période d'éveil (e.g., Drosopoulos, Wagner, & Born, 2005 ; Gais, Lucas, & Born, 2006). Ellenbogen et al. (2006 ; 2009) ont montré qu'en plus d'un simple renforcement basique des traces mnésiques, le sommeil aurait une action spécifique de protection des acquisitions envers les interférences à venir. Schématiquement, après une période d'éveil, les souvenirs resteraient sensibles à l'interférence, mais après une période de sommeil ils y seraient résistants.

Au plan cérébral, un certain nombre d'études ont été réalisées chez l'humain, qui ont montré que lors du sommeil, les patterns d'activation cérébrale observés pendant l'apprentissage sont reproduits, notamment au niveau de l'hippocampe (Ji & Wilson, 2007 ; Maquet et al., 2000 ; Peigneux et al., 2004). Une des hypothèses les plus défendues actuellement est celle de la consolidation systémique : ces réactivations seraient associées au transfert des souvenirs de l'hippocampe au néocortex (dont nous avons déjà parlé dans la section 3.3.1). Il a également été émis l'hypothèse selon laquelle le sommeil permettrait une recalibration synaptique dans les régions cérébrales dans lesquelles les apprentissages de la journée ont conduit à des phénomènes de potentialisation synaptique (Tononi & Cirelli, 2006).

Un plus ample développement des effets du sommeil sur la mémoire serait hors du cadre du présent travail. Pour une revue de ces études, voir Diekelman (2010). Cependant, il est indispensable de mentionner le rôle du sommeil dans notre problématique puisque lorsque les répétitions sont réparties sur plusieurs jours, les IIR contiennent des périodes de sommeil. De plus, le facteur sommeil n'est pas présent lorsque les études se produisent sur une courte échelle de temps. Par conséquent, ce qui se produit lorsqu'un épisode de sommeil est intercalé entre deux répétitions ou entre l'apprentissage et la récupération finale n'est probablement pas l'équivalent « en plus grand » de ce qui se produit dans les études à court terme. Nous reviendrons sur ce point dans la discussion générale.

3. 3. 5. Considérations évolutionnistes

Il nous semble intéressant ici de présenter certaines conceptions qui ont pour objectif de répondre à la question du pourquoi, question moins souvent posée en science que la question du comment : Pourquoi serait-ce utile pour l'individu que les répétitions espacées dans le temps soit associées à une plus grande rétention et/ou à un taux d'oubli plus faible ? La réponse à cette question pourrait nous donner de nouvelles pistes à explorer pour étudier l'effet de la pratique distribuée. Anderson et son équipe (Anderson & Schooler, 1991 ; Anderson & Milson, 1989 ; Anderson, Tweney, Rivardo, & Duncan, 1997) se sont interrogés sur le caractère adaptatif de la mémoire humaine. Les auteurs postulaient en toute logique que le fonctionnement de la mémoire est adapté à l'environnement, et en particulier que la mémoire est sensible aux régularités statistiques présentes dans l'environnement. De plus, comme la mémoire possède une capacité limitée, alors pour pouvoir servir l'individu au mieux, certains de ses contenus doivent être effacés afin de laisser la place aux éléments plus utiles.

Pour ce qui concerne l'effet de pratique distribuée, Anderson et Schooler (1991) ont dans un premier temps tenté de rendre compte des régularités statistiques relatives à la répétition des éléments dans l'environnement. Ils ont pour cela analysé les mots apparaissant plusieurs fois dans trois corpus, dont par exemple les titres du *New York Times* pendant un an. L'idée importante était que l'occurrence d'un mot dans l'environnement - en l'occurrence le journal- doit déclencher chez le sujet -en l'occurrence le lecteur- une remémoration de souvenirs associés à ce mot. En effet, au moment de l'apparition d'un mot donné, le sujet a besoin de pouvoir réactiver des souvenirs et des connaissances liés à ce mot pour comprendre la nouvelle situation. Ainsi, pour être adapté à l'environnement, il faut pouvoir retenir les informations utiles afin de pouvoir les réutiliser plus tard, lorsqu'une nouvelle occurrence surviendra. Leur analyse de corpus a abouti à une description détaillée des moments dans le temps où les mots apparaissaient. Par exemple, le terme 'of' apparaissait très fréquemment et avec de très courts intervalles entre les occurrences. Le terme 'north' n'est apparu qu'une seule fois. Le nom de 'Qaddafi' apparaissait une vingtaine de fois, avec une dizaine d'apparitions très fréquentes pendant une dizaine de jours, puis aucune apparition pendant 60 jours, puis de nouveau une dizaine d'apparitions très fréquentes pendant une dizaine de jours.

Dans un second temps, les auteurs ont souhaité évaluer à quel point la fréquence d'occurrence d'un mot sur une période donnée permettait de prédire la probabilité d'occurrence de ce mot plus tard. Par exemple, ils ont représenté la probabilité d'occurrence d'un mot le jour 101 en fonction de sa fréquence d'apparition pendant les 100 jours précédents. De façon assez spectaculaire, cette fonction s'est révélée linéaire, c'est-à-dire que la probabilité d'occurrence était corrélée de façon directe à la fréquence d'apparition dans la période précédente. Ainsi, l'environnement présente des régularités dans les motifs

de répétition de l'information et, donc, une mémoire adaptative aurait avantage à prédire ces régularités, en faisant sorte de conserver les éléments qui ont une probabilité élevée de se présenter et d'oublier ceux qui ont une probabilité faible.

Concernant la problématique de l'effet de pratique distribuée, les auteurs ont examiné les mots qui n'ont été observés qu'à deux reprises pendant la période étudiée (il y en avait peu en réalité) puis de représenter la probabilité que ces mots apparaissent lors du jour suivant, et ce, en fonction de l'IIR et pour des DR divers. Le but était de pouvoir comparer ces courbes avec les courbes observées par Glenberg (1976, Expérience 1) qui sont présentées dans la Figure 2 dans la section 1.3.1.

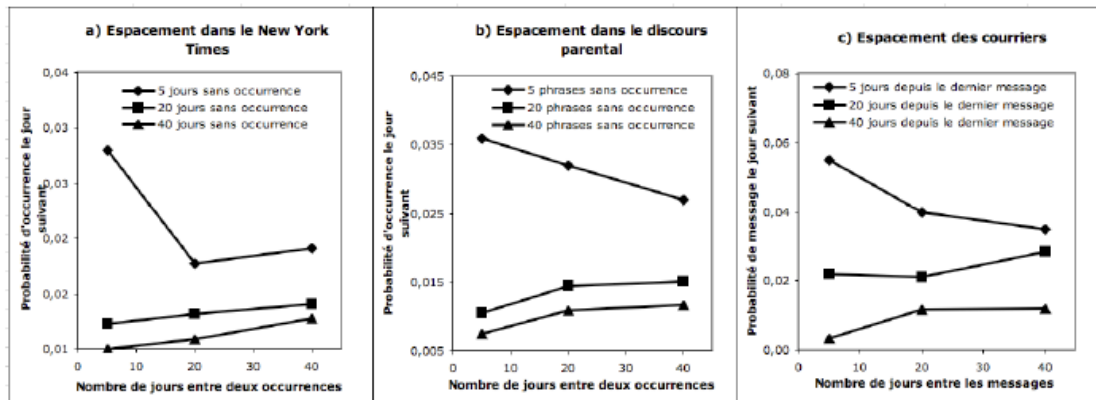


Figure 11 : Reproduction des résultats de l'étude de Anderson et Schooler (1991). En abscisses est représenté l'intervalle entre deux occurrences successives d'un item dans le corpus, et en ordonnées est représentée la probabilité d'apparition de l'item. Les trois figures correspondent à trois corpus différents, respectivement le New York Times à gauche, un corpus de langage d'enfants au centre, et un corpus de courriers électroniques à droite.

Les résultats observés sont présentés dans la Figure 9. De façon surprenante, les courbes étaient globalement similaires à celles de Glenberg (1976, Expérience 1), c'est-à-dire que pour un DR court (représenté par les courbes du haut dans les graphes), la probabilité d'occurrence était globalement élevée mais diminuait avec l'IIR. Au contraire, lorsque le DR était plus long (courbes du bas), la probabilité d'occurrence était plus faible mais elle augmentait avec l'IIR. Autrement dit, dans l'environnement, un item a tendance à se présenter suite à des intervalles constants dans la nature : si les intervalles sont en général courts entre les occurrences d'un item, alors il est probable que cet item réapparaîtra bientôt. Au contraire, si l'item ne réapparaît pas pendant un certain laps de temps, alors il est probable qu'il réapparaîtra de moins en moins. À l'inverse, un item qui se présente avec de longs IIR a peu de chance de se présenter à nouveau immédiatement après une occurrence, il faut plutôt attendre de nouveau un long DR.

Les auteurs ont ensuite étudié les courbes reflétant la probabilité d'occurrence en fonction du DR, courbes qui correspondent aux courbes d'oubli traditionnelles. Ces courbes de probabilité sont décroissantes c'est-à-dire que globalement, plus le DR augmente, moins il est probable que l'item apparaisse. Il est apparu également que les pentes des courbes de probabilité d'occurrence étaient plus importantes pour les items qui s'étaient répétés avec de courts IIR par rapport à de longs IIR. Or, c'est justement ce que l'on observe dans les performances humaines : le taux d'oubli est plus élevé pour des items répétés avec des IIR courts qu'avec des IIR longs (voir la section 3.3.2). Ainsi, les mêmes effets étaient observés entre les probabilités d'occurrences des items dans l'environnement et la rétention des items

répétés. Le fait d'observer un taux d'oubli plus faible suite à un espacement important a un sens et une valeur adaptative : un item qui s'est répété avec des grands intervalles a plus de probabilité de se présenter à nouveau après un long intervalle, il est donc utile de le retenir longtemps. Au contraire un item qui s'est répété avec un court intervalle a des chances de se présenter de nouveau rapidement, mais si ce n'est pas le cas il ne réapparaîtra probablement pas par la suite, et il peut être utile de l'oublier après un certain temps.

Pour conclure, ces observations suggèrent que le fonctionnement mnésique a évolué au cours de millions d'années d'évolution pour être le plus adapté à l'environnement afin de favoriser la survie des individus. Cette perspective donne du sens à l'observation de l'effet de pratique distribuée à la fois chez les animaux les plus rudimentaires (voir la section 3.3.1) et chez les humains. Cela suggère que des lois existent depuis très longtemps dans les systèmes mnésiques primaires, et que, lorsque la mémoire déclarative est progressivement apparue, les mêmes lois ont été transposées à ce nouveau système (e.g., Squire & Kandel, 2002). Ces considérations vont dans le sens de l'idée que des processus basiques, de type neurophysiologiques, sont à l'œuvre dans les effets de pratique distribuée, ce qui n'empêche toutefois pas que d'autres processus, d'ordre psychologique par exemple, tels que la variabilité de l'encodage ou la récupération en phase d'étude soient à l'œuvre également. Ces considérations évolutionnistes sont intéressantes en elles-mêmes, mais peuvent également donner lieu à des nouvelles pistes pour la recherche ou à de nouvelles conceptions. Comme nous l'avons vu dans la section 3.2.3, ce type de réflexion a inspiré le modèle mathématique de Mozer et al. (2009) qui semble constituer une simulation convaincante des effets de pratique distribuée.

Conclusion et problématique

Dans les trois premiers chapitres, nous avons décrit un panorama des effets de l'intervalle inter-répétition sur les performances mnésiques (chapitre 1), ainsi que des hypothèses proposées dans la littérature pour expliquer ces effets, d'une part dans le champ de la psychologie cognitive (chapitre 2) et d'autre part dans le champ de la neuroscience cognitive (chapitre 3). L'exploration de la littérature a révélé que relativement peu de travaux ont porté sur l'effet de l'agencement temporel de plusieurs répétitions d'une information sur une longue échelle de temps, c'est-à-dire sur plusieurs jours. De plus, les conclusions de ces études étaient mitigées, certaines montrant la supériorité de l'agencement expansif (Gay, 1973 ; Tsai, 1927), d'autres l'absence de différence entre les agencements (Cull, 2000). A contrario, un certain nombre d'études récentes ont étudié l'effet de l'agencement sur une courte échelle de temps, c'est-à-dire sur quelques heures au plus (Balota & al., 2006 ; Karpicke & Roediger, 2007 ; Logan & Balota, 2008). Ces études ont globalement montré la supériorité de l'agencement uniforme, et ce d'autant plus que le délai de rétention était long. Ces constatations nous ont amenée à définir les objectifs suivants. Premièrement, il s'agissait de poursuivre les recherches antérieures sur l'effet de l'agencement à long terme afin de tester la généralité de la supériorité de l'agencement expansif sur les agencements uniforme et contractant mise en évidence par Tsai (1927) et Gay (1973). Cette question a fait l'objet des Expériences 1, 2, et 3. Deuxièmement, il s'agissait d'obtenir des éléments d'explication aux effets observés. C'est l'objet de l'Expérience 2, dans laquelle nous avons testé l'hypothèse de la récupération en phase d'étude. De façon quelque peu annexe à la problématique principale, nous avons également testé l'hypothèse de la variabilité de l'encodage dans l'Expérience 4. Troisièmement, il s'agissait d'explorer l'effet du délai

de rétention, facteur qui ne s'est pas révélé très important dans les études sur l'effet de l'agencement temporel sur une longue échelle de temps, mais qui pourtant a montré son importance dans les effets de pratique distribuée.

Deuxième partie : Expériences

Introduction

Au sein de la vaste littérature relative à l'effet de la pratique distribuée (chapitre 1), nous avons constaté qu'un sous-domaine de cette problématique a été relativement peu étudié et a, de plus, fait l'objet de résultats relativement contradictoires. Il s'agit de la question de l'effet de l'agencement à long terme des répétitions d'une même information (Cull, 2000 ; Gay, 1973 ; Tsai, 1927). Etant donné la nature relativement écologique de ce type d'étude et les applications possibles qui en découlent, notamment dans le domaine de l'éducation, nous avons choisi d'explorer cette problématique dans le présent travail.

Nous nous sommes focalisée sur la question suivante : pour un nombre donné de présentations d'un même item verbal, pour une durée donnée de la période d'apprentissage, et pour un apprentissage délibéré du matériel, quel sera l'agencement temporel des répétitions le plus efficace pour la performance mnésique ayant lieu après un délai de rétention donné ? Nous avons choisi de comparer, comme dans Tsai (1927), les trois types suivants d'agencement : 1) Uniforme, c'est-à-dire que les répétitions sont séparées par des intervalles successifs de même durée ; 2) Expansif, c'est-à-dire que les répétitions sont séparées par des intervalles successifs de durée croissante ; 3) Contractant, c'est-à-dire que les répétitions sont séparées par des intervalles successifs de durée décroissante. Les expériences de Tsai ont montré que l'agencement expansif permettait les meilleures performances finales.

Nous avons également souhaité apporter des éléments explicatifs à ces effets à la lumière des théories psychologiques existantes, notamment la théorie de la variabilité de l'encodage (e.g., Glenberg, 1979) et la théorie de la récupération en phase d'étude (e.g., Johnston & Uhl, 1976), développées dans le chapitre 2.

De nombreuses variables expérimentales ont été déterminées de façon relativement arbitraire. En l'occurrence, le nombre de répétitions d'un même item, la durée totale d'apprentissage, la durée du délai de rétention, et les agencements eux-mêmes ont été fixés en partie par commodité en terme d'expérimentation (pour éviter, par exemple, d'expérimenter les samedi et dimanche). Nous n'avons pas souhaité répliquer directement les agencements utilisés dans les études précédentes, préférant élargir le spectre des observations de cet effet afin d'en tester la généralité.

Dans les quatre expériences, un matériel verbal original a été utilisé. Dans l'optique de mimer une situation d'apprentissage écologique tout en permettant un contrôle du matériel expérimental, nous avons créé pour l'occasion des items simulant du vocabulaire d'une langue étrangère. Il s'agissait d'un ensemble de paires associant un mot à un pseudomot, ensemble conçu dans le but d'être homogène. D'authentiques paires de vocabulaire de langue étrangère auraient également pu être utilisées, mais l'emploi de matériel nouveau annulait la possibilité d'interférence éventuelle avec les connaissances antérieures des participants. Les détails de l'élaboration du matériel sont précisés dans la section 4.1.

Dans les trois premières expériences, la variable agencement était une variable intra-sujets. La raison initiale de ce design expérimental était pratique. Etant donné le

caractère contraignant de la procédure pour les participants, l'utilisation d'un tel design nous permettrait de réduire la taille de nos échantillons, par rapport à une situation dans laquelle un groupe de participants serait associé à un agencement. De plus, ce choix s'est révélé judicieux a posteriori car la variabilité dans les performances des participants se sont révélées très importantes, rendant l'observation d'un effet significatif plus facile que dans le cas d'un design inter-sujets.

Ainsi, l'Expérience 1 avait pour principe de comparer trois agencements comportant quatre occurrences de la même paire mot-pseudomot réparties sur sept journées d'apprentissage consécutives. L'agencement uniforme consistait en une présentation des paires lors des jours 1, 3, 5, et 7. L'agencement expansif consistait en une présentation des paires lors des jours 1, 2, 3, et 7. L'agencement contractant consistait en une présentation des paires lors des jours 1, 5, 6, et 7. Le test de rétention était une tâche de rappel indicé qui avait lieu le 9ème jour. Cette expérience avait pour particularité que la phase d'apprentissage se déroulait à domicile, le participant étudiant les items par le biais d'Internet. Le test final avait cependant lieu au laboratoire.

L'Expérience 2 constituait une réplication de l'Expérience 1 intégralement en laboratoire. Des tâches mnésiques ont été réalisées lors du 9ème jour en plus du rappel indicé, en l'occurrence une tâche de rappel libre des pseudomots et une tâche de jugement du nombre d'occurrences des paires. De plus, une tâche de reconnaissance des paires avait lieu en parallèle de l'apprentissage, dans le but d'explorer les processus de récupération en phase d'étude. Enfin des questionnaires de motivation étaient inclus, afin d'éliminer l'hypothèse d'une baisse de la motivation au fil des sessions expérimentales qui aurait pu être responsable de l'avantage de l'agencement expansif.

Pour ces deux premières expériences, nos hypothèses étaient que l'agencement expansif serait associé à des performances plus élevées que les deux autres agencements. Dans l'Expérience 2, nous prédisions également que les paires induisant les taux les plus élevés de reconnaissance pendant la phase d'apprentissage soient les paires les mieux rappelées lors du test final, c'est-à-dire les paires étudiées selon l'agencement expansif.

L'Expérience 3 visait à explorer la généralité de l'effet de l'agencement temporel en allongeant la durée de la phase d'apprentissage, qui durait alors 13 jours, tout en réduisant le nombre d'occurrence de chaque paire à trois. Il s'agissait également de tester si la durée du délai de rétention modulerait éventuellement l'effet de l'agencement temporel, puisqu'il a été montré que la durée du délai de rétention module l'effet de l'intervalle inter-répétition dans le cas où deux occurrences ont lieu (Cepeda et al., 2008). Comme dans l'Expérience 2, des questionnaires de motivation étaient inclus pour tester l'hypothèse alternative d'une motivation décroissante au fil des sessions.

Pour l'Expérience 3, nos hypothèses étaient que l'agencement expansif serait associé aux performances les plus élevées par rapport aux deux autres agencements, et que cet avantage serait accentué avec l'augmentation du délai de rétention. En effet, si l'agencement expansif est celui qui favorise le mieux la consolidation mnésique, alors son efficacité devrait être plus durable et plus résistante au temps par rapport aux autres agencements.

Dans l'Expérience 4, nous nous sommes quelque peu éloignée de l'étude de l'agencement temporel, puisque nous avons pour objectif de tester la validité de la théorie de la variabilité de l'encodage. Le protocole expérimental consistait en la présentation d'une paire lors de deux occurrences séparées par un intervalle d'une durée fixe d'environ 40 minutes. Un groupe réalisait, pendant cet intervalle inter-répétition, un ensemble d'activités

variées, simulant une fluctuation accélérée du contexte, tandis qu'un second groupe réalisait en parallèle une activité monotone, simulant une fluctuation ralentie du contexte. La théorie de la variabilité de l'encodage prédit que le premier groupe devrait présenter de meilleures performances lors du rappel indicé final se produisant après 30 minutes de délai de rétention.

Chapitre 4 : Effet de l'agencement temporel des répétitions à long terme sur la mémoire

[...] an ichthyologist [...] proclaimed that he actively avoided remembering the names of the students he taught. [...] He explained it simply: "Remember a name, forget a fish." This is admittedly a rather simplified view of memory, but it highlights the important point that committing an event to long-term memory is serious business. Typically, the encoding of a lasting memory entails considerable practice and rehearsal [...] Carew (1996)

Dans ce chapitre, nous décrivons trois expériences dont l'objectif était de tester l'effet de l'agencement temporel des répétitions sur plusieurs jours. Ces expériences utilisaient un ensemble original de paires associant un mot et un pseudomot.

4. 1. Création du matériel expérimental

Bien que dans la littérature des paires de vocabulaire étranger réel soient souvent utilisées (e.g., Bahrick et al., 1993 ; Cepeda et al., 2008), nous avons choisi d'utiliser du matériel artificiel pour pouvoir le contrôler, étant donné que nous allions avoir besoin de plusieurs listes. Un ensemble de 96 paires mot-pseudomot a été conçu pour les Expériences 1 et 2, puis six paires ont été ajoutées à cet ensemble pour l'Expérience 3, et un sous-ensemble de ces paires a été utilisé dans l'Expérience 4. Dans cette section nous décrivons comment nous avons élaboré ce matériel dans le but que ce dernier soit le plus homogène possible.

4. 1. 1. Création des pseudomots

Les pseudomots étaient composés de six lettres et comportaient une alternance de consonnes et de voyelles sous la forme CVCVCV. Certaines diphtongues étaient exclues : 'ou', 'ch', 'gr', 'an', etc, ainsi que les lettres les moins fréquentes dans le lexique : h, x, y, z, w, k, q. Les 96 pseudomots ont été conçus afin de ne pas se ressembler entre eux. Parmi les 96 paires mot-pseudomot, certaines étaient destinées à être apprises et testées, c'est-à-dire les paires cibles, tandis que les autres étaient destinées à être apprises mais non testées, c'est-à-dire les paires distractrices.

Les pseudomots cibles ont été constitués en respectant les règles suivantes :

- aucun pseudomot cible ne terminait par la lettre E, la présence de la lettre E en fin de mot réduisant le nombre de syllabes à deux au lieu de trois (par exemple, LATOMU versus LATOME) ;
- une même lettre n'était pas présente deux fois au sein d'un pseudomot cible ;
- aucun pseudomot cible ne correspondait à un mot existant dans une autre langue ;

- aucun pseudomot ne donnait un mot existant quand on remplaçait une de ses lettres par une autre lettre de l'alphabet.

4. 1. 2. Sélection des mots

Les mots appartenant aux paires cibles ont été sélectionnés afin de posséder une fréquence lexicale ainsi qu'une valence d'imagerie élevées et homogènes (Bonin et al., 2003). Les paramètres de valence émotionnelle et de concrétude ont également été pris en compte. Ces valeurs sont présentées dans l'Annexe n°1. Les mots appartenant aux paires cibles étaient concrets et répartis en quatre catégories sémantiques : animal, partie du corps, objet manufacturé, et fruit. Les mots appartenant aux paires distractrices étaient également répartis en 12 catégories sémantiques, parmi lesquelles on retrouvait :

- les quatre catégories des paires cibles, avec des mots cependant moins fréquents que ceux appartenant aux paires cibles ;
- quatre autres catégories d'objets concrets (légumes, par exemple) ;
- quatre catégories de mots abstraits (disciplines scolaires, par exemple).

Les paires distractrices sont présentées dans l'Annexe n°2.

4. 1. 3. Association entre les mots et les pseudomots

L'association entre les mots et les pseudomots constituant les paires a été réalisée de façon à éviter autant que possible un lien éventuel entre le mot et le pseudomot. Un prétest a été réalisé pour s'assurer qu'aucune paire mot-pseudomot cible ne pouvait faire l'objet d'une association mnémotechnique évidente qui aurait biaisé l'apprentissage. Le prétest est détaillé dans la section 4.1.4.

Les paires cibles ainsi constituées ont ensuite été réparties en trois listes comportant huit paires chacune, c'est-à-dire deux exemplaires de chacune des quatre catégories sémantiques. La séparation en trois listes (L1, L2 et L3) a servi au contrebalancement en carré latin dans les Expériences 1, 2, et 3. Les paires cibles sont représentées dans le Tableau 1.

Tableau 1 : L'ensemble des 30 paires mot-pseudomot cibles classées en fonction de leur catégorie sémantique et de leur liste.

Catégorie	Liste 1		Liste 2		Liste 3	
Animal	VACHE	CAJESI	CHEVAL	FASOLU	CHAT	SANUMI
	SOURIS	JAMIDO	CHIEN	POCUGA	OISEAU	RUTOLA
Partie du corps	MAIN	TAGOBU	PIED	LUTISA	OEIL	DERUPA
	BRAS	FIRECO	NEZ	MURONA	JAMBE	TODEGI
Objet	STYLO	COBEMI	CHaise	PUSIRO	TELEVISEUR	VAMELO
	LAMPE	MECUSI	REFRIGERATEUR	JOTUNA	TABLE	DAPIGU
Fruit	FRAISE	VUMOGI	BANANE	NOSAPI	POMME	SONIJA
	RAISIN	GEFIPA	POIRE	BAFISO	ABRICOT	LATOMU
Instrument de musique	PIANO	GOVARU	GUItARE	BUMEFA	TAMBOUR	ROVEBI
	VIOLON	RATIDU	TROMPETTE	NAFECI	ACCORDEON	PIBOJA

Les trois listes possédaient des scores moyens similaires pour les paramètres de concrétude, imagerie, fréquence subjective et valence émotionnelle (Bonin et al., 2003). Ces paramètres sont présentés dans le Tableau 2.

Tableau 2 : Scores moyens des paires mot-pseudomot pour les quatre paramètres d'intérêt en fonction de leur liste.

	Concrétude	Imagerie	Fréquence subjective	Valence émotionnelle
Liste 1	4,85	4,76	3,57	3,65
Liste 2	4,84	4,79	3,64	3,48
Liste 3	4,80	4,72	3,78	3,50

4. 1. 4. Prétests réalisés sur les paires pour en mesurer l'association subjective

Au total, 36 sujets ont passé le prétest. Parmi eux, 15 étudiant(e)s ont été testés sur les 24 paires cibles excluant la catégorie musique (février 2007), et 21 étudiant(e)s sur l'ensemble des 30 paires cibles (février 2008).

Une première association entre les pseudomots et les mots a été réalisée, et ces paires ont été imprimées sur des formulaires. Pour chaque paire, les participant(e)s devaient juger sur une échelle de 0 à 10 la facilité avec laquelle ils pouvaient trouver dans le pseudomot un mot qui pouvait avoir un lien avec le mot associé, tel que 0 = impossible et 10 = évident. Un exemple était donné concernant la paire *maison* - *potrate* : par exemple *potrate* peut évoquer les mots *porte* et *portail* que l'on peut facilement relier au concept de *maison*. Les participants devaient écrire le ou les mot(s) évoqué(s) pour chaque pseudomot puis cocher le niveau de facilité avec laquelle cette association leur était venue à l'esprit entre 0 et 10.

Six paires sur les 30 paires cibles initiales ont dû être modifiées en cours de prétest car les scores étaient plutôt élevés et les mots évoqués étaient souvent les mêmes. Finalement, les 30 paires finales ont été jaugées entre 0,20 et 2,80 / 10 pour la facilité avec laquelle elles évoquaient un autre mot proche, avec cependant une paire avec un score plus élevé, CHAT-SANUMI (4,27).

Les scores moyens d'association pour chaque liste de paires cibles étaient les suivants : 1,92, 1,78, et 2,52 respectivement pour les listes 1, 2, et 3. Ces scores d'association étaient relativement faibles et peu différents entre les listes. Des analyses statistiques plus poussées étaient impossibles car tous les groupes de participants n'ont pas été testés sur exactement les mêmes paires.

4. 2. Expérience 1

4. 2. 1. Introduction

L'étude de la littérature a mis en évidence que peu d'études ont porté sur l'effet de l'agencement temporel des répétitions sur plusieurs jours, et ce, malgré les applications directes possibles pouvant découler de ces études (voir la section 1.4.2). L'objectif de cette première expérience était donc de combler cette lacune de données empiriques de mettant en place un protocole sur le même principe que celui mis en place dans les études de Tsai (1927, Expériences 2 et 3) et de Cull (2000, Expériences 3 et 4). L'Expérience 1 avait donc pour objectif de comparer l'effet de trois types d'agencement temporel des répétitions. Concrètement, nous avons fixé -de façon relativement aléatoire- la durée de la phase d'apprentissage à une semaine, et la durée du délai de rétention à deux jours. Les items à apprendre étaient des paires associant un mot à un pseudomot, tels que décrits dans la section 4.1. Chaque paire était présentée à quatre reprises. La tâche finale était une tâche

de rappel indicé. L'agencement uniforme consistait en des présentations ayant lieu tous les deux jours (X-X-X-X selon la nomenclature décrite dans la section 1.4.2) ; l'agencement expansif consistait en des présentations fréquentes en début de phase d'apprentissage et plus rares à la fin (XXX---X) ; l'agencement contractant consistait en des présentations rares en début de phase d'apprentissage et plus fréquentes à la fin (X---XXX). Pour des raisons pratiques, la phase d'apprentissage se déroulait par l'intermédiaire d'Internet, tandis que la tâche de rappel indicé se déroulait en laboratoire.

Les hypothèses dérivait de la logique de la théorie de la récupération en phase d'étude (voir la section 2.2) : parmi les trois agencements, l'agencement de type expansif est celui qui, logiquement, devrait permettre de favoriser le plus les processus de récupération en phase d'étude, du fait des intervalles inter-répétition (IIR) de durée croissante. En effet, les courts IIR du début de la phase d'apprentissage devraient permettre aux occurrences répétées des items de pouvoir réactiver les souvenirs des occurrences antérieures. À contrario, dans les agencements uniforme et contractant les longs IIR initiaux devraient favoriser le déclin des souvenirs des items et, par conséquent, une moins grande probabilité de pouvoir les récupérer au moment des répétitions. Par conséquent, les performances à la tâche de rappel indicé finale devraient être les plus élevées pour l'agencement expansif que pour les deux autres agencements.

L'hypothèse de la variabilité de l'encodage (voir la section 2.1), qui est la théorie concurrente à celle de la récupération en phase d'étude, n'induit pas a priori de prédiction logique concernant la supériorité d'un agencement sur un autre, étant donné que l'intervalle inter-répétition moyen était identique dans les trois agencements, c'est-à-dire 48h.

4. 2. 2. Méthode

Participants

Trente-six étudiant(e)s, dont 6 hommes et 30 femmes, âgé(e)s entre 18 et 26 ans (âge moyen 21 ans) et de langue maternelle française ont participé à cette étude. Leur vue était jugée satisfaisante à excellente (à l'état normal ou corrigé). Chaque participant ayant terminé l'étude a bénéficié d'un dédommagement de 10 euros.

Les participants étaient recrutés par le biais de petites annonces déposées sur le campus de Lyon 2 à Bron ou sur Internet. Aucun participant n'avait pris part aux prétests. Ils n'étaient pas étudiants en Psychologie cognitive ou Sciences cognitives au-delà de la 2ème année de licence. Un échange de mails a été nécessaire avec de nombreux étudiants pour l'explication du principe de l'expérience. En plus de l'échantillon final de 36 participants, vingt-deux participants supplémentaires ont abandonné l'expérience en cours, une participante a terminé ses sessions après que les 36 participants requis dans le plan aient terminé et n'a donc pas été incluse dans l'échantillon, et deux participants supplémentaires ont réalisé une tâche différente pour le test final (ce qui nous a permis d'ajuster cette tâche).

Les participants étaient aléatoirement assignés à un groupe de contrebalancement parmi A, B, ou C.

Appareils

Pendant la phase d'apprentissage, un serveur Apple Macintosh Xserve doté de 4 Go de RAM était utilisé pour diffuser les stimuli. Les participants utilisaient leur ordinateur personnel pour se connecter au serveur via une connexion haut débit.

Pendant la phase de test, les stimuli étaient présentés par un ordinateur Apple Macintosh de type MacBook doté de 1,7 Go de RAM au moyen du logiciel Psyscope (Cohen, McWhinney, Flatt, & Provost, 1993). Un clavier périphérique était utilisé pour l'enregistrement des réponses.

Matériel

Un ensemble de 96 paires associant un mot et un pseudomot a été utilisé, dont 24 paires cibles (appartenant aux catégories sémantiques *animal*, *partie du corps*, *objet*, et *fruit*) et 72 paires distractrices. Les 24 paires cibles étaient divisées en trois listes L1, L2, et L3 qui ont été présentées dans la section 4.1.3.

Chaque participant recevait une feuille de route rappelant les étapes pour se connecter au site internet, les dates prévues pour les sessions et un tableau pour inscrire les codes de fin de session. La feuille de route est présentée dans l'Annexe n°3.

Design expérimental

La phase d'apprentissage se déroulait pendant sept jours consécutifs comportant chacun une session. Chaque paire cible était présentée quatre fois aux participants, selon l'un des trois agencements temporels suivants :

- dans l'agencement Uniforme (X-X-X-X), une paire donnée était présentée les jours 1, 3, 5, et 7.
- dans l'agencement Expansif (XXX---X), une paire donnée était présentée les jours 1, 2, 3, et 7.
- dans l'agencement Contractant (X---XXX), une paire donnée était présentée les jours 1, 5, 6, et 7.

L'agencement temporel étant une variable intra-sujets, chaque participant était confronté aux trois agencements, et ce, au cours de la même période d'apprentissage. Chaque participant était également confronté aux trois listes de paires, de sorte que les paires d'une liste se répétaient selon agencement différent de ceux d'une autre liste. Par exemple, pour un tiers des participants, la liste L1 était présentée selon l'agencement Uniforme, la liste L2 selon l'agencement Expansif et la liste L3 selon l'agencement Contractant. La liste de paires assignée à un agencement temporel donné était contrebalancée sur l'ensemble des participants selon un plan en carré latin. Le contrebalancement est illustré dans le Tableau 3.

Tableau 3 : Répartition des présentations des listes (L1, L2, et L3) au cours des sessions d'apprentissage (1 à 7) en fonction des groupes de contrebalancement (A, B, ou C) et de l'agencement temporel (Uniforme, Expansif, ou Contractant). La lettre *d* symbolise la présentation de paires distractrices.

Groupe	Agencement	Session						
		1	2	3	4	5	6	7
A	Uniforme	L1	<i>d</i>	L1	<i>d</i>	L1	<i>d</i>	L1
	Expansif	L2	L2	L2	<i>d</i>	<i>d</i>	<i>d</i>	L2
	Contractant	L3	<i>d</i>	<i>d</i>	<i>d</i>	L3	L3	L3
B	Uniforme	L2	<i>d</i>	L2	<i>d</i>	L2	<i>d</i>	L2
	Expansif	L3	L3	L3	<i>d</i>	<i>d</i>	<i>d</i>	L3
	Contractant	L1	<i>d</i>	<i>d</i>	<i>d</i>	L1	L1	L1
C	Uniforme	L3	<i>d</i>	L3	<i>d</i>	L3	<i>d</i>	L3
	Expansif	L1	L1	L1	<i>d</i>	<i>d</i>	<i>d</i>	L1
	Contractant	L2	<i>d</i>	<i>d</i>	<i>d</i>	L2	L2	L2

Dans chaque session d'apprentissage, 24 paires étaient présentées successivement. Dans ce but, des paires distractrices étaient ajoutées afin de combler les lacunes laissées par l'absence de paires cibles dans les sessions d'apprentissage. Les paires distractrices étaient des paires non répétées et non testées, de nature similaire aux paires cibles. Elle sont symbolisées par la lettre *d* dans le Tableau 3.

La séquence des paires, qu'elles soient cibles ou distractrices, était pseudo-randomisée au sein de chaque session. De plus, les règles suivantes ont été respectées afin d'éviter tout effet de séquence dans l'apprentissage :

a) Deux paires appartenant au même agencement n'apparaissaient jamais l'une à la suite de l'autre ;

b) Deux paires appartenant à la même catégorie sémantique n'apparaissaient jamais l'une à la suite de l'autre ;

c) Les trois premières et les trois dernières positions sérielles dans chaque session étaient considérées comme des zones « tampons », éventuellement sujettes à un effet de primauté et de récence. Afin d'équilibrer au maximum ces effets sur les paires cibles, nous avons fait en sorte qu'une paire cible donnée se soit trouvée au maximum une fois dans l'ensemble de ces 14 zones tampon (7 sessions x 2 moments). De plus, la première et la dernière position sérielle étaient occupées par une paire distractrice, hormis pour les sessions composées uniquement de paires cibles, c'est-à-dire les sessions n°1 et n°7. En définitive, 20 paires cibles sur les 24 étaient positionnées dans cette zone tampon, et ce une seule fois.

d) Pour les trois groupes de participants, une même position sérielle d'une session donnée était occupée par une paire d'un même agencement et d'une même catégorie sémantique. Par exemple, pour les trois groupes dans la session n°1, en première position se trouvait une paire de l'agencement Uniforme représentant un objet ; en deuxième position se trouvait une paire de l'agencement Expansif représentant une partie du corps ; et ainsi de suite. D'un groupe à l'autre, seuls les items (issus des listes L1, L2, ou L3) étaient en rotation, mais les associations entre agencement et catégorie sémantique étaient conservées.

e) Une paire particulière avait une position sérielle variable au sein des différentes sessions dans lesquelles elle était répétée. Par exemple, dans le groupe A, la paire BANANE-NOSAPI se trouvait successivement à la position 13, 23, 8 puis 20.

f) D'un groupe à l'autre les mêmes paires distractrices se retrouvaient dans la même session et à la même position sérielle.

La séquence des paires présentées au groupe A est représentée dans l'Annexe n°4.

Procédure

Prétest.

Deux participants ont passé l'expérience dans les conditions réelles mais sans rémunération et avec le test final le 8ème jour (et non le 9ème jour). Il s'agissait de s'assurer du bon fonctionnement du site Internet, d'estimer le temps passé pendant chaque session, d'un entraînement à donner les consignes pour l'expérimentatrice, et d'avoir une estimation du score final. Les résultats ont été les suivants : 9 bonnes réponses sur 24 pour l'une (âgée de 57 ans) et 5 bonnes réponses sur 24 pour le second (âgé de 25 ans). Suite à ces prétests, nous avons décidé de placer le test final le 9ème jour pour obtenir un délai de rétention plus long.

Protocole principal.

L'expérience se déroulait en trois phases principales. Tout d'abord, l'expérimentatrice rencontrait le participant pour une explication des consignes, puis le participant réalisait l'expérience à distance puis, enfin, il revenait au laboratoire pour réaliser le test de rétention.

Première rencontre avec le participant.

Ce rendez-vous avait pour objectifs d'expliquer aux participants le principe de l'expérience ainsi que les consignes et la méthode d'apprentissage, de faire un essai sur un ordinateur, de donner la feuille de route nécessaire pour se connecter, de décider de la date du début de la phase d'apprentissage et de celle du test, et enfin de faire signer le formulaire de consentement. L'explication était en général individuelle mais parfois deux ou trois participants la suivaient en même temps.

L'expérimentatrice expliquait aux participants qu'ils devaient se connecter au site Internet de l'université, depuis leur domicile, une fois par jour, durant 20 minutes environ et en une seule fois, et ce pendant sept jours consécutifs, afin d'apprendre des paires mot-pseudomot. Ensuite, ils devraient revenir à l'université où ils réaliseraient un test portant sur ces paires.

On présentait aux participants un exemple de paire (i.e., CAFETIERE - TABULORI, paire non présentée ultérieurement) en leur expliquant qu'ils devaient apprendre l'association entre le mot et le pseudomot, comme si le pseudomot était la traduction du mot dans une langue imaginaire. Lors du test, un seul des deux éléments de la paire serait présenté et ils auraient à en rappeler l'élément complémentaire, et ce dans les deux sens (i.e., « CAFETIERE - ?? » ou « ?? – TABULORI »). Le sens véritable du test n'était pas précisé car nous voulions que le participant s'efforce de construire en mémoire une représentation globale du pseudomot et non pas, comme cela peut se produire lorsqu'on connaît par avance le sens du test, n'en apprendre que les premières lettres.

Trois stratégies étaient proposées aux participants afin de faciliter leur apprentissage. Premièrement, ils étaient incités à chercher dans le pseudomot un élément ayant un sens et pouvant être associé au mot. Par exemple : TABULORI contient « bu », qui évoque « café », qui est lié à CAFETIERE. Deuxièmement, les participants étaient incités à former des images mentales, par exemple imaginer « Lorie avec une cafetière sur la tête ». Troisièmement, ils étaient incités à élaborer des phrases dans lesquelles le mot était remplacé par le pseudomot (e.g., « Ma tabulori est cassée, je suis en manque de café »). Ces consignes pour faciliter la mémorisation avaient pour but d'homogénéiser les réponses entre les participants et d'éviter de potentiels effets planchers ; en effet, une expérience

du même type menée précédemment avait montré une grande variabilité des scores au test, probablement due, entre autres, aux différences de technique d'apprentissage entre les participants.

Pendant la session quotidienne, les participants auraient à décider eux-mêmes du temps passé sur chacune des 24 paires. Chaque paire apparaîtrait seule sur l'écran et, quand les participants pensaient avoir bien appris une paire, ils pourraient passer à la paire suivante en cliquant sur un lien en bas de la page, sans retour en arrière possible. Les participants étaient informés que 24 paires seraient présentées dans chaque session et que certaines paires seraient répétées au fil des sessions contrairement à d'autres qui ne seraient présentées qu'une fois. Toutes les paires devaient être apprises.

Un essai d'apprentissage était réalisé avec quatre paires (fabriquées pour l'occasion et non revues ultérieurement), lui-même suivi d'un test quelques minutes plus tard.

La procédure pour se connecter au site Internet était décrite et les mots de passe pour accéder aux sessions étaient fournis. L'expérimentatrice insistait sur le fait que les sessions devaient être suivies dans l'ordre, sans oubli, et en respectant le planning convenu afin que le délai de rétention de 48h soit respecté. Les participants étaient informés qu'il pouvaient interrompre leur participation à l'expérience à tout moment en cas de mauvais suivi des sessions ou d'impossibilité diverses. Ils pouvaient contacter l'expérimentatrice pour tout problème ou question. Le formulaire de consentement était rempli et le planning des sessions et du test était décidé.

Phase d'apprentissage.

La phase d'apprentissage avait lieu dans un endroit choisi par les participants (en général leur domicile). Pendant sept jours consécutifs et, dans la mesure du possible, sur le même ordinateur, au même endroit, et à la même heure, les participants se connectaient au site Internet indiqué. Ils visualisaient alors les paires mot-pseudomot dans le but de les apprendre selon les techniques proposées le jour de l'explication. Nous avons demandé aux participants d'être au calme pendant les sessions d'apprentissage (sans téléphone, sans dérangement), de s'installer confortablement, et de se munir de leurs lunettes si besoin.

Avant de commencer chaque session, les participants devaient faire les réglages adéquats de leur navigateur afin d'avoir l'image des paires de mots en plein écran et sans avoir à utiliser la barre de défilement vertical pour atteindre le lien de changement de page.

Pendant la phase d'apprentissage, le mot et le pseudomot étaient présentés simultanément au milieu de l'écran, tel que le mot apparaissait à gauche et le pseudomot à droite (police Arial, taille 48). Les participants devaient cliquer sur un lien hypertexte situé dans le coin inférieur droit de l'écran pour passer à la paire suivante, sans intervalle entre deux pages successives. Une barre de défilement en haut de la page représentait la progression des paires afin que les participants aient une estimation du nombre de paires passées et à venir dans la session.

A la fin de chaque session, les participants devaient noter sur leur feuille de route, le code de fin de session apparaissant à l'écran, afin que l'expérimentatrice puisse s'assurer que les sessions d'apprentissage s'étaient déroulées dans l'ordre. Ils devaient aussi noter la date et l'heure de réalisation de chaque session ainsi que des commentaires éventuels. Les sessions duraient entre 10 minutes et 30 minutes, en fonction des participants et de la facilité de la session. Entre deux sessions d'apprentissage, les participants ne devaient pas volontairement repenser au vocabulaire appris.

Phase de test.

Le test avait lieu le 9ème jour, c'est-à-dire environ 48 h après la dernière session d'apprentissage. Les participants se rendaient individuellement au laboratoire pour passer le test de rappel indicé. Le logiciel Psyscope (Cohen et al., 1993) était utilisé pour la diffusion des items.

Les codes de fin de session étaient vérifiés et la consigne du test était expliquée. Vingt-cinq pseudomots, parmi ceux appris, apparaîtraient à l'écran, un par un. Les participants devaient dire à haute voix le plus rapidement possible le mot correspondant au pseudomot affiché. Il n'y avait pas de limitation de temps pour donner la réponse. Les participants étaient encouragés à répondre le plus rapidement et le plus correctement possible et à donner une réponse claire, sans hésitation. Seule la première réponse donnée était prise en compte. Les participants devaient donner une réponse au hasard quand ils ne connaissaient pas la réponse ou dire « je passe ». Dès que leur réponse était donnée, les participants devaient se préparer à l'essai suivant. La variable dépendante était le nombre de bonnes réponses.

Suite à la consigne, l'expérimentatrice s'installait en arrière du participant, munie du clavier périphérique afin d'enregistrer les réponses. Le test avait lieu et durait entre 3 et 10 minutes.

Les 24 pseudomots cibles étaient testés, avec en plus, pour le premier essai, un des pseudomots issus des paires distractrices, non pris en compte. Les paires testées ainsi que leur ordre de passation étaient les mêmes pour tous les participants quel que soit leur groupe. La séquence des stimuli respectait la règle selon laquelle ni deux mots de la même catégorie sémantique ni deux mots du même agencement temporel ne se trouvait à la suite l'un de l'autre.

Un point de fixation apparaissait au centre de l'écran pendant 500 ms, suivi d'un écran blanc de 50 ms. Le pseudomot s'affichait ensuite au centre de l'écran (police Times New Roman, taille 48) et restait affiché jusqu'à ce que les participants donnent leur réponse à voix haute. Dès que les participants prononçaient un mot, l'expérimentatrice appuyait sur une touche en fonction du type de réponse donnée (i.e., correct : touche 3, et incorrect/passe : touche 2). A ce moment apparaissait un écran blanc pendant 1 seconde puis l'essai suivant était lancé.

Ensuite, les participants remplissaient un questionnaire portant sur la phase d'apprentissage et d'autres questions comme le nombre d'heures de sommeil. Enfin ils remplissaient le reçu attestant du dédommagement financier et le but de l'expérience leur était expliqué.

Deux participants supplémentaires ont réalisé, avant le test de rappel indicé, un test de reconnaissance. Mais leurs commentaires indiquaient que ce test de reconnaissance entraînait une récupération involontaire du mot associé, cette réactivation pouvant avoir une influence sur la tâche de rappel indicé ultérieur. Nous avons donc décidé de ne pas réaliser le test de reconnaissance pour les participants suivants et n'avons pas pris en compte ces participants dans nos analyses.

4. 2. 3. Résultats

Les données utilisées pour les analyses statistiques suivantes n'ont pas été corrigées.

Le nombre de bonnes réponses par participant à la tâche de rappel indicé s'échelonnait de 4 à 24 sur 24, c'est-à-dire de 16,7 % à 100 % de bonnes réponses pour tous les agencements confondus. La moyenne s'élevait à 17,25 bonnes réponses sur 24, c'est-à-

dire 71,9 % de bonnes réponses. Quatre participants ont donné 100 % de bonnes réponses. Les scores moyens sont représentés dans la Figure 12.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant). Cette analyse a révélé un effet significatif de l'agencement temporel : le nombre de bonnes réponses était significativement différent en fonction de l'agencement temporel des répétitions : $F(2,70) = 3.40$, $p < .04$, $\eta^2_p = 0.09$.

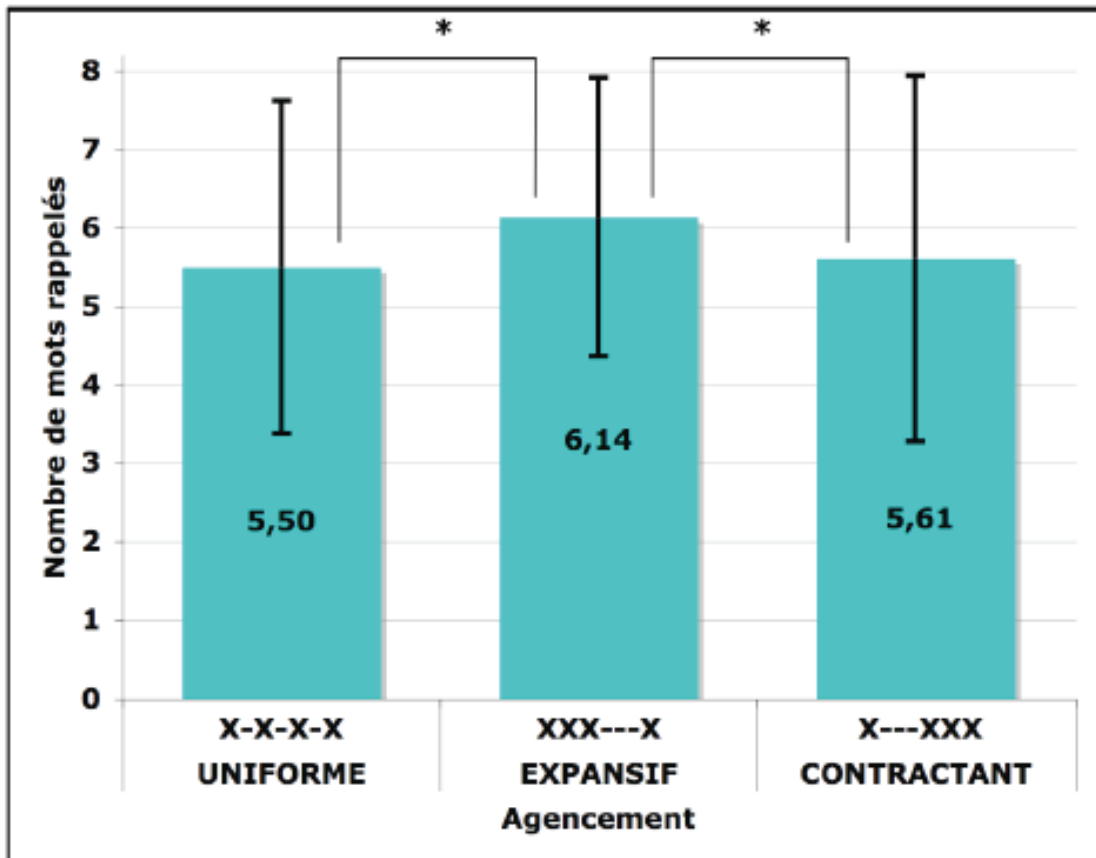


Figure 12 : Nombre moyen de mots rappelés par agencement (Uniforme, Expansif, Contractant) dans la tâche de rappel indicé. Le nombre maximum de bonnes réponses par agencement était de 8. Les barres d'erreurs représentent les écart-types et les astérisques représentent les contrastes significatifs.

Des analyses de contraste ont révélé que les participants produisaient plus de bonnes réponses dans la condition expansive (6,14) que dans la condition uniforme (5,50), $F(1,70) = 5.83$, $p < .02$, et qu'ils produisaient également plus de bonnes réponses dans la condition expansive que dans la condition contractante (5,61), $F(1,70) = 3.98$, $p = .05$. Il n'y avait pas de différence significative entre les conditions uniforme et contractante, $F(1,70) < 1$.

Des analyses secondaires ont été réalisées dans le but de vérifier l'absence d'effets non souhaités, en particulier l'effet de liste et l'effet de groupe.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujets la liste de paires (L1, L2, L3). Cette analyse n'a pas révélé d'effet

significatif de la liste de paires : $F(2,70) < 1$. Ce résultat confirme le fait que les listes étaient homogènes en difficulté, comme nous pensions les avoir construites à l'aide des prétests.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur inter-sujets le groupe de contrebalancement (A, B, C). Cette analyse n'a pas révélé d'effet significatif du groupe de contrebalancement : $F(2,33) < 1$. Ce résultat confirme le fait que les groupes de contrebalancement étaient homogènes en termes de performance pour le rappel indicé.

4. 2. 4. Discussion

Ainsi, l'hypothèse relative à la théorie de la récupération en phase d'étude a été vérifiée car l'agencement expansif a produit les meilleures performances finales, suivi par l'agencement uniforme puis par l'agencement contractant. Les résultats de Tsai (1927 ; voir la section 1.4.2) ont donc été globalement répliqués. Une des limites de cette expérience est cependant le fait que les participants aient réalisé leurs sessions d'apprentissage sur Internet, et donc sans que nous ayons pu contrôler les conditions expérimentales. Ainsi, malgré les précautions mises en place, il est possible que certains participants n'aient pas réalisé l'expérience de façon sérieuse, ce qui aurait pu avoir un impact sur les résultats obtenus. C'est pourquoi l'Expérience 2 a été réalisée, qui constitue une réplification de l'Expérience 1 en laboratoire. Les résultats de ces deux expériences seront discutés ensemble dans la section Discussion de l'Expérience 2.

4. 3. Expérience 2

4. 3. 1. Introduction

L'objectif de l'Expérience 2 était de répliquer l'Expérience 1 en laboratoire, afin d'instaurer d'un contrôle plus important des conditions expérimentales, contrôle qui n'était pas assuré lors des passations via Internet. La procédure expérimentale était similaire à celle de l'Expérience 1 à l'exception de quelques modifications.

Certaines modifications ont par ailleurs été apportées au protocole. Afin d'obtenir des données sur les processus en jeu lors de la phase d'apprentissage, en particulier sur les processus de récupération en phase d'étude, une tâche de reconnaissance a été introduite pendant les sessions d'apprentissage. Ainsi, à la suite de chaque présentation d'item, les participants avaient pour consigne d'indiquer si, selon eux, l'item avait déjà été présenté auparavant.

Par ailleurs, de nouvelles tâches de mémoire explicite ont été introduites lors de la session de test, en plus de celle de rappel indicé : une tâche de rappel libre des pseudomots et une tâche de jugement de fréquence. Ces tâches ont été mises en place dans le but d'obtenir un plus grand nombre de données à l'issue de l'apprentissage. D'autre part, la tâche de jugement de fréquence a été introduite suite à la constatation que certains participants de l'Expérience 1, au moment de l'explication de l'expérience, étaient étonnés d'apprendre que les paires avaient toutes été présentées quatre fois pendant l'apprentissage. Enfin, les participants devaient également évaluer leur niveau de motivation au début et à la fin de chaque session d'apprentissage, et ce dans l'optique suivante. Si l'on part du principe que le niveau de motivation peut induire une efficacité d'encodage plus ou moins importante, alors la supériorité éventuelle de l'agencement expansif pourrait s'expliquer par des effets de motivation s'il s'avérait que le niveau de motivation est significativement plus élevé en début de période d'apprentissage qu'à la fin. La motivation

serait alors plus élevée au moment où la majorité des occurrences concerne l'agencement contractant.

Finalement, les participants étaient invités à revenir deux mois après le début de l'expérience pour une session d'explication des résultats. Nous avons profité de cette session pour leur faire réaliser une tâche de rappel indicé dont ils n'étaient pas informés, dans le but d'explorer les conséquences à long terme des effets de l'agencement temporel.

Les hypothèses étaient similaires à celles de l'Expérience 1. Selon l'hypothèse de la récupération en phase d'étude, l'agencement expansif devrait mener à des performances finales plus élevées que les autres agencements, et ce, dans les trois tâches de mémoire. Par ailleurs, les taux de reconnaissance des paires lors de la phase d'apprentissage, qui sont censés refléter la récupération des traces des occurrences antérieures, devraient également être plus élevés pour les paires présentées selon l'agencement expansif. En effet, les courts IIR du début de l'apprentissage devraient favoriser la reconnaissance. D'autre part, les taux de motivation devraient rester constants au fil des sessions. À J61, la supériorité de l'agencement expansif devrait également être observée. De nouveau, l'hypothèse de la variabilité de l'encodage ne fait pas de prédiction particulière.

4. 3. 2. Méthode

Participants

Trente-huit étudiant(e)s, dont 6 hommes et 32 femmes, âgé(e)s de 18 à 24 ans (âge moyen 20 ans) ont participé à cette étude. Les participants étaient de langue maternelle française sauf une participante polonaise qui parlait le français couramment. Leur vue était jugée satisfaisante à excellente (à l'état normal ou corrigé). Chaque participant ayant terminé l'étude a bénéficié d'un dédommagement de 50 euros. Ils n'étudiaient pas la psychologie cognitive ou les sciences cognitives au-delà de la 3ème année et aucun d'entre eux n'avait participé auparavant à une étude sur la même problématique ni à un prétest. Ils ont été répartis dans les groupes de contrebalancement A, B, C en fonction de leur disponibilité horaire. Les effectifs de chaque groupe de participants étaient de 11, 15, et 12, respectivement pour les groupes A, B, et C.

Les participants étaient recrutés par le biais de petites annonces papier ou par Internet. Des échanges d'emails ont eu lieu avec un grand nombre d'étudiants pour leur expliquer le principe de l'expérience et fixer les dates et horaires de passation. Lors de la première session, 53 participants étaient attendus, huit d'entre eux ne se sont pas présentés, et sur les 45 participants qui ont participé à la première session, 7 n'ont pas pu suivre toutes les sessions et ont donc été supprimés de l'échantillon.

Par ailleurs, l'expérimentatrice s'est assurée que chaque participant a suivi chaque session lors de la plage horaire correspondant à son groupe, ce qui assurait que chaque participant étudiait la liste d'items adéquate.

Appareils

Les stimuli étaient présentés par un ordinateur Apple Macintosh de type MacBook doté de 1,7 Go de RAM au moyen du logiciel Psyscope (Cohen et al., 1993). Les informations étaient projetées sur un écran à l'aide d'un vidéoprojecteur.

Matériel

Les 96 paires mot-pseudomot utilisées dans l'Expérience 1 ont été réutilisées, qui comportaient, pour rappel, 24 paires cibles et 72 paires distractrices. Les 24 paires cibles étaient réparties en trois listes L1, L2, L3 de manière identique à l'Expérience 1.

Des formulaires de reconnaissance ont été fournis à chaque participant lors de chaque session d'apprentissage. Ce formulaire est présenté dans l'Annexe n°8 et contenait :

- en haut de page, un questionnaire de motivation (sur une échelle de 1 à 7).
- en milieu de page, le questionnaire de reconnaissance dans lequel chaque ligne, seulement numérotée, correspondait à chacune des 24 paires présentées pendant la session. Deux cases apparaissaient sur chaque ligne : « vu » ou « pas vu » suivies de trois autres cases pour le jugement du degré de confiance : Pas sûr, Moyennement sûr, ou Sûr.
- en bas de page, un questionnaire de motivation (sur une échelle de 1 à 7) et des remarques éventuelles (facile/difficile, technique de mémorisation utilisée...).

Des carnets étaient élaborés pour les réponses des participants au test final. Chaque carnet comportait 25 pages. La première page était dédiée à l'inscription du nom du participant. Les pages suivantes étaient numérotées de 1 à 24, comportaient une ligne pour inscrire le mot réponse et, en dessous, un questionnaire de degré de confiance avec trois cases à cocher (Pas sûr, Moyennement Sûr, Sûr).

Des formulaires de rappel libre et de jugement de fréquence étaient également fournis. Le formulaire de rappel libre était composé d'un tableau à deux colonnes avec des cases numérotées de 1 à 44. Le formulaire de jugement de fréquence était composé de 39 lignes numérotées associées à chaque fois à une échelle de chiffres allant de 0 à 7.

Design expérimental

Le même design expérimental que dans l'Expérience 1 était utilisé. Les seules modifications apportées ont été les positions sérielles des items au sein des sessions ainsi qu'une tentative de mieux contrôler les effets de récence et de primauté.

Ainsi, une première zone « tampon » a été définie comme constituée de l'ensemble comprenant le premier et le dernier item de chaque session et ce pour toutes les sessions ($7 \times 2 = 14$ emplacements). Dans cette zone tampon, le nombre de paires appartenant à chaque agencement temporel a été rendu égal. Ensuite une deuxième zone tampon comprenait les deux premiers et les deux derniers items de la session ($7 \times 4 = 28$ emplacements) : le même nombre d'items de chaque agencement temporel était placé dans cette zone. Le même contrôle a été réalisé pour la troisième zone tampon, composée des trois premiers et trois derniers items de la session (42 emplacements) et pour la quatrième zone tampon, composée des quatre premiers et des quatre derniers items de la session (56 emplacements). De plus, nous nous sommes assurée que les agencements temporels étaient bien équilibrés entre la partie de début de liste et de fin de liste pour une zone tampon donnée. Enfin, nous avons essayé de limiter au maximum la répétition d'un item donné au sein des zones tampon.

La séquence des paires présentées au groupe A est représentée dans l'Annexe n°5.

Procédure

Principe général.

Chaque participant s'est rendu dans une salle de cours, toujours la même, durant sept jours consécutifs pour la phase d'apprentissage d'une durée de 20 minutes environ chaque jour. Le 9ème jour, chaque participant s'est rendu dans une autre salle pour participer au test portant sur les paires apprises pendant la phase d'apprentissage, et ce pour une durée de 30 minutes environ.

Les sessions d'apprentissage et de test ont été réalisées par groupe de 2 à 14 personnes. Les paires étaient projetées sur un tableau blanc par vidéoprojection.

Les sept sessions d'apprentissage avaient lieu lors de sept jours consécutifs, et pour un groupe donné, à la même heure de la journée. Les 38 participants ont suivi leur apprentissage la même semaine. Une journée expérimentale comprenait six créneaux horaires correspondant à deux créneaux pour chacun des trois groupes A, B, et C. Il était possible pour un membre d'un groupe donné de venir exceptionnellement à l'autre créneau horaire de ce groupe. Quelques sessions ont été rajoutées pour des étudiants seuls afin d'éviter la perte de participants.

Consignes.

Lors des sessions d'apprentissage, rien ne devait être laissé sur la table à part un stylo, les téléphones devaient être éteints, et les participants devaient mettre leurs lunettes si besoin. Chaque jour les participants devaient s'asseoir à la même place. Trois rangs d'étudiants étaient constitués, ceux-ci se trouvaient donc entre 2 et 4 mètres de l'écran.

Lors de la première session d'apprentissage, après remplissage du formulaire de consentement, le principe de l'apprentissage était expliqué aux participants. Il s'agissait exactement des mêmes consignes que dans l'Expérience 1, c'est-à-dire que le sens du test final n'était pas précisé et qu'on les incitait à utiliser trois stratégies d'apprentissage. La première était basée sur la détection d'éléments ayant du sens dans le pseudomot, la seconde sur l'élaboration d'une image mentale, et la troisième sur la création et la répétition mentale d'une phrase contenant le pseudomot. Ces conseils d'apprentissage étaient re-précisés au début des sessions 1 à 4.

Lors de chacune des sessions, 24 paires mot-pseudomot seraient présentées. Les paires défileraient de façon automatique toutes les 30 secondes. Il était annoncé que certaines paires seraient répétées au cours de la semaine, contrairement à d'autres qui n'apparaîtraient qu'une seule fois.

De surcroît à l'apprentissage des paires, les participants devaient juger, pour chacune d'entre elles, si ils l'avaient déjà vu auparavant, c'est-à-dire s'ils constataient qu'elle se répétait ou si elle leur semblait être nouvelle. Ce jugement devait être réalisé à la fin des 30 secondes de présentation consacrées à l'apprentissage, suite à l'apparition de la question « *Aviez-vous déjà vu cette paire auparavant ?* » à l'écran. Les participants devaient cocher sur leur feuille la case « Vu » ou « Pas vu » correspondant à leur réponse, à la ligne adéquate, ainsi que le degré de confiance pour cette réponse sur une échelle comprenant trois possibilités, « Pas Sûr », « Moyennement Sûr », ou « Sûr ». Il a été répété plusieurs fois aux participants qu'il fallait répondre avec le plus d'honnêteté possible, en suivant son impression, et sans regarder la réponse des participants voisins. D'autre part, il fallait répondre au formulaire de reconnaissance quand la question apparaissait à l'écran et non pas avant. Il était également précisé qu'un pseudomot donné était toujours associé au même mot, et réciproquement.

Il était interdit d'écrire les mots ou les pseudomots pendant les sessions et en dehors des sessions. De façon générale, l'expérimentatrice insistait sur le fait que les participants

devaient éviter, dans la mesure du possible, de réviser volontairement les paires en dehors des phases d'apprentissage et éviter de parler avec les autres participants de l'expérience. Suite aux consignes, l'expérimentatrice répondait aux éventuelles questions et la phase d'apprentissage du Jour 1 commençait.

Phase d'apprentissage.

La feuille de reconnaissance était distribuée aux participants. Ceux-ci devaient inscrire leur nom et juger leur niveau de motivation pour cette session sur une échelle de 1 (pas du tout motivé) à 7 (extrêmement motivé).

Quand les participants avaient rempli leur questionnaire de motivation et qu'ils étaient prêts, la phase d'apprentissage commençait. Au centre de l'écran, projeté au tableau, apparaissaient les paires mot-pseudomot pendant 30 secondes chacune, avec le mot à gauche et le pseudomot à droite (police Arial Narrow, caractère gras, taille 55, hauteur des caractères apparents : 5 cm environ). Une croix de fixation centrale apparaissait ensuite pendant 500 ms. La même paire était ensuite présentée à nouveau au même emplacement, avec en haut de l'écran le numéro correspondant à la position sérielle et, au-dessous, la question « *Aviez-vous déjà vu cette paire auparavant ?* ». Cette question restait apparente pendant 7 secondes, pendant lesquelles les participants devaient cocher leur réponse et le degré de confiance correspondant. Ensuite, une croix de fixation centrale apparaissait pendant 2 secondes puis l'essai suivant commençait.

Après l'apparition des 24 paires (phase qui avait duré 16 minutes), les participants répondaient au questionnaire de motivation pour la poursuite de l'expérience et écrivaient d'éventuels commentaires sur le niveau de difficulté, la technique de mémorisation utilisée, et autres. Les feuilles de réponse étaient ramassées et les participants pouvaient quitter la salle.

Entre deux sessions d'apprentissage, les participants ne devaient pas repenser aux paires apprises. Ils ne devaient pas non plus échanger entre eux sur le contenu de l'expérience.

Phase de test.

Une session de test avait lieu le 9ème jour et était annoncée des participants, alors qu'une autre session de test avait lieu le 61ème jour et n'était pas annoncée aux participants.

Lors de J9.

Deux sessions expérimentales de test ont eu lieu le 9ème jour, c'est-à-dire 48h après la 7ème session, et ce dans une salle différente de la salle où l'apprentissage avait eu lieu. Les deux sessions étaient identiques entre elles et étaient organisées pour des raisons de disponibilité horaire des participants. Les participants du premier créneau horaire ne devaient pas dévoiler aux participants du second créneau le contenu de la phase de test. Les participants étaient assis à environ 2 à 5 mètres du tableau. Ils devaient éteindre leur téléphone, mettre leur lunettes, garder sur la table qu'un stylo, et garder le silence jusqu'à la toute fin de la session. Trois tests ont été réalisés successivement : un test de rappel indicé, un test de rappel libre des pseudomots, et un test de jugement du nombre d'occurrences des paires.

Tâche de rappel indicé.

Un carnet de réponse était distribué à chaque participant. Ceux-ci y écrivaient leur nom en première page. Les consignes étaient énoncées.

Une croix de fixation centrale apparaissait pendant 4 secondes. Puis un pseudomot apparaissait seul au centre de l'écran pendant 10 secondes (police Arial Narrow, caractère gras, taille 55, hauteur des caractères apparents : 5 cm environ). En haut de l'écran apparaissait simultanément le numéro correspondant à la page du carnet, de 1 à 24. Les participants devaient écrire le mot correspondant au pseudomot, puis cocher la case correspondant à leur degré de confiance en leur réponse. A la fin des 10 secondes retentissait un bip sonore afin de signaler le moment précis de tourner la page du carnet, et seulement à ce moment-là. Il était interdit de revenir en arrière une fois la page tournée. Les participants étaient incités à répondre au hasard s'ils ne connaissaient pas la réponse.

Les 24 pseudomots cibles étaient testés. Les pseudomots testés ainsi que leur ordre de passation étaient les mêmes pour tous les participants. La séquence des stimuli respectait la règle selon laquelle ni deux mots de la même catégorie sémantique ni deux mots du même agencement temporel ne se trouvaient à la suite l'un de l'autre. La séquence des items était différente de celle de l'Expérience 1.

Les variables dépendantes étaient le nombre de bonnes réponses ainsi que le degré de confiance associé aux bonnes réponses.

Tâche de rappel libre des pseudomots.

Suite à la tâche de rappel indicé, les carnets de réponse étaient ramassés et les feuilles de rappel libre distribuées à l'envers. La consigne était d'écrire lisiblement le plus grand nombre de pseudomots seuls (i.e., sans le mot associé) et le plus correctement possible. Les participants retournaient alors leur feuille et inscrivaient leur nom, et les 4 minutes de rappel libre commençaient. Deux minutes environ s'étaient écoulées entre la fin du test de rappel indicé et le début du rappel libre. Suite au rappel libre, les feuilles de réponse étaient ramassées.

Tâche de jugement de fréquence des présentations.

Les formulaires de jugement de fréquence étaient distribués. A l'écran étaient projetées les paires, et les participants devaient entourer sur le formulaire, pour chacune d'elles, le chiffre correspondant au nombre de fois où ils pensaient avoir vu cette paire au cours des sept sessions de la phase d'apprentissage.

Parmi les 39 paires présentées se trouvaient les 24 paires cibles ainsi que 15 paires distractrices. Ces 15 paires distractrices étaient issues des sessions d'apprentissage n°2 à n°6. Comme dans les tâches précédentes, deux paires successives ne faisaient pas partie du même agencement temporel ni de la même catégorie sémantique.

Suite à la tâche de jugement de fréquence, les participants recevaient le dédommagement de 50 euros.

Phase de test de J 61.

Les participants étaient conviés à une séance d'explication de l'expérience 61 jours après le début de la phase d'apprentissage. Ils n'en avaient pas été avertis mais une tâche de rappel indicé avait lieu ce jour-là. Quatorze participants (3 du groupe A ; 5 du groupe B ; 6 du groupe C) sont revenus. La procédure était identique à celle de la tâche de rappel indicé du 9ème jour, à l'exception de l'ordre des items qui était modifié.

4. 3. 3. Résultats

Les données utilisées pour les analyses statistiques suivantes n'ont pas été corrigées.

Rappel indicé à J9

Le nombre total de bonnes réponses par participant à la tâche de rappel indicé s'échelonnait de 4 à 24 sur 24, c'est-à-dire de 16,7 % à 100 % de bonnes réponses. La moyenne s'élevait à 17,63 bonnes réponses sur 24, c'est-à-dire 73,5 % de bonnes réponses. Quatre participants ont donné 100 % de bonnes réponses. Ces paramètres descriptifs sont très similaires à ceux de l'Expérience 1. Les scores moyens par agencement sont présentés dans la Figure 13.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujets l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant). Cette analyse a révélé un effet significatif de l'agencement temporel : le nombre de bonnes réponses était significativement différent en fonction de l'agencement temporel des répétitions : $F(2,74) = 4.03$, $p < .03$, $\eta^2_p = 0.10$.

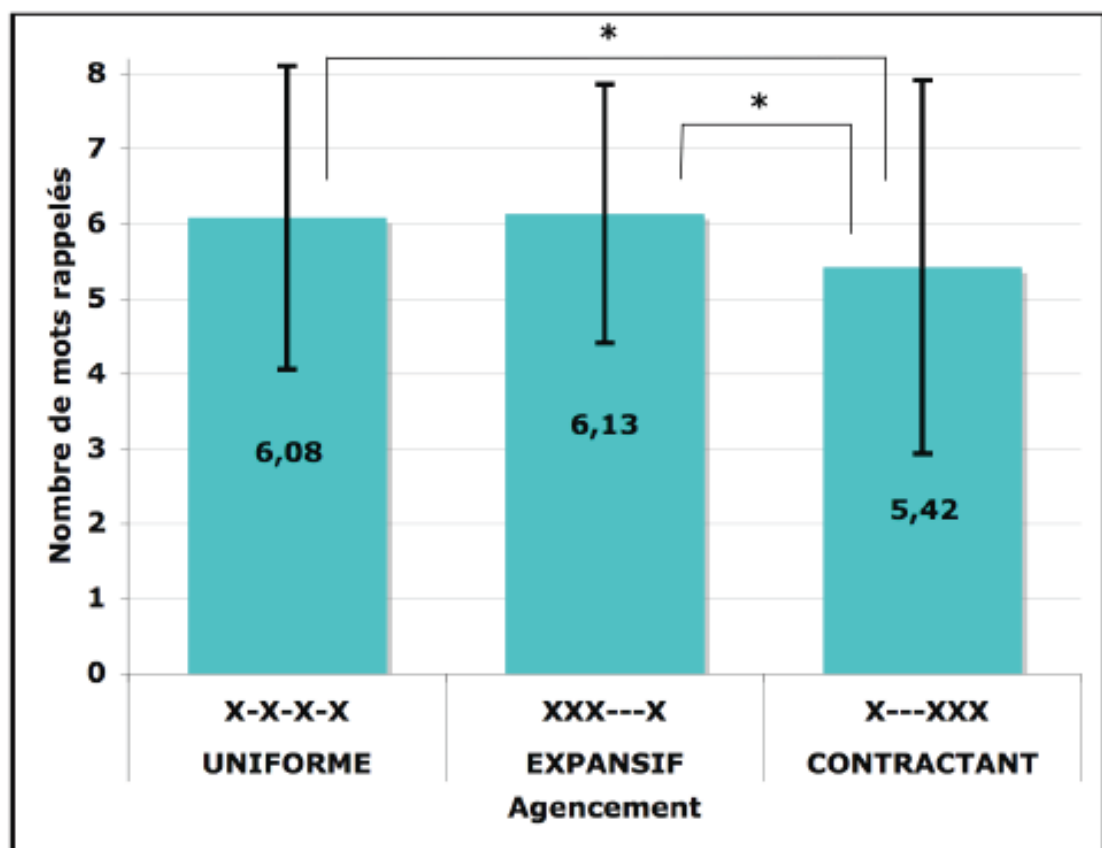


Figure 13 : Nombre moyen de mots rappelés par agencement (Uniforme, Expansif, Contractant) dans la tâche de rappel indicé à J9. Le nombre maximum de bonnes réponses par agencement était de 8. Les barres d'erreurs représentent les écart-types et les astérisques représentent les contrastes significatifs.

Des analyses de contraste ont révélé que les participant produisaient plus de bonnes réponses dans la condition expansive (6,13) que dans la condition contractante (5,42), $F(1,74) = 6.49$, $p < .02$, et qu'ils produisaient également plus de bonnes réponses dans la condition uniforme (6,08) que dans la condition contractante, $F(1,74) = 5.56$, $p < .03$. Il n'y avait pas de différence significative entre les conditions expansive et uniforme, $F(1,74) < 1$.

Des analyses secondaires ont été réalisées dans le but de vérifier l'absence des effets de liste et de groupe.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujets la liste de paires (L1, L2, L3). Cette analyse n'a pas révélé d'effet significatif de la liste de paires : $F(2,74) < 1$. Ce résultat confirme le fait qu'il n'y avait pas de différence de performance en fonction des listes, tout comme dans l'Expérience 1.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur inter-sujets le groupe de contrebalancement (A, B, C). Cette analyse n'a pas révélé d'effet significatif du groupe de contrebalancement : $F(2,35) = 1.04$, $p < .37$. Ce résultat confirme le fait que les groupes de contrebalancement étaient homogènes en termes de performance pour le rappel indicé, tout comme dans l'Expérience 1.

Le degré de certitude associé aux réponses correctes données par les participants, qui comportait trois modalités (Sûr, Moyennement sûr, Pas sûr), a été également analysé. Parmi les réponses correctes données, le pourcentage de réponse de niveau « Sûr » par rapport au total des réponses correctes a été comparé entre les agencements. Il s'élevait à 84,9 % pour l'agencement uniforme, 84,5 % pour l'agencement expansif et 79,0 % pour l'agencement contractant. Une analyse de variance a été réalisée sur ce pourcentage avec comme facteur intra-sujets l'agencement (uniforme, expansif, contractant). (Chaque cellule de l'analyse contenait le pourcentage des réponses correctes associées à un niveau « Sûr » sur le total des réponses correctes pour un sujet donné et un agencement donné.) Cette analyse n'a pas révélé d'effet significatif de l'agencement : $F(2,74) = 1.53$, $p < .23$. Ainsi, le degré de certitude associé aux réponses correctes n'était pas significativement différent entre les agencements.

Rappel libre des pseudomots à J9

Seuls les pseudomots cibles parfaitement orthographiés étaient considérés comme des bonnes réponses.

Le nombre total de pseudomots rappelés par participant s'échelonnait de 2 à 21 sur 24, c'est-à-dire de 8,3 % à 87,5 % de pseudomots rappelés. La moyenne s'élevait à 11,95 pseudomots rappelés sur 24, c'est-à-dire 49,8 %. Les scores moyens par agencement sont présentés dans la Figure 14.

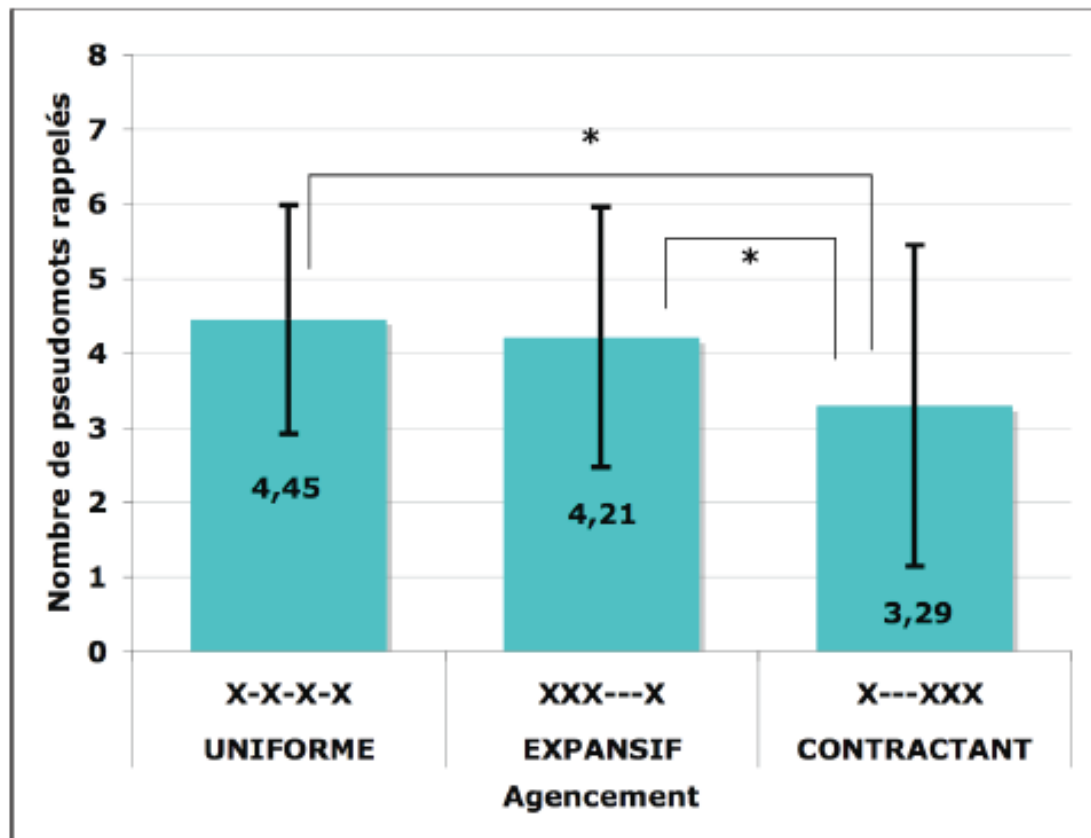


Figure 14 : Nombre moyen de pseudomots rappelés par agencement (Uniforme, Expansif, Contractant) dans la tâche de rappel libre à J9. Le nombre maximum de bonnes réponses par agencement était de 8. Les barres d'erreurs représentent les écart-types et les astérisques représentent les contrastes significatifs.

Une analyse de variance a été réalisée sur le nombre de pseudomots rappelés avec comme facteur intra-sujets l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant). Cette analyse a révélé un effet significatif de l'agencement temporel : le nombre de pseudomots rappelés était significativement différent en fonction de l'agencement temporel des répétitions : $F(2,74) = 7.80, p < .001, \eta^2_p = 0.17$.

Des analyses de contraste ont révélé que les participants produisaient plus de bonnes réponses dans la condition expansive (4,45) que dans la condition contractante (3,29), $F(1,74) = 13.97, p < .001$, et qu'ils produisaient également plus de bonnes réponses dans la condition uniforme (4,21) que dans la condition contractante, $F(1,74) = 8.84, p < .01$. Il n'y avait pas de différence significative entre les conditions expansive et uniforme, $F(1,74) < 1$. Ces résultats sont similaires aux résultats obtenus pour la tâche de rappel indicé.

Jugement de fréquence des paires à J9

La fréquence estimée s'échelonnait de 1 à 7 toutes conditions confondues avec une estimation moyenne de 4,17 occurrences. Les scores moyens par agencement sont présentés dans la Figure 15.

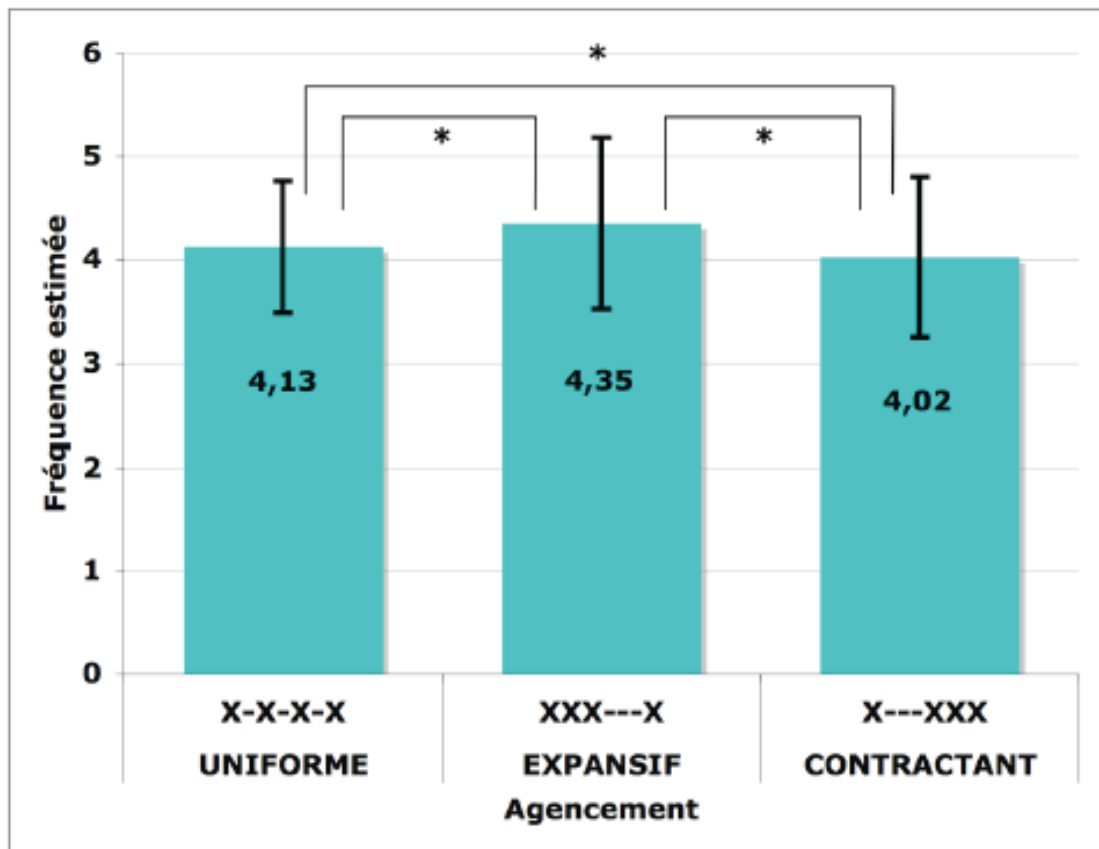


Figure 15 : Nombre moyen d'occurrences estimé par agencement (Uniforme, Expansif, Contractant) dans la tâche de jugement de fréquence à J9. Le nombre réel d'occurrences était de 4. Les barres d'erreurs représentent les écart-types et les astérisques représentent les contrastes significatifs.

Une analyse de variance a été réalisée sur la fréquence estimée avec comme facteur intra-sujet l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant). (Chaque cellule de l'analyse contenait la fréquence estimée moyenne des 8 items d'un agencement donné, et ce pour un participant donné.) Cette analyse a révélé un effet significatif de l'agencement temporel : la fréquence d'occurrence estimée était significativement différente selon l'agencement temporel des répétitions : $F(2,74) = 9.55$, $p < .001$, $\eta^2_p = 0.21$.

Des analyses de contraste ont révélé que les participants estimaient que les paires présentées selon l'agencement expansif (4,35) avaient été présentées un plus grand nombre de fois que celles présentées selon l'agencement uniforme (4,13), $F(1,74) = 8.68$, $p < .01$. Les participants estimaient également que les paires présentées selon l'agencement expansif avaient été présentées un plus grand nombre de fois que celles présentées selon l'agencement contractant (4,02), $F(1,74) = 18.23$, $p < .001$. Enfin, les participants estimaient également que les paires présentées selon l'agencement uniforme avaient été présentées un plus grand nombre de fois que celles présentées selon l'agencement contractant (4,02), $F(1,74) = 18.23$, $p < .001$.

Reconnaissance des paires pendant la phase d'étude

Lors de la quatrième présentation d'une paire cible donnée (P4), c'est-à-dire lors de J7 pour chacun des agencements, tous les participants ont reconnu 100 % des paires comme ayant déjà été présentées. Lors de la troisième présentation d'une paire cible donnée (P3), que ce soit à J3 en condition expansive, à J5 en condition uniforme, ou à J6 en condition contractante, seuls deux participants n'ont pas reconnu 100 % des paires comme ayant déjà été présentées. Le nombre de fois où une paire n'a pas été reconnue s'élevait à 3 sur 912 événements au total, ce qui correspondait à 0,33 % des paires. On peut donc considérer que, globalement, lors de la troisième et de la quatrième présentation des paires, l'intégralité des paires était reconnue et qu'il y avait un effet plafond de la reconnaissance dans ces deux cas. La problématique de la récupération en phase d'étude se pose donc essentiellement pour la deuxième présentation.

Lors de la deuxième présentation d'une paire donnée (P2), c'est-à-dire à J2 en condition expansive, à J3 en condition uniforme, et à J5 en condition contractante, le nombre de paires reconnues par participant pour une condition d'agencement s'échelonnait de 0 à 8 avec une moyenne de 7,21 sur 8, c'est-à-dire 90,1 %. Les taux de reconnaissance moyens par agencement sont présentés dans la Figure 16.

Une analyse de variance a été réalisée sur le nombre de paires reconnues lors de P2 avec comme facteur intra-sujets l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant). Cette analyse a révélé un effet significatif de l'agencement temporel : le nombre de paires reconnues à P2 était significativement différent selon l'agencement temporel des répétitions : $F(2,74) = 5.87, p < .01, \eta^2_p = 0.14$.

Des analyses de contraste ont révélé que les participants reconnaissaient davantage de paires lorsque celles-ci étaient présentées dans l'agencement expansif (7,74) par rapport à l'agencement uniforme (7,03), $F(1,74) = 6.93, p = .01$. Ils reconnaissaient également davantage de paires lorsque celles-ci étaient présentées dans l'agencement expansif par rapport à l'agencement contractant (6,87), $F(1,74) = 10.35, p < .01$. Cependant ils ne reconnaissaient pas davantage de paires lorsque celles-ci étaient présentées dans l'agencement uniforme par rapport à l'agencement contractant, $F(1,74) < 1$.

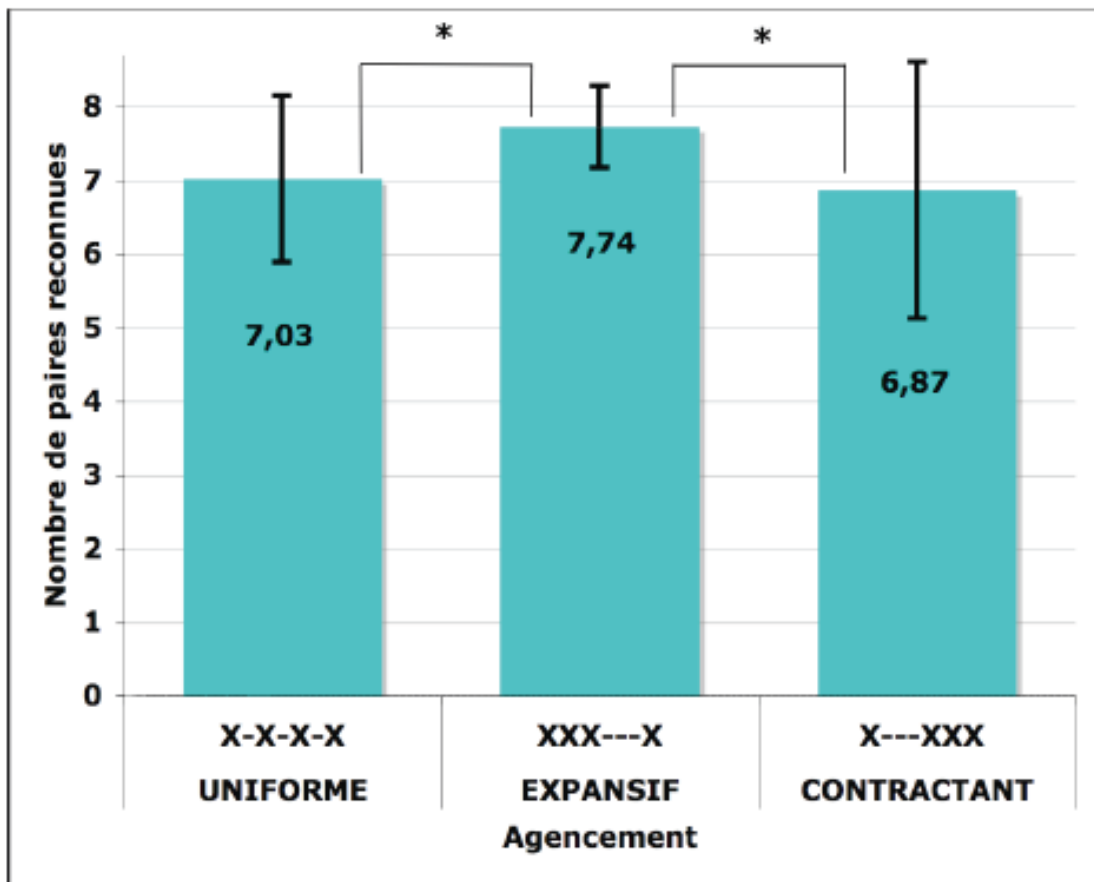


Figure 16 : Nombre moyen de paires reconnues lors de P2 en fonction de l'agencement (Uniforme, Expansif, Contractant). Le nombre maximum de reconnaissances correctes par agencement était de 8. Les barres d'erreurs représentent les écart-types et les astérisques représentent les contrastes significatifs.

Dans cette analyse, chaque modalité du facteur agencement est confondue avec la durée de l'intervalle entre P1 et P2. Ainsi ces résultats reflètent le fait que les participants reconnaissent d'autant mieux les paires présentées pour la deuxième fois que l'intervalle depuis la première présentation est court. Ainsi le taux de reconnaissance est plus élevé lorsque cet intervalle est d'1 jour (agencement expansif) par rapport à des intervalles de 2 ou 4 jours (respectivement, les agencements uniforme et contractant) sans différence significative entre ces deux derniers intervalles.

De nouveau, des analyses secondaires ont été réalisées dans le but de vérifier l'absence des effets de liste et de groupe. Une analyse de variance a été réalisée sur le nombre de paires reconnues à P2 avec comme facteur intra-sujets la liste de paires (L1, L2, L3). Cette analyse n'a pas révélé d'effet significatif de la liste de paires : $F(2,74) = 1,92$, $p < .16$. Ce résultat confirme le fait qu'il n'y avait pas de différence de reconnaissance des paires selon les listes d'items, comme souhaité.

Une analyse de variance a été réalisée sur le nombre de paires reconnues à P2 avec comme facteur inter-sujets le groupe de contrebalancement (A, B, C). Cette analyse n'a pas révélé d'effet significatif du groupe : $F(2,35) < 1$. Ce résultat confirme le fait que les groupes de contrebalancement étaient homogènes en termes de performance pour la reconnaissance à P2, comme souhaité.

Le degré de certitude associé à la reconnaissance des items lors de P2, qui comportait trois modalités (Sûr, Moyennement sûr, Pas sûr), a été également analysé. Parmi les items reconnus, le pourcentage de réponse de niveau « Sûr » par rapport au total des autres réponses a été comparé entre les agencements. Il s'élevait à 81,0 % pour l'agencement uniforme, 89,5 % pour l'agencement expansif et 79,6 % pour l'agencement contractant. Une analyse de variance a été réalisée sur ce pourcentage avec comme facteur intra-sujets l'agencement (Uniforme, Expansif, Contractant). (Chaque cellule de l'analyse contenait le pourcentage des items reconnus avec le niveau « Sûr » sur le total des items reconnus pour un sujet donné et un agencement donné.) Cette analyse a révélé un effet significatif de l'agencement : $F(2,74) = 5.43, p < .01, \eta^2_p = 0.13$. Des analyses de contraste ont révélé que les participants estimaient plus souvent leur niveau de certitude comme étant « Sûr » lorsque, à P2, ils reconnaissaient des paires issues de l'agencement expansif par rapport à l'agencement uniforme ($F(1,74) = 6.85, p = .01$) et par rapport à l'agencement contractant ($F(1,74) = 9.26, p < .01$). Les participants n'estimaient pas plus souvent leur niveau de certitude comme étant « Sûr » lorsqu'ils reconnaissaient des paires issues de l'agencement uniforme par rapport à l'agencement contractant, $F(1,74) < 1$. Ces résultats reflètent le simple fait que plus le temps passe entre P1 et P2, plus les taux de reconnaissance de niveau « Sûr » sont faibles.

Lien entre reconnaissance des paires pendant la phase d'étude et performances au rappel indicé

De façon globale, 76 % des items reconnus à P2 ont fait ultérieurement l'objet d'une réussite au rappel indicé, alors que seulement 52 % des items non reconnus à P2 ont fait ultérieurement l'objet d'une réussite au rappel indicé. Le Tableau 4 présente la répartition des items selon le succès ou l'échec aux tâches de reconnaissance et de rappel indicé, pour tous sujets confondus. Un test de chi-2 d'indépendance réalisé sur ces données révèle un effet significatif ($\chi^2(1) = 21.92, p < .001$), ce qui traduit le fait que l'issue du rappel indicé ultérieur est différente selon que l'item ait été reconnu ou non lors de P2.

Tableau 4 : Répartition des essais selon l'issue de la tâche de reconnaissance à P2 (succès ou échec) et de la tâche de rappel indicé à J9 (succès ou échec) pour tous les participants et tous les items confondus.

Reconnaissance à P2	Rappel indicé		<i>Total</i>
	Succès	Échec	
Succès	623	199	822
Échec	47	43	90
<i>Total</i>	670	242	912

Des analyses supplémentaires sont impossibles car le nombre d'items non reconnus à P2 est nul pour un grand nombre de participants, ce qui rend impossible à la fois le calcul des pourcentages et le calcul de chi-2.

Estimation de motivation pendant la phase d'apprentissage

Afin de déterminer si le niveau de motivation diminuait ou augmentait au fil des sessions, une analyse de tendance a été réalisée d'une part sur les niveaux de motivation estimée en début de session, et d'autre part sur les niveaux de motivation estimée en fin de session. Les niveaux de motivation moyens en fonction des sessions d'apprentissage sont représentés dans la Figure 17.

Motivation en début de session, de la 2ème à la 7ème session.

Les scores de motivation estimés lors du début de la première session n'ont pas été inclus dans l'analyse car ils étaient considérés comme non représentatifs étant donné que le participant n'avait pas encore réalisé de session d'apprentissage. Les scores s'échelonnaient de 1 à 7 avec une moyenne de 5,22. L'analyse de tendance polynomiale n'a révélé aucune tendance significative : pour la tendance linéaire, $F(1,185) = 1.12$, $p < .30$; pour la tendance quadratique, $F(1,185) = 1.11$, $p < .30$, pour les tendances cubique, quartique, et d'ordre 5, $F(1,185) < 1$ dans tous les cas. Ces résultats montrent qu'il n'y a pas d'évolution significative de la motivation pour réaliser l'apprentissage entre les sessions 2 et 7.

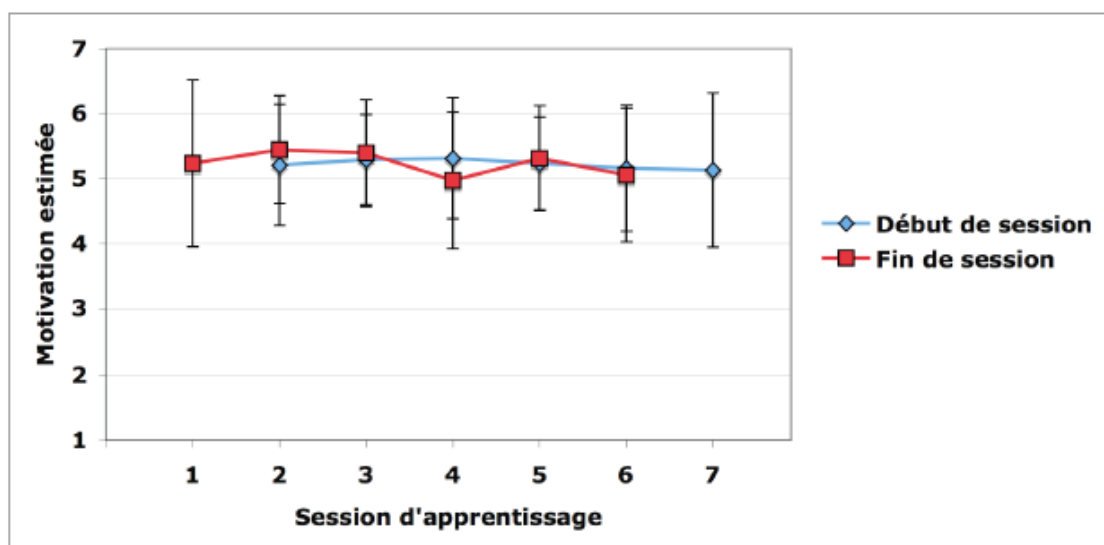


Figure 17 : Niveaux moyens de motivation estimée par les participants en fonction des sessions d'apprentissage (J1 à J7) pour commencer la session (courbe bleue) et pour poursuivre l'expérience (courbe rouge). Les barres d'erreurs représentent les écart-types.

Motivation en fin de session, de la 1ère à la 6ème session.

Les scores de motivation estimés à la fin de la septième session n'ont pas été inclus dans l'analyse car ils étaient considérés comme non représentatifs étant donné que le participant venait de finir son apprentissage. Les scores s'échelonnaient de 1 à 7 avec une moyenne de 5,23. L'analyse de tendance polynomiale a révélé trois tendances significatives parmi les cinq calculées : pour la tendance linéaire, $F(1,185) = 4.39$, $p < .04$; pour la tendance quadratique, $F(1,185) < 1$; pour la tendance cubique, $F(1,185) = 1.63$, $p = .20$; pour la tendance quartique, $F(1,185) = 5.84$, $p < .02$; et pour la tendance d'ordre 5, $F(1,185) = 5.68$, $p < .02$.

La tendance linéaire significative indique une évolution des scores au fil des sessions. L'observation des moyennes dans la Figure 17 indique que les niveaux de motivation, lorsqu'ils sont estimés à la fin de la session, diminuaient au fil des sessions.

Les tendances quadratiques et d'ordre 5 sont à la fois délicates (cf. Howell, 2006, chapitre 12) et peu pertinentes à analyser ici.

Rappel indicé à J61

L'appartenance des 14 participants aux groupes de contrebalancement était répartie de la façon suivante : 3 participants issus du groupe A, 5 du groupe B, 6 du groupe C.

Le nombre total de bonnes réponses par participant s'échelonnait de 8 à 24 sur 24, c'est-à-dire de 33 % à 100 % de bonnes réponses. La moyenne s'élevait à 17,07 bonnes réponses sur 24, c'est-à-dire 71,1 % de bonnes réponses. Un participant a donné 100 % de bonnes réponses. Les scores moyens par agencement sont présentés dans la partie droite de la Figure 18.

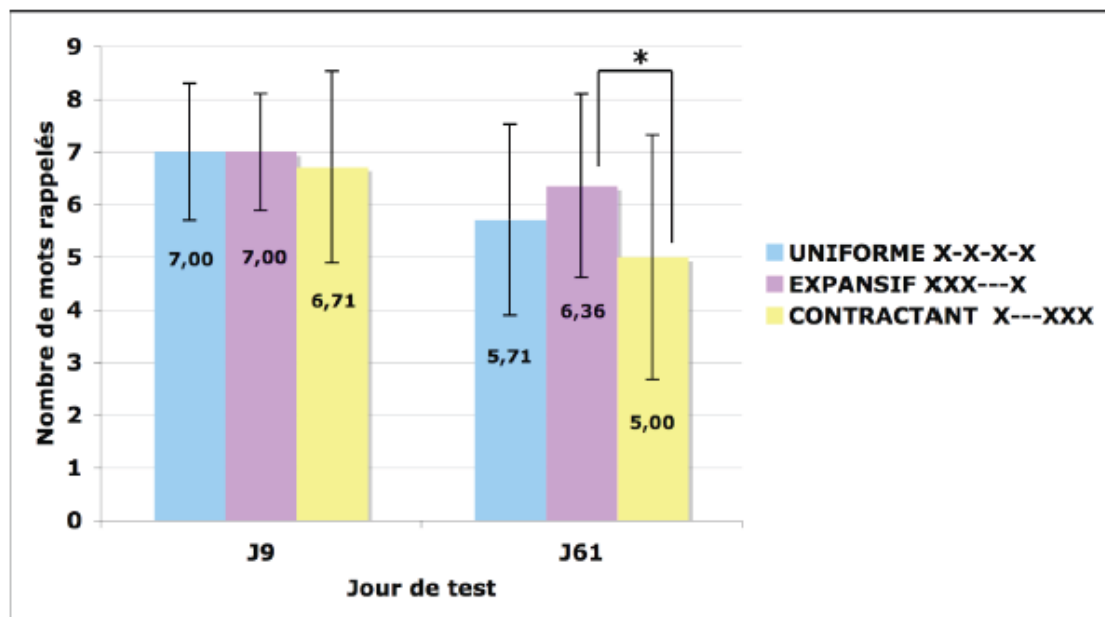


Figure 18 : Nombre moyen de mots rappelés par agencement (Uniforme, Expansif, Contractant) dans la tâche de rappel indicé à J9 et à J61 pour les participants présents à J61. Le nombre maximum de bonnes réponses était de 8. Les barres d'erreurs représentent les écart-types et les astérisques représentent les contrastes significatifs.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujets l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant). Cette analyse a révélé un effet significatif de l'agencement temporel : le nombre de bonnes réponses était significativement différent en fonction de l'agencement temporel des répétitions : $F(2,26) = 7.48, p < .01, \eta^2_p = 0.37$.

Des analyses de contraste ont révélé que les participants produisaient plus de bonnes réponses dans la condition expansive (6,36) que dans la condition contractante (5,0), $F(1,26) = 14.95, p < .001$. La différence de performance entre les conditions expansive et uniforme (5,71) était tendancielle, $F(1,26) = 3.35, p = .078$, ainsi que la différence de performance entre les conditions uniforme et contractante, $F(1,26) = 3.35, p = .052$.

Des analyses secondaires ont été réalisées dans le but de vérifier l'absence des effets de liste et de groupe. L'effet de la liste n'était pas significatif, $F(2,26) < 1$, et l'effet du groupe ne l'était pas non plus, $F(2,11) < 1$.

Afin d'avoir une idée de l'évolution des souvenirs à long terme, nous avons analysé comparativement les résultats des 14 participants à la session de J61, avec les performances de ces mêmes sujets à J9, au rappel indicé. Les scores moyens de ces 14 sujets lors de J9 sont présentés dans la section gauche de la Figure 18. Une analyse de

variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujets l'Agencement. De façon surprenante, ce sous-ensemble de l'échantillon ne présentait pas d'effet significatif de l'agencement temporel à J9, $F(2,26) < 1$. Les moyennes des scores étant élevées (respectivement 7,0 ; 7,0 ; et 6,71 pour les agencements uniforme, expansif et contractant), il se peut qu'un effet plafond soit observé dans ce sous-groupe pour le rappel indicé à J9. Cependant l'effet classique de l'agencement a été retrouvé chez eux à J61, comme si celui-ci s'était démasqué avec le temps.

A titre indicatif, voici les chiffres représentant la diminution des scores entre J9 et J61 en fonction des agencements. Pour l'agencement uniforme, les scores ont diminué en moyenne de 1,29 points, ce qui représente 18,4 % de diminution. Pour l'agencement expansif, les scores ont diminué en moyenne de 0,64 points, ce qui représente 9,2 % de diminution. Pour l'agencement contractant, les scores ont diminué en moyenne de 1,71 points, ce qui représente 25,5 % de diminution. L'analyse de variance réalisée sur la différence de score entre le rappel indicé à J9 et à J61, avec comme facteur intra-sujets l'agencement a révélé un effet marginal de l'agencement, $F(2,26) = 2.61$, $p < .10$. Ainsi, ces résultats suggèrent que l'oubli est le moins important pour les items ayant été appris dans la condition expansive, moyenne pour ceux de la condition uniforme et le plus important pour ceux de la condition contractante.

4. 3. 4. Discussion des Expériences 1 et 2

Rappel du principe des Expériences 1 et 2

Dans ces deux expériences, trois agencements temporels de quatre présentations d'items sur une période d'apprentissage de 7 jours étaient comparés : expansif (XXX---X), uniforme (X-X-X-X), et contractant (X---XXX). Les tâches de rappel indicé (ainsi que de rappel libre et de jugement de fréquence dans l'Expérience 2) se déroulaient le 9ème jour. Dans l'Expérience 1, les participants réalisaient leur phase d'apprentissage à domicile en se connectant chaque jour à Internet, et la tâche de rappel indicé avait lieu en laboratoire. L'Expérience 2 consistait en une réplique de l'Expérience 1 intégralement en laboratoire.

Résumé des résultats

L'Expérience 1 a révélé que l'agencement expansif était plus favorable au rappel indicé à J9 que les agencements uniforme et contractant. L'Expérience 2 a révélé que les agencements expansif et uniforme étaient plus favorables au rappel indicé à J9 que l'agencement contractant, mais sans différence significative entre eux. Dans la tâche de rappel libre, de nouveau l'agencement expansif et uniforme étaient tous deux supérieurs à l'agencement contractant. Enfin, les items présentés selon l'agencement expansif étaient jugés plus fréquents que les items de l'agencement uniforme, eux-mêmes jugés plus fréquents que ceux de l'agencement contractant. Les scores de reconnaissance lors de la seconde occurrence (P2) étaient plus élevés pour l'agencement expansif que pour les agencements uniforme et contractant. Par ailleurs, les items reconnus à P2, tous agencements confondus, avaient plus de chance d'être rappelés dans la tâche de rappel libre à J9 que les items non reconnus. Ainsi, l'hypothèse prédite par la théorie de la récupération en phase d'étude était confirmée (à la limitation près de l'égalité des agencements expansif et uniforme dans l'Expérience 2).

Bien que les scores moyens dans la tâche de rappel indicé se sont révélés très similaires dans les Expériences 1 et 2, parfois au centième près (par exemple, 6.14 et 6.13 sur 8 pour l'agencement expansif), les deux expériences ont mis en évidence un effet

différent de l'agencement temporel des répétitions. D'une façon générale, l'agencement expansif des répétitions était associé à une meilleure performance finale, en valeur absolue, que l'agencement contractant. L'agencement uniforme avait une efficacité soit similaire à celle de l'agencement expansif (Expérience 2, et comme dans Gay, 1973 et Cull, 2000), soit similaire à celle de l'agencement contractant (Expérience 1). La section suivante discute des raisons potentielles de ces différences de résultats.

Causes potentielles de la différence de résultats entre les Expériences 1 et 2

Cette différence inattendue de résultats entre les deux expériences peut être due à de nombreux facteurs. L'interprétation la plus simple de cette différence est la suivante : la condition d'agencement uniforme est passée d'un niveau 'faible' (5.50 sur 8) à un niveau 'élevé' (6.08 sur 8) lors du passage de l'Expérience 1 à l'Expérience 2, les autres agencements ne présentant pas d'évolution.

Premièrement, il est possible que certains participants de l'Expérience 1 n'aient pas réalisé leurs sessions d'apprentissage dans les meilleures conditions. Il est possible par exemple que pour des raisons de manque de temps ou par baisse de motivation, certaines sessions aient été réalisées plus rapidement ou avec moins de sérieux. Cependant, il n'y a pas de raison évidente pour laquelle l'agencement uniforme en ait été plus affecté que les autres agencements.

Deuxièmement, une meilleure procédure de contrôle des effets de primauté et de récence a été mise en œuvre dans l'Expérience 2 par rapport à l'Expérience 1. Cependant, l'observation des positions sérielles respectives des items (voir l'Annexe n°5) dans les sessions d'apprentissage, en fonction des différents agencements, ne révèle pas de défaut flagrant qui nous aurait échappé pendant la constitution des listes dans l'Expérience 1. De nouveau, il n'y a pas de raison évidente pour laquelle l'agencement uniforme ait été plus affecté que les autres agencements par ces éventuelles différences.

Troisièmement, certaines différences de procédure existant entre les deux expériences pourraient expliquer les différences de résultats. Tout d'abord, les temps de présentation des paires en phase d'apprentissage étaient soumis à des contraintes différentes dans les deux cas. Dans l'Expérience 1, ils dépendaient des participants puisque ceux-ci décidaient eux-mêmes de passer de paire en paire, tandis que, dans l'Expérience 2, ils étaient fixés à 30 secondes. On ne peut pas exclure que les participants aient géré le temps imparti de manière différente dans les deux expériences. On peut imaginer par exemple que dans l'Expérience 1, le participant déclenchait le passage à la paire suivante si et seulement s'il avait l'impression qu'il était arrivé au maximum de son apprentissage. A contrario, on peut imaginer que, dans l'Expérience 2, les 30 secondes allouées à chaque paire étaient investies au maximum par le participant, menant à deux situations : 1) son apprentissage pouvait être perçu comme non terminé au bout des 30 secondes, et il était forcé de passer à la paire suivante, ou 2) l'apprentissage de la paire pouvait être perçu comme terminé avant les 30 secondes et il profitait du temps restant pour 'réviser' les paires apprises précédemment dans la session. Cependant, et de nouveau, il n'y a pas de raison évidente pour laquelle l'agencement uniforme ait été plus affecté par ces stratégies que les autres agencements.

D'autre part, les modalités de réponse lors de la tâche de rappel indicé étaient elles aussi soumises à des contraintes différentes. Les réponses étaient données oralement dans l'Expérience 1, alors qu'elles étaient écrites dans l'Expérience 2. Cette différence seule ne nous semble pas pouvoir être responsable de la différence de résultats observés dans les deux expériences. De la même façon, dans l'Expérience 1, aucune limite de temps n'était

imposée, mais les participants devaient répondre le plus vite possible. Dans l'Expérience 2, l'item n'était présenté que pendant 10 secondes et un signal sonore forçait le sujet à tourner la page de son carnet sans retour en arrière possible. De nouveau, cette explication ne nous semble pas pertinente pour expliquer pourquoi, en condition de temps de réponse limitée (Expérience 2), les items de l'agencement uniforme et eux seuls étaient mieux rappelés qu'en condition de temps illimité (Expérience 1). Cette explication pourrait être logique dans le cas contraire, mais pas dans celui-ci.

Une piste plus probable est liée à l'introduction de la tâche de reconnaissance pendant la phase d'apprentissage dans l'Expérience 2. Suite aux 30 secondes de présentation de chacune des paires, les participants devaient juger si la paire présentée avait déjà été rencontrée précédemment. Cette consigne explicite incitant les participants à se demander si les occurrences d'items étaient des répétitions n'était pas donnée dans l'Expérience 1. Ainsi, dans l'Expérience 2, le fait d'insister explicitement sur la reconnaissance des paires a peut-être favorisé les processus de récupération en phase d'étude, par rapport à l'Expérience 1. Peut-être que les participants, pour répondre à cette tâche, ont davantage « sondé » leur mémoire pour tenter de se souvenir des moyens mnémotechniques éventuellement mis en œuvre lors d'une occurrence précédente. Or, c'est justement pour l'agencement uniforme que cette différence aurait pu jouer de la façon la plus prégnante. En effet, on peut supposer que au bout de 24h, c'est-à-dire lors de P2 dans l'agencement expansif, la reconnaissance des items répétés était facile et évidente, et ce dans les deux expériences. À l'inverse, on peut supposer que au bout de 72h, c'est-à-dire lors de P2 dans l'agencement contractant, la reconnaissance des items répétés était difficile et moins probable, et ce, dans les deux expériences. A mi-chemin entre ces deux situations, la reconnaissance des items répétés au bout de 48h, c'est-à-dire lors de P2 dans l'agencement uniforme, était probablement de difficulté intermédiaire ; c'est peut-être spécifiquement sur ces items-là que le fait d'inciter les participants à se souvenir d'anciennes occurrences a pu jouer un rôle favorisant la récupération en phase d'étude, et par conséquent, des performances finales plus élevées. Ainsi, les résultats de l'Expérience 2 étaient possiblement « contaminés » par la tâche de reconnaissance, que nous souhaitons pourtant ne pas constituer d'interférence avec l'apprentissage.

Discussion des résultats pour les autres tâches de mémoire (Expérience 2)

Dans cette partie, nous considérerons les résultats obtenus pour chacune des autres tâches de mémoire dans l'Expérience 2.

Rappel libre.

Concernant la tâche de rappel libre, on peut évidemment suggérer que sa réalisation par les sujets étaient fortement contaminée par la tâche de rappel indicé juste antérieure, et dans laquelle tous les pseudomots cibles étaient présentés. Le fait d'observer le même profil de performance en rappel libre et en rappel indicé, avec cependant des scores plus faibles en rappel indicé, va dans le sens de cette hypothèse de contamination. On ne peut donc pas tirer davantage d'information à partir du résultat de cette tâche et nous ne l'évoquerons plus dans la suite de ce travail.

Jugement de fréquence.

Concernant la tâche de jugement de fréquence, le profil de réponses observé variait par rapport à la tâche de rappel indicé puisque les fréquences estimées étaient décroissantes de l'agencement expansif à l'agencement uniforme, et de l'agencement uniforme à

contractant. Les fréquences estimées, par ailleurs, étaient proches de la valeur réelle 4, tout en la dépassant légèrement. Ce résultat indique que l'agencement temporel influencerait la fréquence d'occurrence subjective des items. On peut interpréter l'avantage de l'agencement expansif selon l'hypothèse de la remémoration récursive (voir la section 2.2). En effet, Hintzman (2004) invoque la notion de remémoration récursive pour expliquer les résultats obtenus dans différentes études concernant les jugements de fréquence. Selon cette hypothèse, proche de celle de la récupération en phase d'étude, le sujet encoderait, lors d'une répétition d'un item, à la fois les processus réalisés pendant cette occurrence, mais également le résultat de la recollection que cette occurrence déclenche. Autrement dit, la trace mnésique correspondant à l'épisode P3 contient le traitement réalisé sur P3 + le souvenir de P2 qui est réactivé + le souvenir de P1 qui était contenu dans le souvenir de P2. De répétition en répétition, ces enchâssements de recollections vont permettre au sujet de pouvoir estimer le nombre d'occurrences qui se sont produites pendant l'apprentissage. Par conséquent, si l'on ne reconnaît pas un item lors d'une occurrence, disons par exemple lors de P2 dans les agencements uniforme ou contractant, la recollection de P1 au moment de P2 ne sera pas contenue dans la trace finale. En définitive, l'estimation du nombre d'occurrences sera logiquement moins élevée que dans une situation où la recollection se produit lors de chaque répétition de l'item.

Nous sommes allée explorer nos données pour vérifier si, selon cette hypothèse, les paires n'ayant pas été reconnues lors de P2 ont été estimées comme moins fréquentes que les paires ayant été reconnues lors de P2. Nous avons calculé pour chaque sujet la moyenne de la fréquence estimée, d'une part pour les paires que le sujet avait reconnues à P2, et d'autre part pour les paires que le sujet n'avait pas reconnues à P2. Nous avons dû éliminer de l'analyse les sujets qui avaient reconnu l'intégralité des paires lors de P2 (effectif restant = 26). Les fréquences estimées moyennes étaient de 4.23 pour les paires reconnues à P2, contre une moyenne de 3.80 pour les paires non reconnues à P2 (tous sujets confondus et tous agencements confondus). Un test de Wilcoxon sur échantillons appariés a révélé que cette différence était significative (nombre de différences non nulles = 25, $T = 37.2$, $z = 2.81$, $p < .01$). Une limite à ces calculs était le fait qu'il y avait une majorité de paires reconnues à P2 et donc un petit nombre de paires non reconnues. Cependant, et de la même façon que pour le rappel libre, on ne peut pas exclure une contamination de la tâche de jugement de fréquence par les performances aux tâches précédentes : pour faire leurs jugements, les sujets pouvaient se baser sur la recollection éventuelle des épisodes d'occurrences, mais également sur leur sentiment de savoir selon une logique proche de la suivante : « plus j'ai l'impression de savoir, plus cela signifie que j'ai dû voir l'item souvent ». Une variable médiatrice pour le jugement de fréquence pourrait être la fluence de traitement (Jacoby & Dallas, 1981 ; Whittlesea, Jacoby, & Girard, 1990 ; Whittlesea & Leboe, 2003). Toutefois, l'observation d'un profil de réponse différent par rapport à la tâche de rappel indicé, en particulier l'observation d'une différence significative entre l'agencement expansif et uniforme, suggère que cette contamination ne semble pas pouvoir expliquer le profil observé. En tout état de cause, il serait intéressant de répliquer ce jugement de fréquence, sans contamination, dans les futures expérimentations.

Reconnaissance des paires pendant l'apprentissage.

Concernant la tâche de reconnaissance des items lors de P2, les items répétés au bout de 24h (agencement expansif) étaient plus souvent reconnus que les items répétés au bout de 48 et 72h (respectivement dans les agencements uniforme et contractant), sans différence significative entre ces deux derniers agencements. Ce résultat reflète naturellement l'oubli qui se produit progressivement suite à la présentation initiale d'une information. Cependant,

l'absence de différence de taux de reconnaissance entre les IIR de 48 h et de 96 h est relativement surprenante, bien que cela reflète le fait que la courbe d'oubli s'aplatit avec le temps (Wixted, 2004) ; en effet, la pente de la courbe d'oubli est la plus forte au début du DR puis est de plus en plus faible avec le passage du temps. On remarquera que les taux de reconnaissance étaient relativement élevés même après 96 h d'intervalle, avec un score de 6.87 sur 8, d'autant plus qu'un grand nombre de distracteurs étaient mêlés aux paires cibles. Il faut toutefois admettre que ce sur quoi les participants basaient leur prise de décision de reconnaissance n'est pas évident. En effet, la reconnaissance d'une paire donnée pouvait être basée sur la paire entière, sur le pseudomot de cette paire, mais également sur le mot seul, étant donné qu'il était précisé aux participants qu'un mot donné était toujours associé au même pseudomot.

L'analyse des taux de reconnaissance à P2 doit être complétée par l'analyse du devenir de ces paires au moment du rappel indicé. La théorie de la récupération en phase d'étude (Johnston & Uhl, 1976 ; Hintzman, 2004) énonce qu'un item dont la trace mnésique est récupérée lors de la répétition bénéficierait de façon positive de cette récupération. À l'inverse, le fait de ne pas récupérer les traces mnésiques associées à un item oblige le sujet à recommencer le traitement de l'item à partir de zéro sans pouvoir bénéficier d'une trace existante en mémoire. Ainsi, la théorie de la récupération en phase d'étude prédit que les items reconnus pendant leur répétition devraient voir leur encodage favorisé par rapport aux items non reconnus, et donc également leur rappel final. C'est justement ce qui a été observé puisque, tous agencements confondus, 76 % des items reconnus à P2 ont été rappelés en rappel indicé à J9, contre seulement 52 % des items non reconnus.

L'étude de ces pourcentages, agencement par agencement, révèle les chiffres suivants. Pour l'agencement uniforme, 78 % des items reconnus à P2 ont été finalement rappelés contre 62 % des items non reconnus. Pour l'agencement expansif, 78 % des items reconnus à P2 ont été finalement rappelés contre 50 % des items non reconnus (dans ce cas ce pourcentage est à prendre avec précaution dans la mesure où il ne reflète que 10 essais au total, tous sujets confondus). Pour l'agencement contractant, 72 % des items reconnus à P2 ont été finalement rappelés contre 43 % des items non reconnus. Ainsi, pour l'agencement contractant, non seulement le taux de reconnaissance lors de P2 était faible dans l'absolu, mais même les items qui avaient été reconnus à P2 étaient moins nombreux que les autres à bénéficier d'un succès au rappel indicé final (les analyses statistiques n'ont pas été réalisées). Les agencements uniforme et expansif présentaient le même pronostic de réussite au rappel indicé pour leurs items reconnus à P2, bien que le taux en valeur absolue de reconnaissance soit inférieur en condition uniforme. En revanche, le devenir des items non reconnus à P2 est plus favorable dans l'agencement uniforme que dans l'agencement expansif. Autrement dit, parmi les items non reconnus à P2, ce sont les items de l'agencement uniforme qui bénéficiaient le plus de la suite de l'agencement, c'est-à-dire des essais P3 et P4.

Ces résultats vont donc dans le sens de la théorie de la récupération en phase d'étude, dans la mesure où les items qui étaient reconnus lors de leur seconde occurrence avaient plus de probabilité d'être rappelés lors de la tâche finale. Ces résultats répliquent donc les observations de Bellezza et al. (1975), Johnston et Uhl (1976) et Melton (1967), dont les répétitions avaient lieu à court terme. Nous n'avons cependant pas mis en évidence de lien causal entre récupération en phase d'étude et rappel indicé final, dans la mesure où nous n'avons fait qu'observer des phénomènes, tandis qu'un lien de causalité ne peut être établi que par la manipulation explicite de la récupération -ou de la non récupération. De plus, si le lien causal était direct, nous aurions observé des profils d'effet d'agencement

similaires dans la tâche de reconnaissance à P2 et dans la tâche de rappel libre final. Au contraire, il y avait une différence significative de reconnaissance lors de P2 entre les agencements expansif d'une part et uniforme et contractant d'autre part, tandis qu'il y avait une différence significative de rappel indicé entre les agencements expansif et uniforme d'une part et contractant d'autre part.

Il est important de mentionner également que ces données peuvent donner lieu à une autre interprétation, selon laquelle les items qui étaient reconnus à P2 avaient plus de chances d'être également rappelés ultérieurement pour la simple raison qu'ils étaient plus faciles à apprendre pour les participants. Cette interprétation est basée sur l'idée d'une corrélation entre le phénomène de reconnaissance et le phénomène de rappel indicé ultérieur : les items qui n'ont pas été reconnus à P2 étaient -et sont restés lors de la suite de la phase d'apprentissage- plus difficiles pour le sujet. A contrario, notre interprétation liée à la théorie de la récupération en phase d'étude est basée sur l'idée d'un lien de causalité entre le phénomène de reconnaissance et le phénomène de rappel indicé ultérieur : le fait de ne pas avoir reconnu l'item à P2 a empêché la récupération en phase d'étude et a privé la trace d'un épisode d'apprentissage (P1). Il est impossible de démêler ces deux interprétations sur la base des données présentées ici, puisqu'il faudrait pouvoir influencer expérimentalement la reconnaissance ou la non reconnaissance des items.

Ces constations nous amènent à des réflexions sur la nature de la récupération dans la théorie de la récupération en phase d'étude. En effet, nous avons postulé que le phénomène de récupération en phase d'étude est reflété expérimentalement par la reconnaissance d'un item comme étant répété. Autrement dit, si un item est reconnu au moment de sa répétition, alors les traces mnésiques associées à cet item sont réactivées en mémoire. On peut cependant s'interroger sur ce postulat. Premièrement, la réactivation et la récupération d'une trace mnésique s'accompagnent-elles obligatoirement d'un sentiment de reconnaissance ? Ne serait-il pas envisageable que la répétition induise, parfois, la réactivation d'éléments mnésiques en deçà du seuil de conscience, ces traces étant alors tout de même renforcées par le nouvel encodage ? D'autre part, on peut imaginer que la récupération des traces mnésiques, plutôt que d'exister selon un mode de type 'tout ou rien', pourrait avoir une intensité variable. Autrement dit, le sujet récupérerait plus ou moins d'éléments de la trace, et/ou récupérerait des traces plus ou moins riches. Ces questions sont à mettre en lien avec les discussions actuelles mettant en opposition les notions de familiarité et de recollection (e.g., Gardiner & Richardson-Klavehn, 2000 ; Yonelinas, 2002). Les données présentées dans ce travail ne permettent pas de répondre à ces questions. Nous pensons que la théorie de la récupération en phase d'étude bénéficierait d'une clarification de ces questions pour devenir une théorie véritablement testable. Nous discuterons de nouveau de ce point dans la discussion générale.

Enfin, on peut discuter des relations entre les scores à la tâche de récupération en phase d'étude et à la tâche de jugement de fréquence. Le jugement de fréquence peut être basé sur des processus du type comptage du nombre d'occurrences pour un item. Lorsque le souvenir d'un item est récupéré avec succès lors d'une nouvelle occurrence de cet item, le « compteur » interne du sujet relatif à cet item s'incrémente d'un point. Lorsqu'une nouvelle occurrence d'un item a lieu mais que le sujet ne reconnaît pas l'item, le compteur interne ne s'incrémente pas. Ainsi, un item qui est reconnu à chacune de ses occurrences est associé à une occurrence estimée plus élevée qu'un item qui n'est pas reconnu à chaque occurrence.

De façon plus subtile, on peut imaginer des processus de comptage indirects dans lesquels, à chaque nouvelle occurrence, le sujet -avant d'incrémenter son compteur interne- ferait une estimation de la fréquence des occurrences précédentes (plutôt que de récupérer

le comptage issu des occurrences précédentes). Cette estimation de fréquence pendant la phase d'étude pourrait être basée sur la richesse des traces récupérées. Ainsi, un item reconnu et dont la richesse des souvenirs récupérés lors de la nouvelle occurrence est importante va mener à une estimation de fréquence plus élevée qu'un item qui induit une récupération de trace moins riche. Une prédiction de cette hypothèse serait que la fréquence estimée devrait être plus faible pour les items qui n'ont pas été reconnus à P2, ce qui était le cas. Ces réflexions sont en lien avec la théorie de la remémoration récursive (Hintzman, 2004). Nous ne développerons pas davantage ces questions car elles vont au delà du cadre du présent travail.

Motivation

Les estimations de motivation ont été demandées aux participants afin de contrôler que des variations de motivation au fil du temps n'étaient pas responsables des différences de performances finales entre les agencements. En effet, puisque notre hypothèse était que les performances seraient les plus élevées pour l'agencement expansif suivi de l'agencement uniforme, suivi de l'agencement contractant, cela aurait pu être dû au fait que la motivation des participants diminuait au fil des sessions. En effet, le niveau de motivation est clairement lié à l'efficacité de l'apprentissage (e.g., de Bonis, 1968). Ainsi, si le niveau de motivation était le plus faible, et donc l'efficacité d'apprentissage également, lors des dernières sessions d'apprentissage, alors l'agencement contractant aurait été celui qui en aurait le plus pâti car c'est en fin de phase d'apprentissage que ses items étaient présentés. Les observations ont montré que, de session en session, le niveau de motivation estimé en début de session ne diminuait pas significativement. En revanche, le niveau de motivation en fin de session diminuait au fil des sessions. Nous interprétons cette diminution de motivation comme étant due à la difficulté de l'expérience pour les participants. En effet, un certain nombre d'entre eux se plaignaient que l'expérience était « difficile » et qu'« ils ne se souviendraient de rien pendant le test ». Cependant, en début de session suivante, la motivation était revenue et, nous le pensons, l'efficacité d'apprentissage également. En tout état de cause, une absence de tendance significative ne signifie pas que l'effet n'existe pas. Ainsi, l'hypothèse d'une diminution de la motivation ne peut pas être complètement exclue.

Rétention à J61

Comme notre hypothèse le prédisait, mais en gardant à l'esprit les réserves liées au faible effectif de l'échantillon testé, les items de l'agencement expansif produisaient de nouveau les meilleures performances deux mois après l'apprentissage. Ces résultats suggèrent donc que la supériorité de l'agencement expansif est un phénomène durable. L'Expérience 3 manipulait justement la durée du délai de rétention afin de tester plus précisément l'évolution de l'effet de l'agencement dans le temps.

4. 4. Expérience 3

4. 4. 1. Introduction

L'objectif de la troisième expérience était double. Il s'agissait en premier lieu de poursuivre l'exploration de la généralité de l'effet de l'agencement temporel observé dans les expériences précédentes, en variant les paramètres des agencements. La durée de la phase d'apprentissage a été allongée à 13 jours et le nombre d'occurrences d'une paire

donnée a été réduit à trois. Les agencements étaient les suivants : uniforme (X-----X-----X), expansif (XX-----X) et contractant (X-----XX).

Il s'agissait en second lieu d'étudier l'influence du délai de rétention (DR) sur les effets de l'agencement temporel. En effet, les études récentes sur l'effet de l'intervalle inter-répétition (IIR) ayant employé deux occurrences d'un même item (voir la section 1.3.2) ont révélé que la durée du DR modifiait la durée de l'IIR optimal. Précisément, plus le DR était long, plus l'IIR optimal était long également. Ainsi, il est possible que, suite à un DR long, l'effet des agencements temporels sur le rappel final ne soit pas le même que suite à un DR court. Nous avons donc comparé trois modalités pour la durée du DR : 2 jours (le test de rétention ayant alors lieu à J15), 6 jours (test à J19), et 13 jours (test à J26). De nouveau, chaque participant était confronté aux trois agencements temporels, mais était cependant testé suite à un seul DR.

Un tel protocole nous permettait de comparer indirectement les théories de la récupération en phase d'étude et de la variabilité de l'encodage. En effet, ces deux théories font des prédictions différentes relativement à l'interaction entre l'agencement et le DR. Selon la théorie de la récupération en phase d'étude, l'agencement expansif devrait aboutir, comme dans les Expériences 1 et 2, à des performances finales plus élevées que pour les autres agencements, et ce, quelle que soit la durée du DR.

Selon la théorie de la variabilité de l'encodage, une interaction entre l'agencement et le DR devrait apparaître. Rappelons les principaux postulats de cette théorie. D'une part, le contexte fluctuerait graduellement avec le temps, c'est-à-dire pendant le (ou les) IIR et également pendant le DR. D'autre part, les éléments contextuels seraient encodés dans les traces aux côtés des représentations des items. Enfin, les traces récupérées lors du rappel seraient celles qui comportent des éléments similaires au contexte instancié à ce moment-là. Ainsi, lorsque le DR est court, le contexte instancié lors du rappel est censé être proche des contextes qui avaient été instanciés à la fin de la phase d'apprentissage. Par conséquent, un DR court devrait favoriser la récupération des traces encodées majoritairement en fin d'apprentissage, c'est-à-dire celles des items de l'agencement contractant. A contrario, lorsque le DR est long, le contexte instancié lors du rappel est censé être très différent des contextes instanciés au cours de l'apprentissage, à cause de l'importante fluctuation contextuelle due au passage du temps. Par conséquent, un DR long devrait favoriser la récupération des traces contenant la plus grande variété de composants différents car elles bénéficient alors d'un plus grand nombre de routes de récupération. En toute logique, les paires apprises selon l'agencement uniforme devraient être celles comportant la plus grande variété de composants contextuels, dans la mesure où elles ont été encodées dans trois contextes bien différents, car pendant des sessions séparées par de longs intervalles : à J1, à J7, et à J13. À l'inverse, les paires apprises selon les agencements expansif et contractant comportent deux occurrences se produisant dans des contextes proches : J1 et J2 pour l'agencement expansif, et J12 et J13 pour l'agencement contractant. Ces deux agencements devraient donc donner lieu à des traces mnésiques moins diversifiées en termes de composants contextuels. En définitive, l'agencement contractant devrait être favorisé pour un DR court, tandis que l'agencement uniforme devrait être favorisé pour un DR long.

4. 4. 2. Méthode

Participants

Soixante-douze étudiant(e)s ont participé à cette expérience, dont 60 femmes et 12 hommes, âgé(e)s entre 18 et 26 ans (âge moyen 20 ans) et de langue maternelle française. Leur vue était jugée satisfaisante à excellente (à l'état normal ou corrigé). Chaque participant ayant terminé l'étude a bénéficié d'un dédommagement de 20 euros ou de 35 euros selon la période d'expérimentation.

Les participants étaient recrutés par le biais de petites annonces ou sur Internet. Aucun participant n'avait pris part aux prétests ni à une expérience précédente. Ils n'étaient pas étudiants en Psychologie cognitive ou Sciences cognitives au delà de la 2ème année de licence. Un échange de mails a été nécessaire avec de nombreux étudiants pour l'explication du principe de l'expérience et convenir des dates et horaires de passation.

L'expérience a été menée en deux temps. La première période d'expérimentation s'est déroulée en mars 2008 et correspondait aux groupes avec dont le DR était de 2 ou 5 jours (respectivement J15 et J19). Les participants étaient alors dédommagés de 35 euros. La seconde période d'expérimentation s'est déroulée en novembre 2009 et correspondait au groupe dont le DR était de 13 jours (J26). Les participants étaient alors dédommagés de 20 euros. Initialement, deux DR devaient être testés lors de la seconde période d'expérimentation, c'est-à-dire J14 et J26, mais les faibles effectifs de participants nous ont amenée à modifier le planning une fois l'expérience commencée de telle sorte à tester uniquement à J26. Une participante prévue pour J14 et n'étant pas disponible à J26 a passé le test à J15, complétant ainsi l'effectif de ce groupe appartenant à la première période expérimentale. Les résultats sont considérés comme issus d'une seule expérience.

Concernant la première période d'expérimentation, lors de la première session, 55 participants étaient attendus, cinq d'entre eux ne se sont pas présentés, et sur les 50 participants qui ont participé à la première session, un participant n'a pas pu suivre toutes les sessions et a donc été supprimé de l'échantillon. Concernant la seconde période d'expérimentation, lors de la première session, 37 participants étaient attendus, deux d'entre eux ne se sont pas présentés, et sur les 35 participants qui ont assisté à la première session, onze n'ont pas pu suivre toutes les sessions et ont donc été supprimés de l'échantillon. De plus, parmi les 24 participants ayant fini l'expérience, les résultats de deux d'entre eux n'ont pas été intégrés à l'analyse du fait d'un comportement manifestement peu sérieux pendant les sessions d'apprentissage et ne garantissant pas un apprentissage correct.

Les participants ont été répartis dans les groupes de contrebalancement A, B, C en fonction de leur disponibilité horaire. Chaque jour de passation expérimentale, au moins 6 sessions étaient organisées pour que chaque participant puisse se rendre disponible à une des sessions correspondant à son groupe expérimental.

Les effectifs finaux de chaque groupe de participants, en fonction du groupe de contrebalancement, sont présentés dans le Tableau 5.

Tableau 5 : Effectifs des groupes selon le jour de test et le groupe de contrebalancement.

Jour de Test	Groupe			Total
	A	B	C	
J15	10	8	9	27
J19	9	8	6	23
J26	7	6	9	22

Appareils

Les stimuli étaient présentés par un ordinateur Apple Macintosh de type MacBook doté de 1,7 Go de RAM au moyen du logiciel Psyscope (Cohen et al., 1993). Les informations étaient projetés sur un écran à l'aide d'un vidéoprojecteur.

Matériel

Les paires mot-pseudomot que dans les Expériences 1 et 2 ont été utilisées de nouveau (24 paires cibles), ajoutées de six paires cibles supplémentaires créées pour l'occasion, de la catégorie sémantique des instruments de musique. Il y avait donc au total 30 paires cibles et 70 paires distractrices (et non plus 72 comme dans les Expériences 1 et 2). Ces 30 paires cibles étaient réparties en 3 listes de 10 paires appelées L1, L2, et L3 (voir la section 4.1.3).

Des formulaires d'évaluation de motivation ont été fournis à chaque participant lors de chaque session d'apprentissage. Un formulaire contenait cinq questions, pour lesquelles les participants devaient évaluer, sur une échelle de 1 à 7 :

- leur niveau de motivation ainsi que leur niveau de concentration / attention pour la session du jour, à remplir avant de commencer l'apprentissage.
- leur niveau de concentration / attention (a posteriori) ainsi que leur performance globale d'apprentissage, pour la session qu'ils venaient juste de finir.
- leur niveau de motivation pour la poursuite de l'expérience.

Les carnets pour les réponses au test de rappel indicé étaient similaires à ceux utilisés dans l'Expérience 2, à l'exception que le nombre de pages destinées aux réponses des participants était de 32.

Des formulaires de rappel libre et de jugement de fréquence similaires à ceux utilisées dans l'Expérience 2 étaient élaborés. Le formulaire de rappel libre était composé d'un tableau à deux colonnes avec des cases numérotées de 1 à 33. Le formulaire de jugement de fréquence était composé de 40 lignes numérotées associées à chaque fois à une échelle de chiffres allant de 0 à 7.

Design expérimental

La phase d'apprentissage a duré 13 jours et a comporté 5 sessions d'apprentissage, ayant lieu les jours 1, 2, 7, 12, et 13. Le Tableau 6 représente la répartition des sessions d'apprentissage et des jours de test. Les participants devaient donc se déplacer six fois.

Tableau 6 : Répartition des sessions d'apprentissage et des sessions de test au fil des jours. Les présentations d'items sont représentées par la lettre P et varient en fonction des agencements. Les sessions de test sont représentées par la lettre T.

	Jour																										
Agencement	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	
Uniforme	P						P						P		T				T								T
Expansif	P	P											P		T				T								T
Contractant	P											P	P		T				T								T

Chaque paire cible était présentée trois fois aux participants, selon l'un des trois agencements temporels suivants :

- dans l'agencement Uniforme (X-----X-----X), une paire donnée était présentée les sessions 1, 3, et 5, c'est-à-dire les jours 1, 7, et 13.

- dans l'agencement Expansif (XX-----X), une paire donnée était présentée les sessions 1, 2, et 5, c'est-à-dire les jours 1, 2, et 13.
- dans l'agencement Contractant (X-----XX), une paire donnée était présentée les sessions 1, 4, et 5, c'est-à-dire les jours 1, 12, et 13.

La variable agencement était, de nouveau, une variable intra-sujets. Le principe de contrebalancement des listes selon les agencements en carré latin était le même que celui utilisé dans les Expériences 1 et 2. Le Tableau 7 présente ce contrebalancement. De nouveau, les paires distractrices avaient pour but de combler les lacunes laissées par l'absence de paires cibles dans les sessions d'apprentissage. Les paires distractrices étaient des paires non répétées et non testées, de nature similaire aux paires cibles. Elles sont symbolisées par la lettre *d* dans le tableau.

Tableau 7 : Répartition des présentations des listes (L1, L2, et L3) au cours des sessions d'apprentissage (J1, J2, J7, J12, et J13) en fonction des groupes de contrebalancement (A, B, ou C) et de l'agencement temporel (Uniforme, Expansif, ou Contractant). La lettre *d* symbolise la présentation de paires distractrices.

Groupe	Agencement	Jour												
		1	2	3	4	5	6	7	8	9	10	11	12	13
A	Uniforme	L1	<i>d</i>					L1					<i>d</i>	L1
	Expansif	L2	L2					<i>d</i>					<i>d</i>	L2
	Contractant	L3	<i>d</i>					<i>d</i>					L3	L3
B	Uniforme	L2	<i>d</i>					L2					<i>d</i>	L2
	Expansif	L3	L3					<i>d</i>					<i>d</i>	L3
	Contractant	L1	<i>d</i>					<i>d</i>					L1	L1
C	Uniforme	L3	<i>d</i>					L3					<i>d</i>	L3
	Expansif	L1	L1					<i>d</i>					<i>d</i>	L1
	Contractant	L2	<i>d</i>					<i>d</i>					L2	L2

Concernant le placement des paires au sein des sessions d'apprentissage, les mêmes règles que dans l'Expérience 2 ont été respectées (succession des catégories sémantiques, succession des agencements, zones tampon, positionnement des paires distractrices). De plus, les deux premières paires présentées dans chaque session étaient des paires distractrices, afin de contrôler un éventuel effet de primauté dans la session.

La séquence des paires présentées au groupe A est représentée dans l'Annexe n°6.

Procédure

Principe général.

Chaque participant s'est rendu dans une salle, toujours la même, aux jours indiqués par l'expérimentatrice pour chacune des cinq sessions, pour une durée de 20 minutes environ à chaque fois. Le 15ème, le 19ème, ou le 26ème jour selon les groupes, chaque participant s'est rendu dans une autre salle pour participer au test portant sur les paires apprises pendant la phase d'apprentissage, et ce pour une durée de 45 minutes environ. Les salles utilisées dans la première (Octobre 2008) et la seconde période d'expérimentation (Novembre 2009) étaient différentes.

Les sessions d'apprentissage et le test ont été réalisées par groupes de 1 à 16 personnes. Les paires étaient projetées sur un tableau blanc par vidéoprojection.

Chaque jour où une session avait lieu, deux créneaux horaires, toujours les mêmes, étaient possibles pour les participants d'un groupe donné (A, B ou C), afin de satisfaire à leurs contraintes de disponibilité. Cependant, il était recommandé aux participants de venir, dans la mesure du possible, toujours au même créneau. A plusieurs reprises, des sessions avec un seul participant ont été aménagées par suite d'oubli de l'horaire ou de problème de disponibilité.

Contrairement à l'Expérience 2, il n'y avait pas de tâche de reconnaissance à réaliser pendant la phase d'apprentissage, mais il y avait des auto-évaluations en début et en fin de chaque session.

Suite à la passation de l'expérience proprement dite, les participants étaient conviés à une séance d'explication de l'expérience 60 jours après le début de la phase d'apprentissage. Ils n'en avaient pas été avertis mais une tâche de rappel indicé avait lieu ce jour-là. Douze participants (9 ayant passé la session de test à J15, 3 ayant passé le test à J19, aucun ayant passé le test à J26) sont revenus. La procédure était identique à celle de la tâche de rappel indicé, à l'exception de l'ordre des items qui était modifié.

Consignes.

Les mêmes consignes que dans l'Expérience 1 et 2 ont été données aux participants. En particulier, les trois méthodes d'apprentissage, basées sur la recherche de sens dans le pseudomot, la fabrication d'images mentales, et la répétition mentale, ont été proposées. Ces conseils d'apprentissage étaient re-précisés au début des sessions 1 à 4.

Lors de chacune des sessions, 32 paires pseudomot-mot seraient présentées. Les paires défileraient de façon automatique toutes les 25 secondes. Il était annoncé que certaines paires seraient répétées au cours de la semaine contrairement à d'autres qui n'apparaîtraient qu'une seule fois.

Phase d'apprentissage.

En premier lieu, le formulaire de motivation était distribué aux participants. Ceux-ci devaient inscrire leur nom et répondre aux deux premières questions, c'est-à-dire évaluer leur niveau de motivation, puis de concentration/attention pour la session, sur une échelle de 1 à 7.

Quand les participants avaient répondu et qu'ils étaient prêts, la phase d'apprentissage commençait. Au centre de l'écran, projeté au tableau, apparaissaient les paires mot-pseudomot pendant 25 secondes chacune, avec le mot à gauche et le pseudomot à droite (police Arial Narrow, caractère gras, taille 55, hauteur des caractères apparents : 5 cm environ). Ensuite, une croix de fixation centrale apparaissait pendant 2 secondes puis la paire suivante apparaissait comme la première, et ainsi de suite.

Après l'apparition des 32 paires (phase qui avait duré 15 minutes), les participants répondaient immédiatement aux trois dernières questions du formulaire de motivation, c'est-à-dire évaluer leur niveau de concentration/attention a posteriori, leur performance d'apprentissage a posteriori, et leur motivation à poursuivre l'expérience, sur une échelle de 1 à 7. Les formulaires étaient ramassés et les participants pouvaient quitter la salle.

Phase de test.

Lors de J15, J19, ou J26.

Chaque jour où une session de test avait lieu se sont déroulées deux à quatre sessions de test, pour des raisons de disponibilité des participants. Toutes les sessions de test étaient

identiques entre elles et avaient lieu dans une salle différente de celle où s'était déroulé l'apprentissage. Les participants des premiers créneaux horaires ne devaient pas dévoiler aux participants des créneaux suivants le contenu de la phase de test. Les participants étaient assis à environ 2 à 5 mètres du tableau de projection. Ils devaient éteindre leur téléphone, mettre leur lunettes, ne garder sur la table qu'un stylo, et garder le silence jusqu'à la toute fin de la session. Trois tests ont été réalisés : une tâche de rappel indicé, une tâche de rappel libre des pseudomots, et une tâche de jugement du nombre d'occurrences des paires.

Tâche de rappel indicé.

Un carnet était distribué à chaque participant. Ceux-ci y écrivaient leur nom en première page. Les consignes étaient ensuite énoncées.

Une croix de fixation centrale apparaissait pendant 4 secondes. Un pseudomot apparaissait ensuite seul au centre de l'écran pendant 10 secondes. En haut de l'écran apparaissait le numéro correspondant à la page du carnet, de 1 à 32. Les participants devaient écrire le mot correspondant au pseudomot, puis cocher la case correspondant à leur degré de confiance pour leur réponse. À la fin des 10 secondes retentissait un bip sonore afin de signaler le moment précis de tourner la page du carnet. Il était interdit de revenir en arrière une fois la page tournée. Les participants étaient incités à répondre au hasard s'ils ne connaissaient pas la réponse.

Les 30 pseudomots cibles étaient testés. Les pseudomots n° 1 et 32 étaient des pseudomots issus de paires distractrices. Les pseudomots testés ainsi que leur ordre de passation étaient les mêmes pour tous les participants. La succession des items au cours du test respectait les mêmes règles que celles de l'apprentissage (i.e., pas deux fois le même agencement à la suite, ni deux fois la même catégorie sémantique à la suite).

Tâche de rappel libre des pseudomots.

Suite à la tâche de rappel indicé, les carnets de réponse étaient ramassés et les feuilles de rappel libre distribuées à l'envers. La consigne était d'écrire lisiblement le plus grand nombre de pseudomots seuls (i.e., sans le mot associé), et ce le plus correctement possible. Les participants retournaient leur feuille, inscrivaient leur nom, et les 4 minutes de rappel libre commençaient. Une à deux minutes environ s'étaient écoulées entre la fin de la tâche de rappel indicé et le début de la tâche de rappel libre. Suite au rappel libre, les feuilles de réponse étaient ramassées.

Tâche de jugement de fréquence des présentations.

Les formulaires de jugement de fréquence étaient distribués. À l'écran étaient projetées les paires mot-pseudomot, et les participants devaient entourer, pour chacune d'elles, le chiffre correspondant au nombre de fois où ils pensaient avoir vu cette paire au cours des cinq sessions de la phase d'apprentissage.

Parmi les 40 paires présentées se trouvaient les 30 paires cibles et 10 paires distractrices. Comme dans les tâches précédentes, deux paires successives ne faisaient pas partie du même agencement ni de la même catégorie sémantique.

Suite à la tâche de jugement de fréquence, les participants recevaient le dédommagement de 35 ou de 20 euros.

Lors de J 60.

Les participants de la première période expérimentale étaient conviés à une séance d'explication de l'expérience 60 jours après le début de la phase d'apprentissage. Ils

n'en avaient pas été avertis mais une tâche de rappel indicé avait lieu ce jour-là. Douze participants (3 du groupe A-J15 ; 3 du groupe B-J15 ; 1 de groupe B-J19 ; 3 du groupe C-J15 ; et 2 du groupe C-J19) sont revenus. La procédure était identique à celle présentée auparavant, à l'exception de l'ordre des items qui était modifié. Cette session n'a pas eu lieu pour la deuxième période expérimentale.

4. 4. 3. Résultats

Les données utilisées pour les analyses statistiques suivantes n'ont pas été corrigées.

Rappel indicé à J15, J19, ou J26

Le nombre total de bonnes réponses par participant à la tâche de rappel indicé s'échelonnait de 0 à 25 sur 30, c'est-à-dire de 0 % à 83,3 % de bonnes réponses. La moyenne s'élevait à 10,14 bonnes réponses sur 30, c'est-à-dire 33,8 % de bonnes réponses. Deux participants, qui appartenaient au groupe testé à J26, ont donné 0 % de bonnes réponses. Les scores moyens en fonction des agencements et du jour de test sont présentés dans la Figure 19.

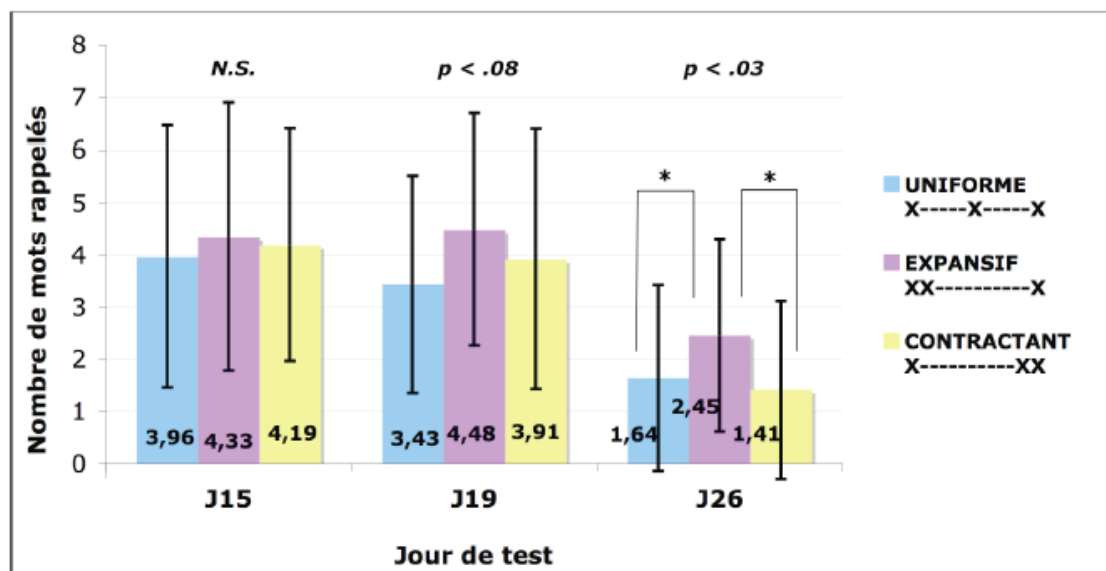


Figure 19 : Nombre moyen de mots rappelés par agencement (Uniforme, Expansif, Contractant) en fonction du délai de rétention (Jour de test à J15, J19, ou J26) à la tâche de rappel indicé. Le nombre maximal de bonnes réponses était de 10. Les barres d'erreurs représentent les écart-types. La significativité de l'effet de l'agencement pour chaque Jour de test est représentée, et les astérisques représentent les contrastes significatifs.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant) et comme facteur inter-sujet le Jour de test (avec trois modalités : J15, J19, J26). Cette analyse a révélé un effet significatif de l'agencement temporel, $F(2,138) = 4.51$, $p < .02$, $\eta^2_p = 0.06$, ainsi qu'un effet significatif du facteur Jour de test, $F(2,69) = 11.53$, $p < .001$, $\eta^2_p = 0.25$. L'interaction Agencement * Jour de test ne s'est pas révélée significative, $F(4,138) < 1$.

Des analyses secondaires ont été réalisées dans le but de vérifier l'absence des effets de liste et de groupe. Une analyse de variance a été réalisée sur le nombre de bonnes

réponses avec comme facteur intra-sujet la liste de paires (L1, L2, L3). Cette analyse n'a pas révélé d'effet significatif de la liste de paires : $F(2,142) = 1.61, p < .21$. Ce résultat confirme le fait qu'il n'y avait pas de différence de performance en fonction des listes, tout comme dans les Expériences 1 et 2.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur inter-sujet le groupe de contrebalancement (A, B, C). Cette analyse n'a pas révélé d'effet significatif du groupe de contrebalancement : $F(2,69) < 1$. Ce résultat confirme le fait que les groupes de contrebalancement étaient homogènes en termes de performance pour le rappel indicé, tout comme dans les Expériences 1 et 2.

Bien que l'interaction ne se soit pas révélée significative, nous avons exploré les effets de l'agencement temporel pour chaque modalité du jour de test, qui concernaient des groupes de participants différents. Nous justifions ce choix en invoquant les trois raisons suivantes. D'une part, le profil général de l'effet d'agencement (avantage numérique de l'agencement expansif) est similaire pour les trois jours de test, ce qui rend difficile l'obtention d'une interaction significative. D'autre part, le fait que les groupes correspondant aux trois jours de test sont indépendants peut également masquer l'interaction. Enfin, les analyses de l'effet de l'agencement pour chaque DR étaient planifiées à l'avance.

Effet de l'agencement à J15.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement. Cette analyse n'a pas révélé d'effet significatif de l'agencement temporel, $F(2,52) < 1$. Ainsi, chez les participants ayant réalisé le test de rappel indicé lors de J15, il n'y avait pas de différence de performance selon que l'agencement temporel des items ait été uniforme (3,96), expansif (4,33), ou contractant (4,19).

Effet de l'agencement à J19.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement. Cette analyse a révélé un effet tendanciel de l'agencement temporel, $F(2,44) = 1.70, p < .08$. Ainsi, chez les participants ayant réalisé le test de rappel indicé lors de J19, les différences de performance entre les agencements temporels étaient presque significatives, avec dans l'ordre décroissant l'agencement expansif (4,48), contractant (3,91), et uniforme (3,49).

Effet de l'agencement à J26.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement. Cette analyse a révélé un effet significatif de l'agencement temporel, $F(2,42) = 4.19, p < .03$.

Des analyses de contraste ont révélé que les participants produisaient plus de bonnes réponses dans la condition expansive (2,45) que dans la condition uniforme (1,64), $F(1,42) = 4.85, p < .04$, et que dans la condition contractante (1,41), $F(1,42) = 7.92, p < .01, \eta^2_p = 0.17$. Il n'y avait pas de différence significative entre les conditions uniforme et contractante, $F(1,42) < 1$.

Rappel libre à J15, J19, ou J26

Seuls les pseudomots cibles parfaitement orthographiés étaient considérés comme des bonnes réponses.

Le nombre total de bonnes réponses par participant à la tâche de rappel indicé s'échelonnait de 0 à 15 sur 30, c'est-à-dire de 0 % à 50 % de bonnes réponses. La moyenne s'élevait à 6,79 bonnes réponses sur 30, c'est-à-dire 22,6 % de bonnes réponses. Deux participants, qui appartenaient au groupe testé à J26, ont donné 0 % de bonnes réponses. Les scores moyens en fonction des agencements et du jour de test sont présentés dans la Figure 20.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant) et comme facteur inter-sujet le Jour de test (avec trois modalités : J15, J19, J26). Cette analyse a révélé un effet significatif de l'agencement temporel, $F(2,138) = 3.73, p < .03, \eta^2_p = 0.05$, ainsi qu'un effet significatif du facteur Jour de test, $F(2,69) = 7.53, p < .005, \eta^2_p = 0.18$. L'interaction Agencement * Jour de test ne s'est pas révélée significative, $F(4,138) < 1$.

Bien que l'interaction ne se soit pas révélée significative, nous avons exploré les effets de l'agencement temporel pour chaque modalité du jour de test, qui concernaient des groupes de participants différents.

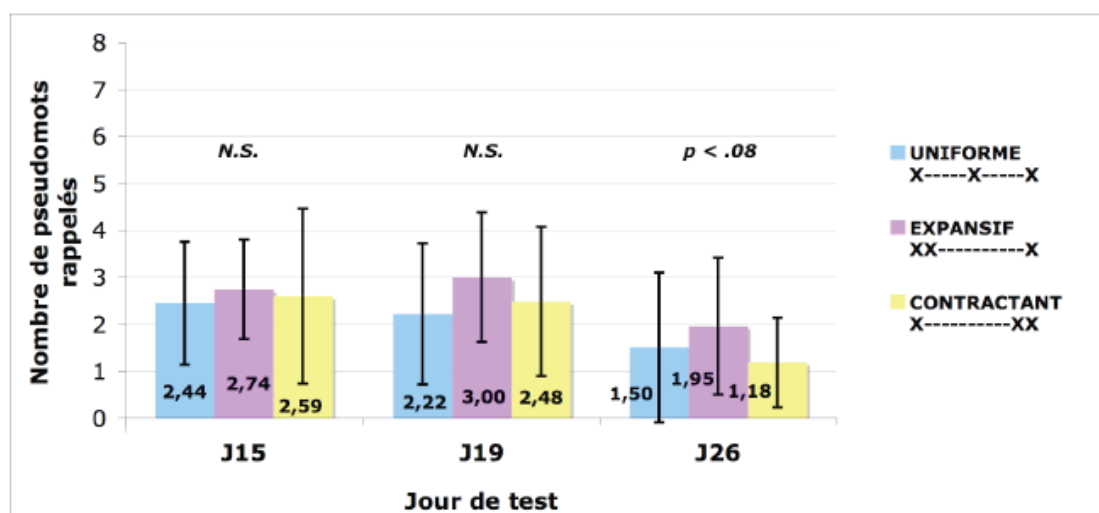


Figure 20 : Nombre moyen de pseudomots rappelés par agencement (Uniforme, Expansif, Contractant) en fonction du délai de rétention (test à J15, J19, ou J26) à la tâche de rappel libre. Le nombre maximal de bonnes réponses était de 10. Les barres d'erreurs représentent les écart-types. La significativité de l'effet de l'agencement pour chaque Jour de test est représentée.

Effet de l'agencement à J15.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement. Cette analyse n'a pas révélé d'effet significatif de l'agencement temporel, $F(2,52) < 1$. Ainsi, chez les participants ayant réalisé le test de rappel libre lors de J15, il n'y avait pas de différence de performance selon que l'agencement temporel des items ait été uniforme (2,44), expansif (2,74), ou contractant (2,59).

Effet de l'agencement à J19.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement. Cette analyse n'a pas révélé d'effet significatif de l'agencement temporel, $F(2,44) = 1.90$, $p < .17$. Ainsi, chez les participants ayant réalisé le test de rappel libre lors de J19, il n'y avait pas de différence de performance selon que l'agencement temporel des items ait été uniforme (2,22), expansif (3,0), ou contractant (2,48).

Effet de l'agencement à J26.

Une analyse de variance a été réalisée sur le nombre de bonnes réponses avec comme facteur intra-sujet l'Agencement. Cette analyse a révélé un effet tendanciel de l'agencement temporel, $F(2,42) = 3.79$, $p < .08$, $\eta^2_p = 0.05$. Ainsi, les différences de scores observés entre les agencements uniforme (1,50), expansif (1,95), et contractant (1,18), étaient proches de la significativité.

Jugement de fréquence des paires à J15, J19, ou J26

La fréquence estimée s'échelonnait de 1 à 7 toutes conditions confondues avec une estimation moyenne de 2,92 occurrences. Les fréquences estimées moyennes en fonction de l'agencement et du jour de test sont présentées dans la Figure 21.

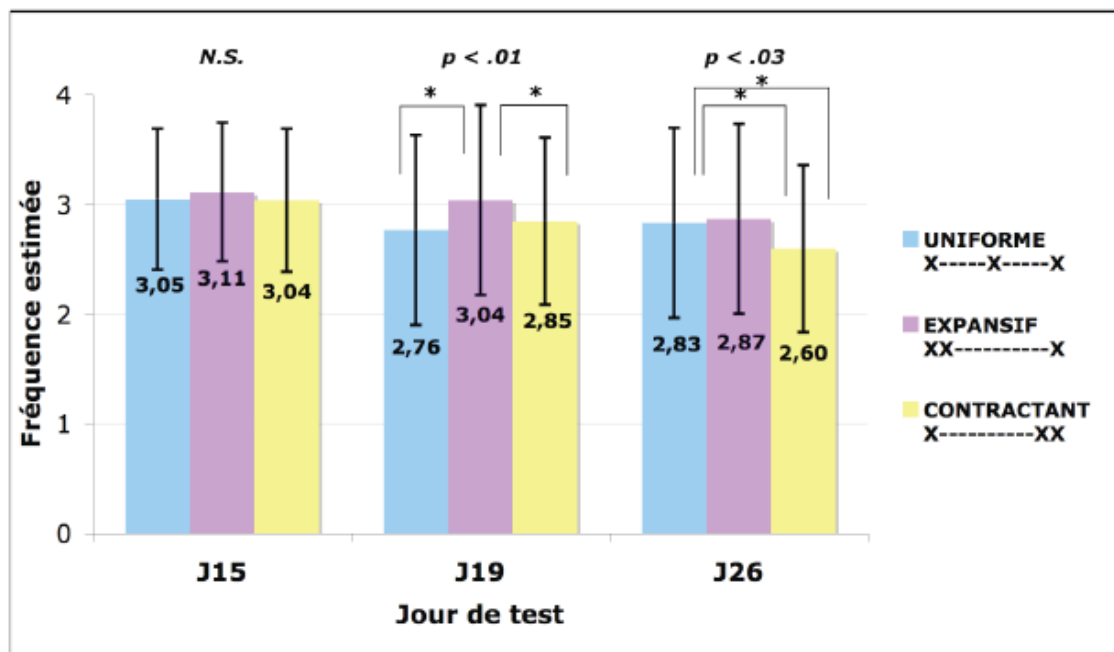


Figure 21 : Nombre moyen d'occurrences estimé par agencement (Uniforme, Expansif, Contractant) en fonction du délai de rétention (test à J15, J19, ou J26) dans la tâche de jugement de fréquence. Le nombre réel d'occurrences était de 3. Les barres d'erreurs représentent les écart-types. La significativité de l'effet de l'agencement pour chaque Jour de test est représentée et les astérisques représentent les contrastes significatifs.

Une analyse de variance a été réalisée sur la fréquence estimée avec comme facteur intra-sujet l'Agencement (avec trois modalités : Uniforme, Expansif, Contractant) et comme facteur inter-sujets le Jour de test (avec trois modalités : J15, J19, J26). (Chaque cellule de

l'analyse contenait la fréquence estimée moyenne des 10 items d'un agencement donné, et ce pour un participant donné.) Cette analyse a révélé un effet significatif de l'agencement temporel, $F(2,138) = 6.40$, $p < .01$, $\eta^2_p = 0.08$, mais pas d'effet du facteur Jour de test, $F(2,69) = 1.20$, $p = .30$. L'interaction Agencement * Jour de test était tendanciellement significative, $F(4,138) = 2.34$, $p = .058$.

Nous avons exploré les effets de l'agencement temporel pour chaque modalité du jour de test, qui concernaient des groupes de participants différents.

Effet de l'agencement à J15.

Une analyse de variance a été réalisée sur la fréquence estimée avec comme facteur intra-sujet l'Agencement. Cette analyse n'a pas révélé d'effet significatif de l'agencement temporel, $F(2,52) < 1$. Ainsi, chez les participants ayant réalisé la tâche d'estimation de fréquence lors de J15, il n'y avait pas de différence d'estimation selon que l'agencement temporel des items ait été uniforme (3,05), expansif (3,11), ou contractant (3,04).

Effet de l'agencement à J19.

Une analyse de variance a été réalisée sur la fréquence estimée avec comme facteur intra-sujet l'Agencement. Cette analyse a révélé un effet significatif de l'agencement temporel, $F(2,44) = 5.11$, $p = .01$, $\eta^2_p = 0.10$.

Des analyses de contraste ont révélé que les participants estimaient comme plus fréquents les items ayant été présentés dans la condition expansive (3,04) que dans la condition uniforme (2,76), $F(1,44) = 9.68$, $p < .01$, et que dans la condition contractante (2,85), $F(1,44) = 4.78$, $p < .01$. Il n'y avait pas de différence significative entre les conditions uniforme et contractante, $F(1,44) < 1$.

Effet de l'agencement à J26.

Une analyse de variance a été réalisée sur la fréquence estimée avec comme facteur intra-sujet l'Agencement. Cette analyse a révélé un effet significatif de l'agencement temporel, $F(2,42) = 4.15$, $p < .03$, $\eta^2_p = 0.17$.

Des analyses de contraste ont révélé que les participants estimaient comme plus fréquents les items ayant été présentés dans la condition expansive (2,87) et uniforme (2,83) que dans la condition contractante (2,60), respectivement $F(1,42) = 7.04$, $p < .02$ et $F(1,42) = 5.29$, $p < .03$. Il n'y avait pas de différence significative entre les conditions expansive et uniforme, $F(1,42) < 1$.

Estimations de motivation, concentration et performance pendant la phase d'apprentissage

Les niveaux estimés de motivation, de concentration et de performance ont été considérés d'une façon globale, c'est-à-dire sans distinguer les participants selon le jour où ils ont été testés.

Chaque paramètre a été analysé indépendamment. Les niveaux de motivation de début de session n'ont pas pris en compte les résultats de la 1^{ère} session car les participants n'avaient pas encore réalisé de session d'apprentissage. De la même façon, les niveaux de motivation pour poursuivre l'expérience n'ont pas pris en compte les résultats de la session 5

car il ne restait alors que la session de test. Les moyennes des estimations sont présentées dans la Figure 22. (Les mêmes figures, mais représentant les niveaux en fonction des groupes de Jour de test, sont présentées dans l'Annexe n°9).

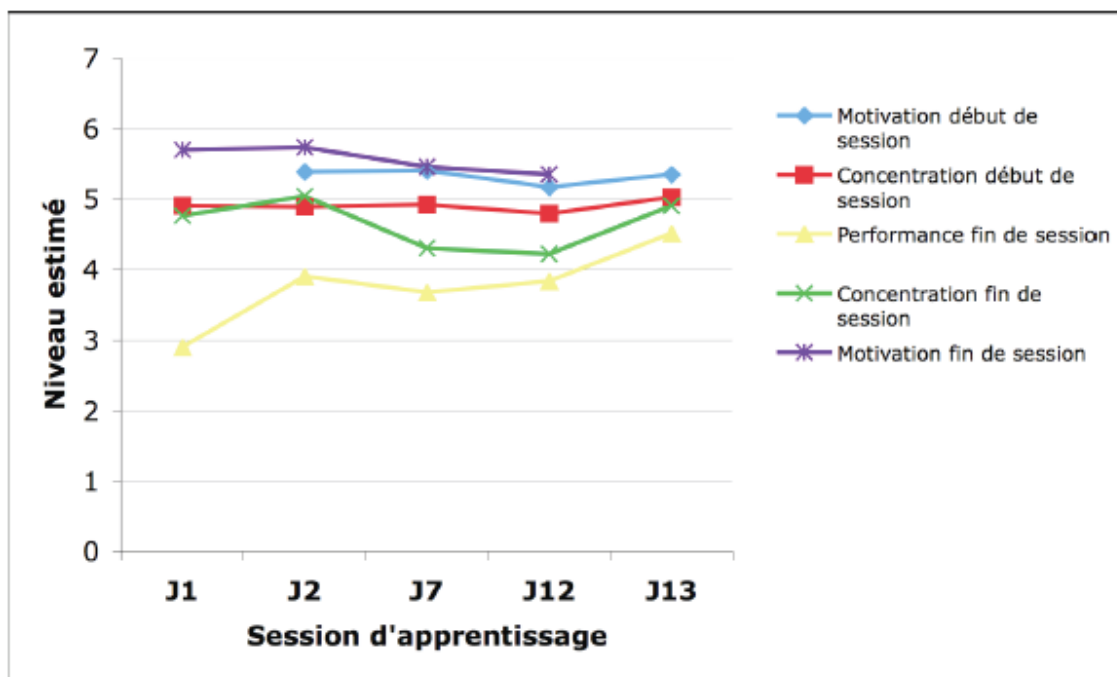


Figure 22 : Niveaux moyens de motivation estimée (en début et en fin de session), de concentration estimée (en début et en fin de session), et de performance estimée (en fin de session), en fonction des sessions d'apprentissage (J1 à J13). Les écart-types ne sont pas représentés pour ne pas surcharger la figure.

Niveau estimé de motivation avant de commencer la session d'apprentissage.

Les niveaux de motivation s'échelonnaient de 2 à 7 avec une moyenne de 5,33. L'analyse de tendance polynomiale n'a révélé aucune tendance significative : $F(1,213) < 1$ pour les tendances linéaire et quadratique, et $F(1,213) = 2.35$, $p < .13$ pour la tendance cubique. Ce résultat indique que le niveau de motivation avant de commencer la session ne diminuait pas significativement au fil des sessions d'apprentissage.

Niveau estimé de concentration avant de commencer la session d'apprentissage.

Les niveaux de concentration s'échelonnaient de 2 à 7 avec une moyenne de 4,91. L'analyse de tendance polynomiale n'a révélé aucune tendance significative, $F(1,284) < 1$ pour les tendances linéaire, quadratique, cubique, et quartique. Ce résultat indique que le niveau de concentration estimé moyen ne variait pas significativement au fil des sessions d'apprentissage.

Niveau estimé de performance a posteriori pour l'apprentissage.

Les niveaux de performance s'échelonnaient de 1 à 7 avec une moyenne de 3,77. L'analyse de tendance polynomiale a révélé une tendance linéaire significative, $F(1,284) = 79.56$,

$p < .001$, ainsi qu'une tendance cubique significative, $F(1,284) = 24.51$, $p < .001$. Les autres tendances n'étaient pas significatives, $F(1,284) < 1$ pour la tendance quadratique, et $F(1,284) = 2.39$, $p < .13$ pour la tendance quartique.

Ce résultat indique que les estimations des participants sur leur niveau de performance augmentaient au fil des sessions d'apprentissage.

Niveau estimé de concentration a posteriori pour l'apprentissage.

Les niveaux de concentration s'échelonnaient de 1 à 7 avec une moyenne de 4,65. L'analyse de tendance polynomiale a révélé une tendance quadratique significative, $F(1,284) = 9.08$, $p < .01$, ainsi qu'une tendance cubique significative, $F(1,284) = 18.89$, $p < .001$. Les autres tendances n'étaient pas significatives, $F(1,284) = 1.75$, $p < .20$ pour la tendance linéaire, et $F(1,284) = 2.07$, $p < .16$ pour la tendance quartique.

Niveau estimé de motivation pour poursuivre l'expérience.

Les niveaux de motivation s'échelonnaient de 2 à 7 avec une moyenne de 5,56. L'analyse de tendance polynomiale a révélé une tendance linéaire significative, $F(1,213) = 14.88$, $p < .001$. Les autres tendances n'étaient pas significatives : $F(1,213) < 1$ pour la tendance quadratique, et $F(1,213) = 2.02$, $p < .16$ pour la tendance cubique. Ce résultat indique que le niveau de motivation des participants pour poursuivre l'expérience diminuait significativement au fil des sessions.

Rétention à J60

En tâche de rappel indicé à J60, le nombre total de bonnes réponses par participant s'échelonnait de 5 à 18 sur 30, c'est-à-dire de 16,7 % à 60 % de bonnes réponses. La moyenne s'élevait à 10,25 bonnes réponses sur 30, c'est-à-dire 34,2 % de bonnes réponses.

L'analyse de variance de contrôle sur le nombre de bonnes réponses portant sur l'effet du groupe (A, B, C) s'est révélée significative, $F(2,9) = 5.16$, $p < .04$, indiquant un biais dans la constitution du sous-ensemble de participants, et donc que les résultats seraient à prendre avec précaution. L'analyse de variance de contrôle portant sur l'effet du jour de test (J15, J19) n'était pas pertinente étant donné que seuls 3 participants appartenaient à la modalité J19.

L'analyse de variance sur le nombre de bonnes réponses portant sur le facteur Agencement n'a pas révélé d'effet significatif, $F(2,22) < 1$.

Ces résultats ne sont pas informatifs dans la mesure où l'échantillon avait un effectif faible et que toutes les combinaisons des facteurs Groupe et Jour de test n'étaient pas représentées.

Cependant, afin d'avoir une idée de la rétention des items sur le long terme, nous avons analysé comparativement les résultats des 12 participants à la session de J60 avec les performances de ces mêmes participants à J15 ou J19, au rappel indicé. Les performances moyennes des participants lors des deux phases de test sont représentées dans la Figure 23.

Le profil de l'effet d'agencement lors du test à J15 ou J19 n'est pas représentatif du pattern observé dans l'ensemble de l'échantillon à ce moment-là, ce qui confirme l'existence d'un biais d'échantillonnage à J60. À titre indicatif, voici les chiffres représentant la diminution des scores entre J15-J19 et J60 en fonction des agencements. Pour l'agencement uniforme, les scores ont diminué en moyenne de 0,25 points, ce qui

représente 7,1 % de diminution. Pour l'agencement expansif, les scores ont diminué en moyenne de 0,67 points, ce qui représente 15,7 % de diminution. Pour l'agencement contractant, les scores ont diminué en moyenne de 1 point, ce qui représente 22,6 % de diminution. Ainsi, ces résultats suggèrent, avec toutes les réserves émises auparavant, que l'oubli est le moins important pour les items ayant été appris dans la condition uniforme, moyenne pour ceux de la condition expansive et le plus important pour ceux de la condition contractante. Nous ne discuterons pas davantage de ces données.

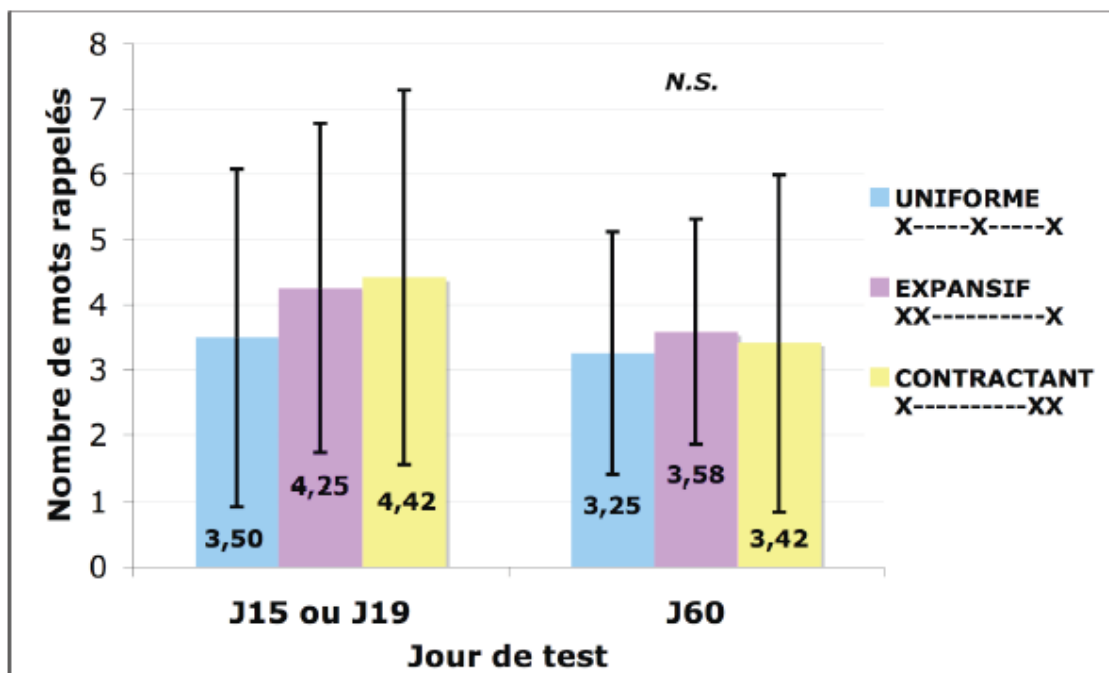


Figure 23 : Nombre moyen de mots rappelés par agencement (Uniforme, Expansif, Contractant) dans la tâche de rappel indicé à J15 ou J19 et à J60 pour les participants présents à J60. Le nombre maximum de bonnes réponses était de 10. Les barres d'erreurs représentent les écart-types. La significativité de l'effet de l'agencement à J60 est représentée.

4. 4. 4. Discussion

Rappel du principe de l'Expérience 3

Dans cette expérience, trois agencements temporels de trois présentations d'items sur une période d'apprentissage de 13 jours étaient comparés : les agencements expansif (XX-----X), uniforme (X-----X-----X), et contractant (X-----XX). Une tâche de rappel indicé, une tâche de rappel libre des pseudomots, et une tâche de jugement de fréquence se déroulaient le 15ème, le 19ème, ou le 26ème jour, selon les participants. De plus, ceux-ci devaient à chaque session d'apprentissage estimer un ensemble de paramètres liés à la motivation, à la concentration et à la performance d'apprentissage.

Résumé des résultats

Cette expérience a montré, dans la lignée des Expériences 1 et 2 et de Tsai (1927, Expériences 1 et 2), la supériorité de l'agencement expansif sur les agencements uniforme et contractant en rappel indicé. Le second résultat important est que cet avantage de

l'agencement expansif apparaissait progressivement avec l'allongement de la durée du DR. En effet, l'effet de l'agencement temporel n'était pas significatif chez les participants ayant réalisé le test à J15, il était tendanciel chez les participants ayant réalisé le test à J19, et il était significatif chez les participants ayant réalisé le test à J26.

Les performances en tâche de rappel libre présentaient le même profil que pour la tâche de rappel indicé, avec des performances moins élevées, mais sans présenter de significativité.

La tâche de jugement de fréquence a également révélé un profil comparable aux performances de rappel indicé : à J15, l'effet de l'agencement n'était pas significatif ; à J19, l'effet était significatif, avec une estimation plus élevée pour les paires issues de l'agencement expansif par rapport aux autres agencements ; enfin, à J26, la fréquence estimée était plus élevée pour les paires issues des agencements expansif et uniforme que pour celles de l'agencement contractant.

Des estimations des niveaux de motivation, de concentration et de performance ont été demandées aux participants pour contrôler que des variations de ces paramètres au fil des sessions ne puissent pas expliquer les différences de performances finales entre les agencements. Les observations ont montré que, de session en session, ni le niveau de motivation en début de session ni le niveau de motivation en fin de session ne diminuaient significativement. Ainsi, il n'y a pas de raison de penser qu'une diminution du niveau de motivation au fil des sessions d'apprentissage puisse expliquer que l'agencement contractant (et dans une moindre mesure, l'agencement uniforme) soient associées à un apprentissage moins efficace et donc à des performances mnésiques finales moins élevées. Les niveaux de concentration avant et après l'apprentissage ne présentaient pas non plus d'évolution significative au fil des sessions. Cependant, et comme dans l'Expérience 2, l'absence de significativité ne signifie pas l'absence de l'effet.

De façon intéressante, le niveau de performance estimée augmentait clairement de session en session, indiquant que les participants se sentaient de plus en plus à l'aise dans leur apprentissage, et nous proposons deux raisons à cela. Premièrement, la répétition des items au fil des sessions a probablement rassuré les sujets sur la difficulté de la tâche qu'on leur demandait. Deuxièmement, les sujets ont probablement maîtrisé de plus en plus les techniques mnémoniques au fur et à mesure des sessions.

Implications des résultats

L'hypothèse de supériorité de l'agencement expansif, prédite par la théorie de la récupération en phase d'étude, a donc été globalement vérifiée. Ainsi, avec les agencements relativement extrêmes comparés dans cette expérience, cette expérience apporte un élément empirique supplémentaire en faveur de la généralité de l'avantage de l'agencement expansif. L'hypothèse contradictoire de la supériorité de l'agencement contractant pour un DR court, puis uniforme pour un DR long, prédite par la théorie de la variabilité de l'encodage, n'a donc pas été vérifiée.

Un effet inattendu et très intéressant est apparu. L'effet de l'agencement temporel s'est révélé progressivement à mesure que le DR s'allongeait, aboutissant à une situation dans laquelle, à J26, l'agencement expansif était plus efficace que les deux autres agencements. Nous choisirons d'interpréter ce résultat de la façon suivante : l'agencement expansif favoriserait la création de traces mnésiques plus résistantes dans le temps, c'est-à-dire moins sensible à l'oubli et/ou aux interférences, que les autres agencements. En effet, les performances pour les trois agencements étaient similaires peu après la fin

de l'apprentissage (J15), puis divergeaient progressivement en faveur de l'agencement expansif (J19 et J26).

Avant de discuter plus avant de la notion de taux d'oubli, il nous semble utile de discuter de deux points méthodologiques : le premier relatif à la différence de rémunération accordée aux différents sous-groupes, et le second relatif à la possible contamination des tâches mnésiques entre elles.

Points méthodologiques

Différence de rémunération entre les deux périodes expérimentales.

Il peut être important de mentionner que tous les participants de l'expérience n'ont pas bénéficié du même niveau de rémunération puisque les 22 participants de la deuxième période expérimentale (J26) ont reçu un dédommagement de 20 euros, tandis que les 50 participants de la première période expérimentale (J15 et J19) ont reçu un dédommagement de 35 euros. Ce sont des raisons purement budgétaires qui sont à l'origine de cette différence, et il aurait été plus approprié de fournir le même dédommagement à tous les sujets. Il est possible que la moindre rémunération accordée lors de la deuxième période expérimentale ait résulté en un biais de sélection des participants volontaires lors du recrutement, ceux-ci étant moins motivés ou plus enclins à abandonner l'expérience une fois celle-ci commencée. Cependant, nous ne voyons pas de raison pour laquelle l'effet de l'agencement temporel aurait pu être différent chez ces participants-là à cause du dédommagement. Le fait d'observer un effet de l'agencement à J26 similaire à celui de J19 (mais en plus accentué) nous conforte dans cette idée. Par ailleurs, l'évolution des niveaux de motivation, de concentration et de performance au fil des sessions n'était pas différente dans les trois groupes (bien que les analyses comparatives n'aient pas été réalisées).

Contamination des tâches entre elles lors du test final.

Des critiques liées au fait que la tâche de rappel indicé aurait probablement influencé les performances des sujets lors des tâches de rappel libre et de jugement de fréquence doivent être faites, tout comme dans l'Expérience 2.

De la même façon que dans l'Expérience 2, les items présentés selon l'agencement expansif étaient globalement jugés comme ayant été présentés plus souvent que les items des autres agencements. Ainsi, le fait d'observer globalement des estimations de fréquence plus élevées pour l'agencement expansif nous semblent cohérent avec la théorie de la récupération en phase d'étude et de la remémoration récursive (Hintzman, 2004, 2008). En effet, l'augmentation progressive de la durée des IIR devrait favoriser les processus de récupération en phase d'étude, et ces récupérations successives des occurrences antérieures d'un même item devraient favoriser une estimation correcte du nombre d'occurrences de l'item. Au contraire, l'incapacité à récupérer les traces des occurrences antérieures devrait faire diminuer l'estimation du nombre d'occurrences. Cependant, nous n'avons pas de données à disposition, contrairement à l'Expérience 2, sur les taux de reconnaissance pendant les présentations et ne pouvons donc pas vérifier que les items reconnus se voient attribuer une fréquence d'occurrence plus élevée.

Toutefois, la durée du DR modulait l'effet principal de l'agencement, puisque les trois agencements ne présentaient pas de différence lorsque l'estimation avait lieu à J15, alors que l'effet était significatif à J26. Or, il est logique de supposer que les phénomènes de récupération en phase d'étude devraient être similaires chez tous les participants quel que

soit leur DR, puisqu'ils ont tous suivi la même phase d'apprentissage. Ainsi, la théorie de la récupération en phase d'étude ne peut pas expliquer pourquoi l'effet de l'agencement n'est pas significatif pour J15. Au contraire, cette observation va plutôt dans le sens d'une contamination des tâches mnésiques préalables sur la tâche de jugement de fréquence. Par exemple, il est possible que les participants aient réalisé leur estimation de fréquence en se basant -plus ou moins explicitement- sur le sentiment de familiarité évoqué en eux par les paires, sentiment explicité par l'issue des tâches mnésiques précédentes. On peut imaginer ce genre de discours intérieur : « Comme je me souviens bien de cet item, c'est que j'ai dû le voir un grand nombre de fois ». Pour explorer cette hypothèse, nous avons déterminé, pour chaque participant et tous agencements confondus, la fréquence moyenne estimée pour les items qui ont été correctement rappelés à la tâche de rappel indicé (3,61) d'une part, et pour les items qui n'avaient pas été rappelés (2,74) d'autre part. Nous avons dû éliminer un sujet qui n'avait donné aucune réponse correcte au rappel indicé (effectif restant = 71). Un test de Wilcoxon sur échantillons appariés a révélé que cette différence était significative (nombre de différences non nulles = 70, $T = 39,5$, $z = 7,04$, $p < .001$). Ainsi, les items qui étaient rappelés étaient également jugés comme plus fréquents. Cela ne signifie toutefois pas nécessairement que le sujet se base explicitement sur ses performances aux tâches de rappel pour réaliser son estimation de fréquence. Cela peut aussi refléter le fait qu'une propriété mnésique, comme par exemple la force de la trace, pourrait être impliquée à la fois dans le rappel indicé et dans le jugement de fréquence. Nous ne pouvons pas dissocier ces deux hypothèses.

Différence de taux d'oubli entre les différents agencements

Afin de mieux se représenter les taux d'oubli induits par chacun des agencements, nous avons transformé la Figure 19 en représentant le délai de rétention en abscisses, et les différents agencements par différentes courbes. La Figure 24 présente ces données. Chaque courbe représente les niveaux de rétention pour un agencement donné au fur et à mesure du passage du temps pendant le DR.

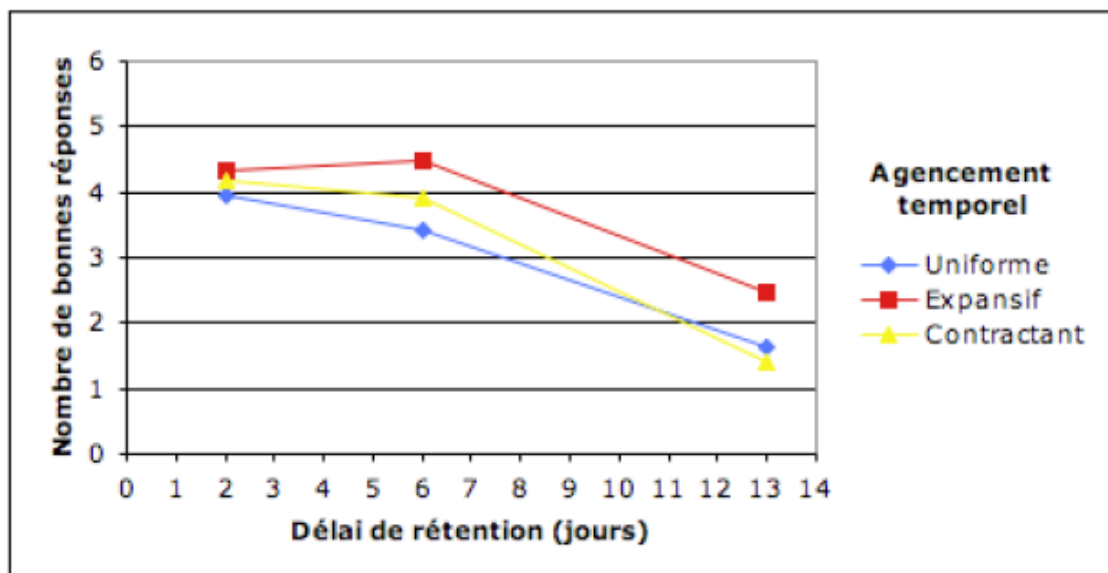


Figure 24 : Résultats de l'Expérience 3 avec en ordonnées le nombre de bonnes réponses et en abscisses le délai de rétention, en fonction de l'agencement temporel (couleur de la courbe).

La Figure 24 permet de visualiser que la courbe de rétention la plus plate est globalement celle de l'agencement expansif, puisque les diminutions de scores entre J15 et J26 sont de 2.33, 1.88, et 2.78, respectivement pour les agencements uniforme, expansif, et contractant. On remarquera que l'on n'observe pas la forme habituelle de la courbe d'oubli ; au lieu d'être fortement décroissante au début du DR puis de moins en moins fortement décroissante (e.g., Wixted & Ebbesen, 1991), la courbe est initialement relativement plate entre J15 et J19 puis diminue plus fortement ensuite. En toute logique, cette « anomalie » est due au fait que, globalement, les groupes de participants (J15, J19, et J26) n'étaient pas complètement équivalents. Une autre anomalie, due probablement à la même raison que la précédente, est l'augmentation de la performance entre J15 et J19 pour l'agencement expansif seul. D'autre part, comme seulement trois points sur cette courbe de rétention sont disponibles, il est difficile de comparer les pentes des agencements entre elles. Selon l'intervalle considéré, ce n'est pas toujours le même agencement qui produit la courbe la plus plate. Par exemple, entre J15 et J19, ce sont les items de l'agencement expansif qui étaient les moins oubliés (augmentation de 0.14), tandis qu'entre J19 et J26, ce sont ceux de l'agencement uniforme qui étaient les moins oubliés (diminution de 1.80). L'agencement contractant présente le taux d'oubli le plus important quand on le considère d'une façon générale (diminution de 2.78 entre J15 et J26).

Nous pensons que la théorie de la remémoration récursive, telle que développée chez Hintzman (2004 ; 2010) et Benjamin et Tullis (2010) pourrait expliquer ces résultats : la récupération en phase d'étude pourrait permettre non seulement de faire augmenter la rétention finale des items, mais également de ralentir leur taux d'oubli. Nous développerons ce point dans la discussion générale, à la section 6.2.2.

Il aurait été intéressant, du point de vue de cette théorie, de tester la reconnaissance des paires pendant la phase d'apprentissage et en particulier lors de P2. Il est probable que des différences très importantes de taux de reconnaissance soient apparues entre les agencements à P2, étant donné les longs IIR employés dans cette expérience. Il est possible également que toutes les paires ne soient pas reconnues lors de P3, contrairement à l'Expérience 2.

D'une façon générale, et de façon intéressante, les résultats sont apparus cohérents avec les prédictions du modèle de Mozer et al. (2009), décrites dans Lindsey et al. (2009). Or ce modèle simule une combinaison des hypothèses de la variabilité de l'encodage et de la récupération en phase d'étude. Cette expérience apporte donc indirectement des arguments expérimentaux en faveur d'un modèle mixte.

Il est nécessaire de rappeler ici que ces courbes ne sont pas réellement des courbes de rétention dans la mesure où des sujets différents appartiennent aux trois groupes de sujets (J15, J19, J26). En effet, pour pouvoir déterminer correctement des courbes et des taux d'oubli, la procédure adaptée est de tester à différents intervalles un sous-ensemble du matériel expérimental pour un même sujet (comme l'ont fait par exemple Litman & Davachi, 2008). Il est nécessaire pour confirmer ces observations, qui ne font que *suggérer* des pentes d'oubli différentes, de répliquer ces résultats avec une méthode appropriée. Toutefois, le fait de tester chaque participant plusieurs fois sur un sous-ensemble du matériel augmente la difficulté du protocole, puisque les participants doivent alors se rendre disponibles pour des sessions supplémentaires de test.

4. 4. 5. Discussion méthodologique pour les Expériences 1 à 3

Tâches mnésiques

La première critique des protocoles expérimentaux porte sur le fait d'avoir enchaîné dans la session expérimentale finale plusieurs tâches mnésiques, en l'occurrence, un rappel indicé, un rappel libre, et un jugement de fréquence, portant sur le même matériel. Cette critique a déjà été discutée, mais mérite d'être reproduite ici. Il est indéniable que la réalisation de la première tâche a influencé le résultat des tâches suivantes, puisque dans la tâche de rappel indicé, l'ensemble des pseudomots cibles ont été présentés. Ainsi, lors de la tâche de rappel libre des pseudomots qui a suivi, les participants pouvaient s'aider des pseudomots qu'ils venaient juste de voir, et les résultats étaient probablement différents de ceux obtenus si la tâche avait été réalisée en premier. De la même façon, la tâche de jugement de fréquence des paires pouvait être influencée par la réalisation des tâches antérieures. Nous avons souhaité réaliser ces tâches additionnelles par rapport à la tâche initiale de rappel indicé de façon quelque peu exploratoire, étant donné le faible nombre de données récoltées pour un protocole expérimental si complexe, et dans le cas où des résultats apparaîtraient uniquement dans les dernières tâches. Il faut noter par ailleurs que dans l'étude de Cepeda et al. (2008), deux tests étaient réalisés sur le même matériel (i.e., un test de rappel puis un questionnaire à choix multiple) et les résultats étaient analysés de façon indépendante. En particulier, le fait d'ajouter la tâche de jugement de fréquence provenait de la constatation que certains participants de l'Expérience 1 semblaient très surpris d'apprendre, lors du débriefing, que chaque paire avait été présentée exactement quatre fois. Un paradigme idéal aurait été de faire porter les différentes tâches sur un sous-ensemble différent du matériel, mais nous avons préféré conserver le caractère primordial de la tâche de rappel indicé. Ces remarques nous ont conduite, d'ailleurs, à considérer presque exclusivement dans nos conclusions les résultats à la tâche de rappel indicé.

Interférences entre les items

Nous avons mis en avant le fait que, dans nos expériences, l'agencement temporel était une variable intra-sujets contrairement aux expériences réalisées antérieurement, en particulier Gay (1973) et Tsai (1927). Ce type de design intra-sujets nous a permis de recourir à un plus faible effectif de participants, étant donné la grande variabilité des scores entre les sujets. Cependant, il est possible que ce design intra-sujets ait pu favoriser des phénomènes d'interférences entre les items cibles. En effet, dans nos expériences, chaque sujet était confronté à tous les agencements en même temps. Imaginer que des interférences ont pu se produire revient à imaginer, par exemple, que le fait d'apprendre *chat-sanumi* dans un agencement donné a pu perturber l'apprentissage de *chien-pocuga* dans un autre agencement, et que ces interactions ont pu être davantage favorables à un agencement plutôt qu'à un autre. En particulier, on pourrait imaginer que c'est uniquement ce jeu d'interférences qui a amené à l'avantage général de l'agencement expansif puisque c'est dans cette condition que les items sont appris « en premier », si l'on met de côté le fait que tous les items cibles aient été vus le premier jour de la phase d'étude. Il est possible que les items vus les premiers ensuite, à J2, aient été favorisés par rapport aux autres, selon une sorte d'« effet de primauté à long terme », c'est-à-dire bénéficiant d'un statut mnésique particulier parce que vu au début. Ensuite, l'apprentissage des autres items aurait pu pâtir d'interférences proactives dues aux items de l'agencement expansif. De la même façon, l'apprentissage des items de l'agencement expansif auraient pu également subir des interférences rétroactives de la part des items des autres agencements. Ainsi, on ne peut pas exclure l'existence de tels phénomènes d'interférences pro- et rétroactives, ni exclure que ces effets aient participé à avantager ou désavantager certains agencements par rapport aux autres. Il est également possible que certains de ces effets se soient mutuellement annulés. Le fait de globalement répliquer les observations des expériences

réalisées en inter-sujets (Gay, 1973 ; Tsai, 1927) nous pousse à considérer ces effets comme minimes. Enfin, il est difficile d'expliquer comment ces effets d'interférences, s'ils étaient la source des résultats observés, pourraient expliquer la modulation de l'effet de l'agencement par le délai de rétention observée dans l'Expérience 3.

Une autre source d'interférence pourrait également exister du fait de l'utilisation des items distracteurs. En effet, afin de pouvoir réaliser une comparaison des agencements en intra-sujets, il a été nécessaire d'inclure des paires distractrices afin de présenter le même nombre d'items dans chacune des sessions. Premièrement, selon nous, l'inconvénient éventuel apporté par les paires distractrices, sous forme d'interférence pro et rétroactive, est de moindre importance que la situation alternative dans laquelle un nombre différent d'items est présenté dans chaque session. En effet, dans une session comportant moins d'items, on peut penser que chaque item aurait pu bénéficier de plus de ressources attentionnelles de la part des sujets. De plus, moins d'interférences auraient eu lieu dans les sessions les plus courtes, favorisant les items présentés dans ces sessions, par rapport aux items présentés dans des sessions plus longues. Deuxièmement, et comme nous l'avons déjà mentionné, une expérience menée en inter-sujets aurait nécessité un très grand nombre de sujets, étant donné la grande variabilité entre les participants du point de vue des performances. Une telle étude n'aurait effectivement pas présenté les problèmes d'interférences mettant en jeu les paires cibles et les paires distractrices.

Absence de contrôle de l'apprentissage à P1

Il existe une différence importante entre la procédure d'apprentissage utilisée dans les Expériences 1 à 3 et celle utilisée dans un certain nombre d'études récentes sur l'effet de la pratique distribuée à long terme (e.g., Bahrick et al., 1993 ; Cepeda et al., 2008 ; 2009 ; Cull, 2000). En effet, ces études suivent les recommandations préconisées par Bahrick (1979), en faisant en sorte de s'assurer que les items, lors de la première session d'apprentissage, sont appris jusqu'à un certain critère. Par exemple, dans la première session d'apprentissage de l'étude de Cepeda et al. (2008), les mots swahili étaient présentés avec le mot anglais correspondant, puis testés en rappel indicé, un par un, avec feedback jusqu'à ce que chaque item ait bénéficié d'une réponse correcte. Dans nos expériences, un tel contrôle n'avait pas lieu puisqu'il s'agissait de simples présentations de l'item, ressemblant en cela à la procédure classique utilisée dans les études plus anciennes sur l'effet d'espacement, comme par exemple dans Glenberg et Lehman (1981).

Ainsi, il est possible que nous ayons pu observer d'autres effets de l'agencement temporel si nous avons utilisé cette procédure. Cependant, nous ne pensons pas que les résultats auraient été drastiquement différents pour la raison suivante : comme les items de tous les agencements ont été présentés le premier jour, de façon mélangée dans la liste, il est raisonnable de postuler que l'apprentissage lors première session ait été globalement de même niveau pour tous les items, indépendamment de leur agencement. Toutefois, de nouvelles expérimentations seraient nécessaires pour vérifier ces hypothèses.

Effet de la consigne d'apprentissage

Il est pertinent de s'interroger sur les effets qu'ont pu avoir les consignes d'apprentissage sur les comportements des participants. Il s'agissait pour eux d'apprendre les paires mot-pseudomot en s'aidant de différentes stratégies connues pour aider la mémorisation : chercher dans le pseudomot des éléments rappelant des mots réels, fabriquer une image mentale reliant ces éléments avec le mot, et fabriquer mentalement des phrases dans lesquelles le mot était remplacé par le pseudomot. Initialement, notre objectif en proposant

ces stratégies aux sujets était d'homogénéiser les performances dans la mesure où les participants connaissant déjà ces techniques auraient été favorisés par rapport aux participants naïfs. L'utilisation de ces techniques a-t-elle pu influencer l'effet de l'agencement temporel observé ? On pourrait avancer que l'avantage de l'agencement expansif peut provenir du fait que les participants auraient utilisé ces moyens mnémotechniques davantage au début de la phase d'apprentissage qu'à la fin, oubliant progressivement de les mettre en pratique au fil des sessions. Toutefois, l'utilisation de ces techniques a été rappelée aux sujets au début de chaque session d'apprentissage. De plus, dans l'Expérience 3, les participants estimaient que leur performance augmentait de session en session, ce qui va également à l'encontre de cet argument. Cependant, on ne peut pas affirmer avec certitude que les mêmes résultats auraient été observés si on avait laissé aux sujets toute liberté dans le choix de leur stratégie. Il serait intéressant d'avoir des informations sur les stratégies réellement utilisées, item par item, en direct (comme par exemple dans Bahrack & Hall, 2005), mais cela a nécessairement pour conséquence de modifier la phase d'apprentissage.

D'autre part, nous avons également donné comme consignes aux sujets de ne pas « réviser » volontairement les items en dehors des sessions d'apprentissage, et cette consigne était répétée lors de chaque session. Rien ne nous garantit que les sujets ont réellement respecté ces consignes. Peut-être que certains ont tout de même repensé aux paires une fois sortis de la session expérimentale, volontairement ou de façon automatique. Dans cette perspective, la supériorité de l'agencement expansif pourrait s'expliquer par le fait que ces réactivations du matériel aient pu avoir lieu davantage en début de phase d'apprentissage, bien que nous n'ayons aucun indice nous incitant à le penser. En tout état de cause, il est impossible de contrôler les réactivations mentales se produisant chez les sujets, et cela fait partie des caractéristiques d'un protocole expérimental écologique.

Chapitre 5 : Effet de la variation du contexte

After more than a century of research on episodic memory, there is still little agreement on mechanisms that underlie the effects of repetition. Hintzman (2010)

5. 1. Expérience 4

5. 1. 1. Introduction

L'Expérience 4 avait pour objectif de tester la théorie de la variabilité de l'encodage comme explication de l'effet de pratique distribuée. La théorie de la variabilité de l'encodage (e.g., Glenberg, 1979 ; voir le chapitre 2) postule qu'un long intervalle inter-répétition (IIR) est bénéfique à la récupération finale d'un item, et ce pour les raisons suivantes. Premièrement, selon Glenberg (1979), les éléments contextuels sont encodés dans la trace mnésique représentant l'item et participent à la réactivation de cette trace lors de la récupération ultérieure de l'item. Deuxièmement, le passage du temps a pour conséquence de faire fluctuer graduellement les éléments contextuels. Troisièmement, la différence de contexte entre les épisodes P1 et P2 est donc plus importante pour un long IIR que pour un court IIR. Finalement, puisqu'il est bénéfique d'encoder les items dans des contextes les plus

variés possible, alors un IIR long mènera à des performances de récupération ultérieures plus élevées.

Cette théorie a été développée et testée principalement dans le cadre d'études à court terme, c'est-à-dire dans lesquelles des items sont présentés au sujet au sein d'une session expérimentale, et en général, au sein d'une liste. Dans cette liste, un item donné est répété, et durant l'intervalle séparant ses occurrences (i.e., l'IIR), un nombre variable d'autres items est présenté. Les items sont ensuite testés quelques minutes plus tard par une tâche de rappel. Dans ce type d'études, on conçoit sans difficulté que le contexte fluctue de façon continue tout au long de l'expérience. En effet, au fil des minutes, des éléments environnementaux peuvent se modifier graduellement (par exemple, la luminosité), l'état interne du sujet fluctue également (par exemple, son niveau de stress diminue et sa faim augmente), et enfin la suite des items présentés tout au long de l'expérience constitue également un contexte qui évolue.

On peut toutefois s'interroger sur le fait que la théorie de la variabilité de l'encodage est une explication valide sur le long terme. En effet, dans le cadre d'une étude sur le long terme où les répétitions ont lieu lors de jours différents, il semble moins évident de penser que le contexte fluctue régulièrement entre les occurrences. Par exemple, il ne semble pas certain que le contexte environnemental fluctue davantage entre P1 et P2 lorsque l'IIR est de sept jours par rapport au cas où l'IIR est d'un jour. Il est même possible d'envisager qu'au contraire, après un IIR de sept jours, le contexte à P2 soit plus proche de celui de P1 qu'après un IIR d'un jour. En effet, si l'intervalle P1-P2 dure 7 jours, alors P1 et P2 se déroulent le même jour de la semaine ; or les jours de la semaine sont, pour la plupart des individus, marqués par des habitudes hebdomadaires -dans le cas des étudiants par exemple, les mêmes cours ou loisirs sont suivis chaque semaine, les mêmes lieux sont fréquentés, les mêmes déplacements sont réalisés, etc. Au contraire, deux journées successives peuvent être très différentes, par exemple lorsqu'on compare un jour de cours comme le vendredi à un jour de fin de semaine comme le samedi. En tout état de cause, on peut penser que des IIR de l'ordre de plusieurs heures interrompent la continuité de la fluctuation contextuelle puisque entre les sessions expérimentales, les participants aux expériences retrouvent leurs activités habituelles, leur « vie réelle », contrairement aux expériences sur le court terme. L'évolution graduelle du contexte n'a donc pas lieu a priori dans ce cas. Un autre élément nous conduit à nous interroger sur les effets de contexte comme explication des effets de pratique distribuée à long terme. Il s'agit de l'observation que, dans les expériences étudiant l'effet d'intervalle sur une longue échelle de temps, les expérimentateurs prennent soin de réaliser les sessions d'apprentissage dans des contextes expérimentaux les plus homogènes possibles, c'est-à-dire en évitant au maximum de changer le contexte environnemental entre les sessions. Ainsi, l'explication de l'effet de la pratique distribuée à long terme par le principe de la variabilité de l'encodage nous semble sujet à questionnement.

Avant d'aborder plus avant le lien entre les effets de répétition et la variabilité de l'encodage, il est nécessaire de faire une parenthèse concernant les effets généraux du contexte environnemental sur la mémoire. On postule que la trace épisodique d'un souvenir contient l'élément dit focal, c'est-à-dire les informations sur lesquelles l'individu a porté son attention (e.g., les items à apprendre), mais également des éléments contextuels sur lesquels il n'a pas nécessairement porté son attention. Un certain nombre d'études ont porté sur l'effet de la congruence entre l'environnement d'apprentissage et l'environnement de récupération sur les performances mnésiques dans des situations où une seule occurrence de l'item survient et où l'item est testé dans un deuxième temps (pour une revue, voir

Smith & Vela, 2001). Une grande majorité de ces études a montré que les performances mnésiques sont plus élevées lorsque le contexte d'encodage et le contexte de récupération sont similaires (e.g., Godden & Baddeley, 1975 ; Smith, 1979). Cet effet est lié à la notion de spécificité de l'encodage développée par Tulving et Thompson (1973), selon laquelle les éléments environnementaux présents lors de la récupération serviraient d'indices pour aider la récupération des traces contenant des éléments similaires. Nous avons déjà vu cette notion dans la théorie de la variabilité de l'encodage développée par Glenberg (1979). Cependant, dans certains cas la supériorité de la congruence du contexte n'a pas été observée (Fernandez & Glenberg, 1985).

Définition du « contexte »

Ces effets font apparaître la nécessité de définir précisément ce que l'on entend par contexte expérimental. Nous avons vu dans le chapitre 2 que selon Bower (1972), le contexte est constitué de nombreux éléments de nature différente. On peut distinguer les éléments extérieurs composant l'environnement physique (i.e., la pièce où a lieu l'expérience, la luminosité, les items adjacents à l'item cible dans la liste, etc...) mais également les éléments psychologiques, internes au sujet (i.e., les sensations physiologiques comme la faim, l'état émotionnel, le questionnement sur les buts de l'expérience, etc...). Glenberg (1979) considérait que les composants dits contextuels sont ceux qui contiennent les éléments de l'environnement physique, le temps, et les états affectifs et cognitifs du sujet. Mais il ne faisait pas particulièrement la distinction entre les différents types. Concernant les éléments psychologiques, ou plus généralement l'état interne du sujet, Eich (1995) proposait que le fait d'observer de meilleures performances mnésiques lorsque le lieu de récupération est le même que le lieu de l'encodage pourrait en fait être médié, non pas par le lieu en tant que tel, mais par l'humeur (*mood*) du participant. Ce ne serait pas les éléments environnementaux (e.g., les caractéristiques physiques de la pièce) qui seraient encodés dans les traces et qui serviraient d'indices de récupération, mais les états d'esprit associés aux lieux. Les expériences réalisées par Eich suggèrent que l'effet de concordance entre lieu d'encodage et lieu de récupération n'est pas dû à la plus ou moins grande similarité entre les lieux mais à la similarité entre les émotions (*mood*) évoquées par les lieux. En effet, lorsque les lieux étaient différents lors de l'encodage et de la récupération, le rappel était plus élevé lorsque ces deux lieux évoquaient des émotions similaires. À l'inverse, quand les lieux étaient différents, le rappel était facilité si l'état émotionnel était similaire (la similarité subjective entre les états lors de la première session et lors de la deuxième session était estimée par le biais d'un questionnaire). Smith (1995) considérait plutôt que c'est la situation mentale (*mental set*) dans son ensemble, et non pas simplement l'humeur, qui serait responsable des effets de l'environnement sur la mémoire.

Cette très succincte revue de la littérature suggère qu'il n'est pas évident que tous les types d'éléments contextuels (environnementaux, internes, etc...) jouent le même rôle fonctionnel dans les processus mnésiques. Nous allons donc nous pencher sur cette question dans le cas spécifique où des répétitions de stimuli entrent en jeu.

Effet du contexte en situation de répétition

Cette question a justement fait l'objet d'une étude séminale réalisées par Smith, Glenberg, et Bjork (1978). Dans l'Expérience 1, les auteurs faisaient apprendre une même liste de mots lors de deux épisodes, P1 et P2, séparés par trois heures d'intervalle puis testée lors d'une tâche de rappel libre trois heures plus tard. Dans ces expériences, l'intervalle P1-P2 (IIR) était maintenu constant. Deux contextes environnementaux distincts pouvaient

être instanciés pour les épisodes d'étude, C1 ou C2. Ces deux contextes différaient de diverses façons ; l'immeuble et la pièce où se déroulait l'apprentissage, la luminosité, le mode sensoriel de présentation, et l'habillement de l'expérimentateur étaient différents dans C1 et C2. Deux conditions étaient comparées : une condition de *constance* dans laquelle le même contexte était instancié à P1 et P2 (C1-C1 ou C2-C2) et une condition de *changement* dans laquelle le contexte variait (C1-C2 ou C2-C1). Le contexte instancié lors du rappel libre était neutre (C3). Les performances de rappel ont montré que la condition la plus favorable au rappel était la condition de changement, c'est-à-dire que la variation du contexte lors des épisodes d'encodage favorisait la rétention des items. Glenberg (1979, Expérience 3) a montré le même type de résultats.

Trente ans plus tard, Isarida et Isarida (2010) se sont interrogés sur l'effet de l'amplitude de la modification des contextes entre les occurrences. Dans leurs quatre expériences, une liste de mots était présentée à deux reprises séparées par un IIR donné. Ils faisaient varier ou gardaient constant le contexte environnemental, tout comme Smith et al. (1978) et, de plus, comparaient deux conditions :

- la condition de *contexte de lieu simple*, dans laquelle seul le lieu différait entre P1 et P2 ;
- la condition de *contexte de lieu complexe*, dans laquelle davantage de paramètres étaient modifiés entre P1 et P2 : le lieu, l'environnement social (le participant étant seul ou en groupe), et la tâche d'encodage.

Les auteurs ont montré que dans la condition de *contexte de lieu simple*, la condition de *changement* était plus avantageuse que la condition de *constance* pour la rétention finale, ce qui constituait une réplication des effets observés par Smith et al. (1978) et Glenberg (1978, Expérience 3). Cependant, cet effet était inversé en condition *contexte de lieu complexe* ; ainsi, lorsque la manipulation portait sur un plus grand nombre de caractéristiques de la situation expérimentale, il était plus avantageux de garder constant l'ensemble de ces paramètres entre P1 et P2 plutôt que de les faire varier.

De façon intéressante, Isarida et al. (2010) interprétaient ces résultats en termes de variabilité de l'encodage et de récupération en phase d'étude. Selon les auteurs, la récupération en phase d'étude est rendue plus difficile lorsqu'un grand nombre de paramètres sont modifiés, et en particulier la tâche d'encodage, ce qui est délétère pour la performance finale. En condition de variation du *contexte de lieu simple*, la récupération en phase d'étude est censée ne pas être différente dans les conditions de changement et de constance. Il est alors avantageux de faire jouer la variabilité de l'encodage, et c'est pourquoi on observait l'avantage de la condition de variation. En condition de variation du *contexte de lieu complexe*, la récupération en phase d'étude est rendue plus difficile entre P1 et P2. Ainsi il vaut mieux favoriser cette récupération qui est primordiale à la performance finale. C'est pourquoi on observe l'avantage de la condition de constance. Ainsi, de nouveau, réapparaît l'idée qu'une variabilité de l'encodage maximum est souhaitable, mais uniquement dans la mesure où la récupération en phase d'étude est réalisée, ce qui rejoint les notions de modèles mixtes développés dans le chapitre 2 (e.g., Young & Bellezza, 1982).

Pour résumer, si la théorie de la variabilité de l'encodage est bien la cause de l'effet d'espacement, alors dans le cas d'un apprentissage sur plusieurs sessions, ce ne peut pas raisonnablement être l'environnement physique de passation des sessions qui est l'élément qui fluctue ; en effet, on fait en sorte que les contextes environnementaux soient justement les plus similaires possibles de session en session. On peut donc penser que c'est plutôt la fluctuation du contexte interne, autrement dit l'état psychologique du sujet, qui devrait

avoir une influence ; par exemple, le participant peut avoir de plus en plus faim au fur et à mesure du déroulement de l'expérience, il peut être de moins en moins anxieux, avoir un cheminement de pensée concernant les buts de l'expérimentateur, etc... Il n'a pas été réalisé à notre connaissance d'étude dans laquelle sont comparées une condition dans laquelle est induit un état mental différent lors de P1 et P2 avec une condition dans laquelle est induit un état mental similaire lors de P1 et P2.

De plus, il est également possible de considérer que les expériences qui manipulent l'IIR manipulent en réalité deux facteurs confondus : la durée de l'intervalle, mesurée en valeur absolue d'une part, et la fluctuation aléatoire de l'environnement et de l'état psychologique du sujet d'autre part. Nous proposons donc dans cette expérience de séparer ces deux facteurs habituellement confondus. Une piste pourrait être d'induire artificiellement ces contextes internes lors de P1 et P2. Cependant, dans l'Expérience 4, nous avons choisi de mettre en œuvre une autre procédure, basée sur une induction indirecte, dont le principe est décrit dans la section suivante.

5. 1. 2. Principe de l'expérience

Un protocole original d'une durée d'environ 1h45 a été élaboré, dans lequel les stimuli à apprendre étaient répétés après un IIR de 40 minutes, puis la rétention de ces stimuli était testée suite à un DR de 30 minutes. L'expérience se composait des cinq étapes classiques des études sur l'effet de pratique distribuée, représentées dans la Figure 25.

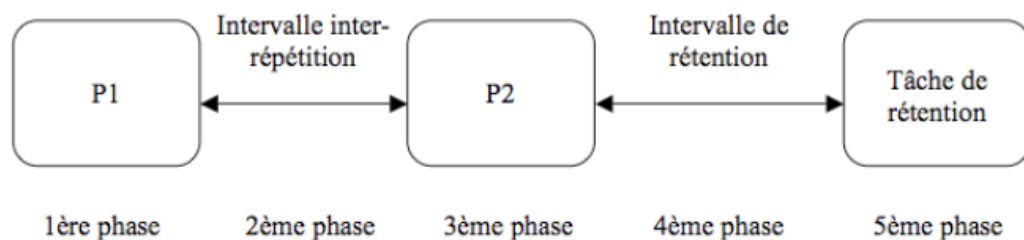


Figure 25 : Schéma des différentes phases de l'Expérience 4.

Deux conditions expérimentales étaient comparées, qui variaient non pas sur la durée de l'IIR, mais sur les événements se produisant pendant cet intervalle. De cette façon nous cherchions à simuler, dans une première condition, une importante fluctuation de l'état psychologique du sujet (correspondant selon nous à une longue période de temps), tandis que dans une seconde condition nous simulons une faible fluctuation de l'état psychologique (correspondant selon nous à une courte période de temps). Ces deux conditions expérimentales se distinguaient par les activités réalisées pendant l'IIR :

- Dans le groupe Fluctuation Forte, les participants réalisaient des activités de nature diverse, mettant en jeu plusieurs modalités sensorielles et motrices.
- Dans le groupe Fluctuation Faible, les participants réalisaient une activité monotone.

Dans les deux groupes, les activités étaient réalisées dans un environnement autre que celui des sessions expérimentales proprement dites. Lors de chaque session expérimentale, les deux conditions étaient testées, et ce avec deux échantillons homogènes de participants. En effet, il était important de rendre égaux par ailleurs tous les autres facteurs tels que le moment de la journée, la luminosité, le déroulement des apprentissages, la durée des intervalles, etc..., pour pouvoir comparer les performances des deux groupes.

Le matériel consistait en un ensemble de paires associant un mot et un pseudomot, comme dans les Expériences 1 à 3. Le test final était une tâche de rappel indicé dans laquelle le pseudomot était présenté et le participant devait donner le mot associé.

5. 1. 3. Hypothèses

Les sujets du groupe Fluctuation Forte, qui réalisaient pendant l'IIR un ensemble d'activités simulant un long passage du temps, devraient bénéficier d'un différentiel de contexte psychologique entre P1 et P2 plus important que les sujets du groupe Fluctuation Faible, dont les activités pendant l'intervalle simulaient un plus court passage de temps. L'encodage devrait donc être plus varié entre P1 et P2 dans la condition Fluctuation Forte que dans la condition Fluctuation Faible. La théorie de la variabilité de l'encodage prédit que le groupe Fluctuation Forte devrait donc présenter lors du test final de meilleures performances de rappel.

5. 1. 4. Méthode

Avant-propos

Six sessions expérimentales de 1h45 chacune ont été organisées. Les passations se déroulaient par groupe. Trois salles expérimentales étaient utilisées pour chaque session. Trois expérimentateurs animaient simultanément les sessions : Carolina Saletti, étudiante en Master 1 de Psychologie Cognitive, Jonathan Fayard, étudiant en Master 1 de Sciences Cognitives, et l'auteur du présent travail. C.S. et E.G. animaient la phase 2 pour le groupe « Fluctuation importante », tandis que J.F. animait la phase 2 pour le groupe « Fluctuation faible ». A tout de rôle, J.F. et E.G. animaient les phases 1, 3, 4, et 5 qui se déroulaient avec l'intégralité des participants dans la première salle expérimentale, un unique expérimentateur animant l'intégralité d'une session.

Comme dans les Expériences 1 à 3, des consignes mnémotechniques étaient suggérées aux participants afin d'essayer d'homogénéiser l'ensemble des scores.

Participants

Soixante-et-un étudiant(e)s, dont 13 hommes et 48 femmes, âgés de 18 à 26 ans (âge moyen 20 ans) ont participé à cette étude. Tous étaient de langue maternelle française. Aucun n'avait participé auparavant à une étude sur la même problématique ni sur le même matériel. Leur vue était jugée satisfaisante à excellente (à l'état normal ou corrigé). Chaque participant bénéficiait, à l'issue de sa participation à l'expérience, à une session de groupe dans laquelle les résultats leur étaient présentés ainsi qu'une « mini-formation » pratique sur « comment optimiser sa mémorisation dans le cadre de ses études », en guise de dédommagement non financier pour leur participation. De plus, les 25 étudiants de 3ème année de Licence de Psychologie bénéficiaient également d'un demi-point à leur note pour une discipline de leur cursus. Quatre participants additionnels ont également passé l'expérience mais leur âge les excluait de l'échantillon et leurs résultats n'ont pas été pris en compte. Les participants ont été recrutés par voie de petites annonces papier et par annonces faites lors de cours pour les étudiants de Psychologie. Des échanges d'emails ont eu lieu avec un grand nombre d'étudiants pour leur expliquer le principe de l'expérience et fixer les dates et horaires de passation.

Pour chaque session expérimentale, un ensemble de sept à treize sujets réalisait l'expérience. Chaque participant avait échangé un certain nombre de courriels avec E.G.

pour convenir d'un créneau horaire et fournir certaines informations telles que l'âge, le sexe, le niveau d'études actuel, la langue maternelle, les coordonnées téléphoniques. Ces informations permettaient de constituer, avant de commencer chaque session, les deux groupes de participants, manuellement, de façon à obtenir des groupes les plus similaires entre eux en termes d'âge, de sexe, et d'études suivies. L'attribution d'un groupe donné à une condition expérimentale ou à l'autre était aléatoire.

De plus, suite au dépouillement des questionnaires finaux, nous avons exclu de l'échantillon les participants ayant annoncé qu'ils avaient révisé volontairement en dehors des sessions d'apprentissage, ce qui nous a amenée à exclure quatre participants, tous issus du groupe Fluctuation Faible.

Ainsi, toutes sessions expérimentales réunies, et après exclusion des quatre participants qui avaient révisé volontairement, les deux groupes de sujets présentaient les caractéristiques présentées dans le Tableau 8.

Tableau 8 : Tableau synthétique des caractéristiques des groupes Fluctuation Faible et Forte, avant le commencement de l'expérience.

Caractéristiques	Groupe Fluctuation Faible	Groupe Fluctuation Forte
Effectif	27	30
Proportion d'hommes	5 / 27 = 18,5 %	7 / 30 = 23,3 %
Age moyen	21,1 ans	20,6 ans
Ecart-type de l'âge	1,78	1,75
Proportion de L3 psychologie	16 / 27 = 59%	19 / 30 = 63 %
Proportion hors psychologie ou sciences cognitives	3 / 27 = 11%	4 / 30 = 13 %

Appareils

Trois salles expérimentales appelées A, B, et C, étaient utilisées pendant une session donnée.

En salle A, les stimuli étaient présentés par un ordinateur Apple Macintosh de type MacBook Pro doté de 2,7 Go de RAM au moyen du logiciel Superlab Pro (version 4.0, Cedrus). Les informations étaient projetées sur un écran à l'aide d'un vidéoprojecteur.

En salle B, des morceaux musicaux étaient diffusés par le biais d'un lecteur CD relié à un amplificateur.

Matériel

Phases 1, 3, et 5 : tâche mnésique.

Un ensemble de vingt paires pseudomot-mot parmi celles utilisées dans les Expériences 1 et 2 ont été utilisées, réparties en cinq exemplaires des quatre catégories sémantiques suivantes : animal, partie du corps, objet, et fruit. Le Tableau 9 présente ces vingt paires. Par rapport aux listes présentées dans la section 4.1.3, quatre paires ont été supprimées. En effet, suite au prétest (présenté dans la section suivante) la paire considérée comme la plus « facile » de chacune des catégories sémantiques (au vu des résultats de l'ensemble des Expériences 1, 2, et 3) a été supprimée.

Tableau 9 : Les 20 paires pseudomot-mot cibles utilisées dans l'Expérience 4.

Animal		Partie du corps	
CAJESI	VACHE	TAGOBU	MAIN
JAMIDO	SOURIS	FIRECO	BRAS
FASOLU	CHEVAL	LUTISA	PIED
POCUGA	CHIEN	DERUPA	OEIL
RUTOLA	OISEAU	TODEGI	JAMBE
Objet		Fruit	
COBEMI	STYLO	VUMOGI	FRAISE
MECUSI	LAMPE	GEFIPA	RAISIN
PUSIRO	CHAISE	BAFISO	POIRE
VAMELO	TELEVISEUR	SONIJA	POMME
DAPIGU	TABLE	LATOMU	ABRICOT

Des carnets ont été élaborés pour les réponses des participants au test final. Chaque carnet comportait 21 pages. La première page était dédiée à l'inscription du nom du participant. Les pages suivantes étaient numérotées de 1 à 20, chacune comportait une ligne pour inscrire le mot réponse et, en dessous, un questionnaire de degré de confiance avec trois cases à cocher (Pas sûr, Moyennement Sûr, Sûr).

Un formulaire de consentement et un formulaire de renseignements ont également été élaborés. Sur le formulaire de renseignements étaient demandés la date de naissance, l'âge, la langue maternelle, les langues parlées, la qualité de la vue.

Des questionnaires d'évaluation de motivation et de fatigue ont été fournis à chaque participant lors du début des phases 1, 3, et 5. Il s'agissait d'une feuille contenant deux questions, pour lesquelles les participants devaient évaluer, sur une échelle de 1 à 7 :

- leur niveau de motivation, tel que 1 = Pas du tout motivé, et 7 = Extrêmement motivé ;
- leur niveau de fatigue, tel que 1 = Pas du tout fatigué et 7 = Extrêmement fatigué.

Enfin, un questionnaire de débriefing était proposé aux participants après l'expérience. Ce questionnaire permettait d'interroger le sujet sur les stratégies mnémoniques utilisées, s'il avait révisé volontairement les items pendant les intervalles, s'il (ou elle) s'était ennuyé pendant l'expérience, s'il pensait que les activités autres que l'apprentissage avaient influencé son apprentissage. Le questionnaire est présenté dans l'Annexe n°10.

Phases 2 : intervalle inter-répétition.

Dans le groupe Fluctuation Faible, chaque participant recevait une feuille comportant 4 grilles de sudoku de niveau enfant, débutant, moyen et difficile, ainsi que leurs solutions et un crayon à papier.

Dans le groupe Fluctuation Forte, chaque participant recevait une feuille présentant des points alignés, une feuille blanche pour le dessin et le pliage, un puzzle de 20 pièces sans motif, une feuille comportant 6 opérations de calculs à 3, 4 ou 5 chiffres, des feuilles de papier avec des noms de métiers inscrits, et un crayon à papier.

Phase 4 : délai de rétention.

Une vidéo de 20 minutes a été utilisée, composée de plusieurs séquences de 3 à 5 minutes présentant des paysages variés sur fond sonore. Aucun être humain, animal ou objet manufacturé n'apparaissait dans la vidéo. Des captures d'écran de cette vidéo sont présentées dans l'Annexe n°11.

Procédure

Prétest.

Six participants ont passé l'expérience dans les conditions réelles (décrites dans la section protocole principal). Il s'agissait d'une part de permettre aux trois expérimentateurs de s'entraîner à la passation et, d'autre part, d'ajuster le matériel expérimental. Dans ce prétest, 24 items cibles étaient présentés. Le temps de présentation de chaque paire à P1 et P2 était de 15 sec, et le temps de présentation de chaque pseudomot pendant la tâche de rappel indicé était de 10 s.

Les nombre de réponses correctes s'étendaient de 1 à 11 sur 24 avec une moyenne de 4,67. Ces scores étant particulièrement bas, et les participants ayant trouvé l'expérience trop difficile, nous avons réalisé les modifications suivantes : 1) le nombre d'items cibles a été réduit à 20, 2) le temps de présentation à P1 et P2 a été allongé à 20 s, 3) le temps de présentation lors du rappel indicé a été allongé à 15 s. Nous avons également amélioré la feuille de sudoku afin d'avoir une plus grande variété de niveaux, et avons modifié la vidéo de la phase 4 pour la rendre moins monotone.

Protocole principal.

Avant de commencer chaque session, deux groupes de même effectif étaient constitués parmi les participants inscrits. Les groupes étaient constitués de façon à être les plus homogènes possibles au regard des paramètres suivants : sexe, âge, études en cours, effectif. L'assignation de chaque groupe à l'une des deux conditions expérimentales était décidée de façon arbitraire.

L'expérience était composée de cinq phases.

Première phase (P1).

Les participants étaient assis à une table, dans une salle de type salle de classe (appelée la salle A). Les informations étaient projetées à l'écran.

Après signature du formulaire de consentement et remplissage de la feuille de renseignement, les consignes de l'expérience étaient présentées. Les participants étaient informés que la session se déroulerait en cinq phases qui seraient expliquées au fur et à mesure de l'expérience. Ils ne devaient pas parler, ni écrire, ni copier sur leur voisin. Concernant la première phase, les participants étaient informés qu'ils devaient apprendre un ensemble de paires associant un pseudomot et un mot, dans le but de pouvoir répondre à un test qui aurait lieu à la fin de la session. Dans ce test, ils devraient donner le mot en réponse au pseudomot qui apparaîtrait seul à l'écran. Le test aurait lieu dans ce sens-là uniquement. Un exemple de paire, non testée ultérieurement, était présenté (i.e., LETOBARI - NUAGE). On recommandait aux participants d'utiliser des stratégies mnémotechniques consistant à trouver des éléments dans le pseudomot pouvant être mis en relation avec le mot (e.g., « BAR » contenu dans le pseudomot est une unité de mesure de pression atmosphérique, qui peut être facilement relié au nuage). Il était dit aux participants qu'ils n'auraient pas à rapporter les moyens mnémotechniques employés.

Les participants étaient informés qu'ils verraient une vingtaine de paires pendant 20 secondes chacune. Ils estimaient ensuite leur niveau de fatigue et de motivation puis les questionnaires étaient ramassés.

Vingt-quatre paires étaient ensuite projetées à l'écran, une à une, pendant une durée de 20 s chacune. Le pseudomot apparaissait à gauche et le mot à droite (police Arial

Narrow, caractère gras, taille 55, hauteur des caractères apparents : 5 cm environ). Les deux premières et les deux dernières positions sérielles étaient constituées de paires distractrices, appartenant aux mêmes catégories sémantiques que les paires cibles, dans le but de limiter les effets de primauté et de récence. L'ordre de présentation était pseudo-aléatoire et suivait les règles suivantes :

- Les 20 paires cibles étaient subdivisées en quatre blocs de cinq paires, que nous appellerons successivement A, B, C, et D. (Lors de P2, la séquence des blocs devenait C, A, D, B, afin d'homogénéiser les effets de position sérielle).
- Une même catégorie sémantique n'apparaissait jamais deux fois de suite pendant la séquence. Nous avons essayé au maximum de limiter les associations évidentes possibles entre deux paires successives, par exemple CHEVAL suivi de ŒIL (car les chevaux ont des yeux).

Cette première phase durait environ 8 minutes. Les participants ne devaient pas parler ni écrire les mots. Pendant la projection des paires, les trois expérimentateurs étaient présents dans la salle, à l'écart du lieu de projection.

Il était ensuite demandé aux participants d'essayer de ne pas réviser volontairement les paires de vocabulaire en dehors des périodes d'apprentissage.

Deuxième phase : intervalle inter-répétition.

Les membres des deux groupes expérimentaux étaient désignés, puis participants et expérimentateurs quittaient la salle A pour rejoindre la salle B ou C selon le groupe. Ce déplacement durait environ une minute.

En salle C, les membres du groupe « Fluctuation Faible » étaient assis autour d'un bureau central et devaient remplir des grilles de sudoku pendant une durée d'environ 40 minutes.

En salle B, les membres du groupe « Fluctuation Importante », devaient réaliser six activités successives, chaque activité étant réalisée alternativement debout et assis pendant une durée d'environ 3 minutes chacune :

- Pas de danse (debout),
- Jeu des points (assis),
- Puzzle (debout),
- Calculs (assis),
- Sortie en file indienne avec imitation de gestes (debout),
- Dessiner sur une musique (assis),
- Jeu des mimes (debout),
- Pliage d'un bateau en papier (assis).

Le détail de ces activités est précisé dans l'Annexe n°11.

Lorsque les activités du groupe « Fluctuation Importante » étaient terminées, le groupe sortait de nouveau, l'expérimentatrice E.G. prévenait le groupe « Fluctuation Faible » qui rejoignait le premier groupe, et tous les participants retournaient en salle A.

Troisième phase (P2).

Les participants étaient de nouveau assis en salle A à la même place que lors de la première phase. Ils estimaient de nouveau leur niveau de motivation et de fatigue, puis les questionnaires étaient ramassés.

Les participants étaient ensuite informés qu'ils allaient revoir les mêmes paires que précédemment et qu'ils devaient en profiter pour perfectionner leur apprentissage. Les 24 paires pseudomot-mot étaient présentées de nouveau pendant 20 s chacune. L'ordre de présentation était différent de l'ordre précédent : 1) la succession des blocs était de type CADB ; 2) tout comme lors de P1, deux paires successives n'étaient pas de la même catégorie sémantique ; et 3) une succession de deux paires observée à P1 n'était pas répliquée lors de P2.

Cette phase durait également environ 8 minutes et se déroulait de la même façon que la première phase. On rappelait aux participants qu'ils ne devaient pas chercher à réviser volontairement les paires en dehors des périodes d'apprentissage.

Quatrième phase : délai de rétention.

Toujours dans la salle A, on informait les participants qu'une vidéo leur serait projetée pendant un certain temps et qu'ils devaient regarder les images sans rien faire de particulier. Ils pouvaient en profiter pour se relaxer. La vidéo durait 20 minutes.

Cinquième phase : tâche de rappel indicé.

Les participants estimaient ensuite une dernière fois leur niveau de motivation et de fatigue. Les questionnaires étaient ramassés et les carnets de réponse étaient distribués. Les consignes du test étaient expliquées. Ensuite, les 20 pseudomots cibles étaient présentés seuls à l'écran, un par un, pendant 15 secondes chacun. L'ordre d'apparition des pseudomots était pseudo-aléatoire, suivait les règles déjà mentionnées, et différait de l'ordre utilisé dans les conditions précédentes. Les participants devaient écrire le mot français correspondant puis juger de la certitude de leur réponse (Pas Sûr, Moyennement Sûr, Sûr). Un numéro apparaissait également avec le pseudomot, pour que les participants inscrivent leur réponse sur la bonne page du carnet. Entre deux pseudomots successifs, un signal sonore retentissait pendant 2 s, indiquant qu'il fallait tourner la page. Il était interdit de revenir en arrière une fois la page tournée.

Débriefing.

Les participants rendaient leurs carnets de réponse puis répondaient au formulaire de débriefing. Ils étaient ensuite libres de quitter la salle.

5. 1. 5. Résultats

Les données utilisées pour les analyses statistiques n'ont pas été corrigées.

Rappel indicé

Le nombre de bonnes réponses par participant à la tâche de rappel indicé s'échelonnait, pour le groupe Fluctuation Faible, de 1 à 20 sur 20, c'est-à-dire de 5 % à 100 % de bonnes réponses, et pour le groupe Fluctuation Forte de 4 à 19 sur 20, c'est-à-dire de 20 % à 95 %. La moyenne du nombre de bonnes réponses s'élevait à 12,41 bonnes réponses sur 20 pour le groupe fluctuation faible (62 %, écart-type 5,60), et à 13,57 pour le groupe fluctuation forte (68%, écart-type 4,35). Une participante a donné 100 % de bonnes réponses.

Un test t de Student a été réalisé pour comparer les moyennes des deux groupes, et n'a pas révélé de différence significative, $t(48.96) = -0,87$, $p < .40$.

Le degré de certitude associé aux réponses correctes données par les participants, qui comportait trois modalités (Sûr, Moyennement sûr, Pas sûr), a été également analysé. Le nombre de bonnes réponses associées à un niveau 'Sûr' s'élevait à 10,70 (écart-type

5,64) pour le groupe fluctuation faible, et à 12,07 (écart-type 4,19) pour le groupe fluctuation importante. Cette différence n'était pas significative, $t(47,74) = -1.03$, $p = .30$.

Motivation et fatigue

Nous souhaitions nous assurer que les manipulations expérimentales n'influençaient ni le niveau de motivation des participants, ni leur niveau de fatigue. La Figure 26 présente les scores moyens de motivation et de fatigue en fonction des périodes et selon les groupes.

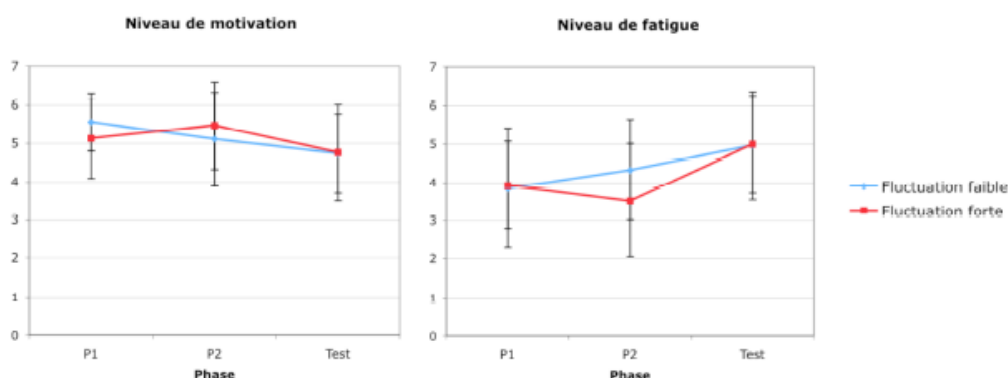


Figure 26 : Niveaux estimés de motivation (figure de gauche) et de fatigue (figure de droite) en fonction des phases de l'expérience (P1, P2, et test) et des groupes (Fluctuation faible ou forte).

Tout d'abord, les niveaux estimés de motivation et de fatigue ont été comparés entre les deux groupes expérimentaux au début de l'expérience. Le niveau moyen de motivation estimé (sur une échelle de 1 à 7) était de 5,56 (écart-type 0,75) dans le groupe Fluctuation Faible, et de 5,13 (écart-type 1,04) pour le groupe Fluctuation Forte. Cette différence s'est révélée tendanciellement significative ($t(52.60) = 1.77$, $p = .083$) ce qui n'était pas du tout attendu étant donné qu'à ce moment-là de l'expérience les participants ne savaient rien du groupe auquel ils avaient été rattachés. En ce qui concerne l'estimation de la fatigue, les niveaux moyens estimés (sur une échelle de 1 à 7) étaient de 3,85 (écart-type 1,56) dans le groupe Fluctuation Faible, et de 3,93 (écart-type 1,14) pour le groupe Fluctuation Forte. Cette différence entre les groupes n'était pas significative, $t(47.26) = 0.23$, $p < .85$. Les scores ne différaient pas non plus entre les groupes au moment du test ($t(54.53) = 0.08$, $p = .93$ et $t(52.63) = 0.10$, $p = .92$, respectivement pour la motivation et la fatigue).

Dans un deuxième temps, l'évolution des niveaux de motivation et de fatigue entre les phases P1 et P2 a été étudiée en fonction des groupes. Les niveaux de motivation étaient respectivement de 5,11 (écart-type 1,22) et de 5,47 (écart-type 1,14) pour les groupes Fluctuation Faible et Fluctuation Forte, tandis que les niveaux de fatigue étaient de 4,33 (écart-type 1,30) et de 3,53 (écart-type 1,48) respectivement. Ainsi, les scores de motivation ont augmenté chez les participants du groupe Fluctuation Forte et ont diminué chez les participants du groupe Fluctuation Faible. Une analyse de variance a été réalisée sur ces deux estimations individuellement, avec comme facteur intra-sujets la période expérimentale (P1 et P2) et comme facteur inter-sujets le groupe expérimental (Fluctuation Faible ou Forte). Concernant les niveaux de motivation, l'interaction s'est révélée significative ($F(1,55) = 4.30$, $p < .001$), indiquant que les activités réalisées pendant l'IIR n'avaient pas le même effet sur l'évolution du niveau de motivation entre P1 et P2 : les activités réalisées en condition de Fluctuation Forte rendait les participants

davantage motivés tandis que l'activité réalisée en condition de Fluctuation Faible rendait les participants moins motivés. Concernant les niveaux de fatigue, l'interaction s'est également révélée significative ($F(1,55) = 8.44, p < .01$), indiquant que les activités réalisées pendant l'IIR n'avaient pas le même effet sur le niveau de motivation : les activités réalisées en condition de Fluctuation Faible rendait les participants davantage fatigués.

Bien entendu, ces effets n'étaient pas du tout recherchés initialement puisque nous souhaitions que les deux types d'activités soient similaires du point de vue de leur effet sur la motivation et sur la fatigue.

Débriefing

Une synthèse des résultats du questionnaire de débriefing est présentée dans le Tableau 10 en fonction des groupes.

Tableau 10 : Tableau synthétique des caractéristiques des groupes Fluctuation Faible et Forte relativement à l'expérience.

Caractéristiques	Groupe Fluctuation Faible	Groupe Fluctuation Forte
Difficile	39% facile - 7% moyen - 50% difficile	37% facile - 17% moyen - 47% difficile
Utilisation de stratégie	1 non	1 non
Pourcentage d'utilisation d'une stratégie	11% 100 ; 54% 80 ; 19% 60 ; 8% 40 ; 8% 20	23% 100 ; 33% 80 ; 33 % 60 ; 0% 40 ; 10% 20
Réminiscences	48 % oui ; 52 % non	57 % oui ; 43 % non
Effet de l'activité entre les répétitions sur apprentissage	non 37% - distraction 50% - fatigant 12% - démotivant 4%	non 26% - distraction 37% - fatigant 7% - motivant 7% - démotivant 3%
Ennui	non 7% - vidéo 81% P1 7% - sudoku 7%	non 10% - vidéo 83% P1 3% - P2 7%
Appréciation	beaucoup 60% ; un peu 37% ; pas trop 4%	beaucoup 72% ; un peu 28%

La comparaison a posteriori des paramètres entre les deux groupes ne révèle pas de différence flagrante entre les deux groupes. Deux observations cependant pourraient avoir un certain intérêt : 1) les sujets du groupe Fluctuation Forte auraient eu davantage de réminiscences des items en dehors des phases d'apprentissage (57% versus 48%) ; 2) les sujets du groupe Fluctuation Forte auraient davantage été motivés par les activités réalisées pendant l'IIR (7% versus 0%, ce qui est cohérent avec les analyses des scores de motivation) et auraient davantage apprécié de participer à l'expérience (72% versus 60 % pour la réponse 'beaucoup'). Nous ne développerons pas plus avant la discussion de ces résultats.

5. 1. 6. Discussion

Résumé des résultats

L'expérience avait pour objectif de tester la théorie de la variabilité de l'encodage en comparant une condition de fluctuation faible dans laquelle était simulé -par le biais d'un plus ou moins grand nombre d'activités variées- un court IIR, avec une condition de fluctuation forte dans laquelle était simulé un long IIR.

Bien que le score moyen de rappel libre en condition de fluctuation importante (13,57 sur 20) ait été plus élevé de 9,3 % que celui en condition de fluctuation faible (12,41), la différence ne s'est pas révélée significative. Ainsi, nous n'avons pas validé l'hypothèse issue de la théorie de la variabilité de l'encodage. De plus, sont apparus des phénomènes non désirés liés à la motivation des sujets : les sujets de la condition Fluctuation Faible présentaient une diminution de leur motivation entre P1 et P2 tandis que les sujets de la condition Fluctuation Forte présentaient une diminution de leur motivation entre P1 et P2. Or, nous souhaitions que tout soit égal par ailleurs, y compris les niveaux de motivation et de fatigue, qui peuvent justement influencer l'efficacité de l'encodage.

Interprétations

Le fait de n'avoir pas montré que la variabilité des activités réalisées pendant l'IIR influençait significativement le rappel indicé final peut s'interpréter de différentes façons relativement à la théorie de la variabilité de l'encodage.

Selon la première interprétation possible, l'hypothèse de la variabilité de l'encodage est correcte et l'expérience est adaptée au test de cette hypothèse, mais la trop grande variabilité des scores a eu pour effet d'empêcher d'atteindre la significativité de la différence. Il faudrait donc augmenter de façon drastique les effectifs des deux groupes de participants pour avoir une chance d'obtenir un effet. Une possibilité serait également, avant de commencer l'expérience proprement dite, de faire passer aux participants une tâche d'apprentissage et de rappel sur un matériel similaire à celui de l'expérience ; à partir des résultats à cette tâche, ne participeraient à l'expérience que ceux ayant des performances moyennes, dans une fourchette déterminée au préalable. On pourrait également envisager un protocole dans lequel le facteur fluctuation serait manipulé en intra-sujets, mais cela amènerait d'autres biais à contrôler également (effet de séquence par exemple).

Selon la seconde interprétation possible, l'hypothèse de la variabilité de l'encodage est correcte, mais l'expérience n'est pas adaptée pour la mettre en évidence. Plusieurs éléments expérimentaux pourraient être ajustés pour pouvoir mieux tester l'hypothèse. Par exemple, une tâche de rappel libre pourrait être mieux à même de mettre en évidence des effets de contexte (en effet, en reconnaissance, les indices extérieurs jouent un rôle plus faible qu'en rappel indicé où ils jouent un rôle essentiel. De même, la tâche de *sudoku* n'est peut-être pas si monotone que nous le pensions initialement, et l'effet induit par la tâche de *sudoku* est probablement différent selon les personnes en fonction de leur habitude d'en réaliser, de leur attirance pour les jeux de chiffres, etc... Il se peut également que les deux groupes, du fait des activités réalisées, aient été confrontés de façon différente à des interférences avec l'apprentissage des paires. En tout état de cause, l'effet que nous cherchions à mettre en évidence est très subtil et peut être facilement noyé par d'autres facteurs entrant en jeu dans l'expérience.

Selon la troisième interprétation possible, la différence (non significative) entre les deux conditions serait due non pas à la fluctuation de l'état psychologique des sujets, qui serait plus importante en condition de Fluctuation Forte, mais à un niveau de motivation plus élevé induit par les tâches -assez ludiques- réalisées dans cette condition par rapport à la condition de Fluctuation Faible. En effet, on imagine aisément qu'un participant plus motivé sera plus efficace dans l'encodage lors de P2, ce qui sera la cause de ses performances finales plus élevées, et non pas la fluctuation de contexte. De plus, certains participants ayant appartenu au groupe Fluctuation Faible nous ont fait part d'une certaine « frustration » due à la longue durée de la tâche de *sudoku*. Tout un ensemble de facteurs peuvent également différer entre les deux conditions : niveau d'arousal, niveau de fatigue, état psychologique tel que

l'anxiété, l'ennui. On est donc confronté à un problème de confusion de variable avec de nombreux autres facteurs, qui ferait en sorte que le participant se trouve simplement dans de meilleures conditions pour apprendre à moment de P2 dans la condition de Fluctuation Forte.

Enfin, selon la quatrième interprétation possible, l'hypothèse de la variabilité de l'encodage est fausse et ne joue pas de rôle dans les effets de pratique distribuée. Ainsi, l'intervalle de temps séparant P1 et P2 serait le facteur responsable des effets de pratique distribuée, la fluctuation du contexte étant un phénomène annexe et qui pourrait dans certains cas avoir une influence. Selon cette interprétation, les activités réalisées par le sujet pendant l'IIR n'influeraient pas sur la variabilité de l'encodage, ou alors la variabilité de l'encodage induite par ces activités différentielles n'aurait pas d'influence sur la rétention. Si seul le temps est le facteur responsable de l'effet de pratique distribuée, cela va dans le sens des hypothèses de type neurophysiologique qui postulent que les phénomènes de consolidation synaptique et systémique (voir la section 3.3) à l'œuvre dans l'effet de pratique distribuée nécessitent une période incompressible pour se réaliser. Ces effets pourraient éventuellement être modulés par d'autres phénomènes comme la variabilité de l'encodage.

Nous ne pouvons pas départager ces quatre interprétations sur la base des résultats de l'Expérience 4.

Il est nécessaire également d'ajouter ici que, dans la mesure où 1) l'IIR était constant, 2) les caractéristiques de présentation des stimuli étaient identiques entre P1 et P2, 3) que les temps de présentation étaient relativement longs, 4) que les participants étaient informés lors de P2 que les items de P1 seraient répétés, et 5) au vu des scores plafonds de reconnaissance 24h plus tard observés dans l'Expérience 2, que les phénomènes de récupération en phase d'étude de P1 lors de P2 devraient se trouver autour des 100% sans différence particulière entre les deux groupes de participants. Si cette intuition est correcte, alors nous avons manipulé, comme nous le souhaitons, uniquement la variabilité de l'encodage sans interférer avec la récupération en phase d'étude. Toutefois nous n'avons pas de données tangibles permettant d'étayer notre intuition.

Considérations sur la théorie de la variabilité de l'encodage

Nos réflexions sur la théorie de la variabilité de l'encodage nous ont amenée à nous interroger en profondeur sur cette théorie, nous amenant à constater qu'il peut exister différentes formulations des hypothèses induites par la notion de variabilité de l'encodage. Nous dénombrons trois d'entre elles, qui se distinguent par les phénomènes en jeu lors de la répétition. Ces trois formulations ne font pas l'objet, à notre connaissance, de distinction particulière dans la littérature.

Premièrement, on peut considérer qu'entre les deux groupes expérimentaux de la présente expérience, un nombre différent d'éléments ou de composants sont actifs en mémoire lors de P2, et composent le contexte mental (*mental set*, Smith, 1995). Dans le groupe *fluctuation forte*, les participants viennent d'être confrontés à différents lieux, diverses interactions, plusieurs types d'informations (métiers, chiffres), des éléments musicaux, tandis que dans le groupe *fluctuation faible*, les participants viennent d'être confrontés à un plus petit nombre de lieux, de type d'information, etc. Or, les éléments de ce contexte mental plus ou moins latent pourraient être encodés dans les traces mnésiques associées aux éléments focaux, rendant les traces plus riches dans le groupe fluctuation forte que dans le groupe fluctuation faible.

Deuxièmement, le fait de réaliser un plus ou moins grand nombre d'activités variées entre P1 et P2 a pour conséquence de faire en sorte que l'état psychologique lors de P2 est de nature plus ou moins différente de celui de P1. On peut citer par exemple le niveau d'anxiété, la sensation de bien-être, le niveau d'arousal, la valence émotionnelle, la faim, etc. Ces éléments qui constituent le contexte mental pourraient être encodés dans les traces mnésiques associées aux éléments focaux, rendant les traces plus riches dans le groupe fluctuation forte que dans le groupe fluctuation faible.

Ces deux premières interprétations ne sont pas mutuellement exclusives, et sont davantage reflétées dans les articles récents sur l'effet de pratique distribuée (e.g., Cepeda et al., 2006).

Troisièmement, la façon d'encoder les informations est influencée par le contexte mental instancié au moment de l'encodage. C'est par conséquent la façon d'encoder l'élément cible, et non pas les composants associés à l'élément cible, qui va varier entre P1 et P2 d'une façon plus ou moins importante. Par exemple, si la personne est d'humeur joyeuse elle va formuler des stratégies mnémotechniques évoquant davantage des caractéristiques positives, qu'une personne d'humeur triste. Ainsi, selon l'hypothèse de la variabilité de l'encodage, en condition de fluctuation faible, l'encodage va être similaire à P1 puisque l'état mental n'aura pas beaucoup fluctué (encodage de type A lors de P1 et de type A lors de P2), tandis qu'en condition de fluctuation importante, l'encodage va être différent de P1 (encodage de type A lors de P1 et de type B lors de P2). Cette conception est proche des conceptions développées par Bower (1972).

Ces trois conceptions s'appliquent également à l'effet de la pratique distribuée, et nécessitent d'être testées expérimentalement. Selon que telle ou telle conception soit correcte, des recommandations différentes pourront être faites pour favoriser l'apprentissage et la rétention finale.

5. 1. 7. Conclusion

La théorie de la variabilité de l'encodage est une des théories les plus populaires pour expliquer l'effet de pratique distribuée. Pourtant, dans ce chapitre ainsi que dans le chapitre 2 nous avons vu qu'une condition d'encodage varié lors des diverses occurrences de l'item n'étaient pas toujours plus bénéfique à la rétention finale qu'une condition d'encodage constante. De plus, lorsque les répétitions ont lieu lors de jours différents, la continuité de la fluctuation du contexte est raisonnablement rompue, et la proportionnalité entre la variation contextuelle et la durée de l'IIR semble sujette à discussion. Enfin, dans les expériences sur l'effet de la pratique distribuée, il y a une confusion de variable entre la durée de l'IIR et la fluctuation du contexte. Notre expérience n'a pas pu mettre en évidence d'effet significatif de l'intensité de la fluctuation contextuelle interne ; il se révèle difficile de manipuler cette variation de contexte interne, étant donné la subtilité des processus en jeu. Il est plus aisé de manipuler le contexte environnemental externe.

De façon intéressante, dans cette section consacrée à la théorie de la variabilité de l'encodage, nous avons rencontré de nouveau, en particulier dans Isarida et Isarida (2010), la notion de récupération en phase d'étude. Comme nous l'avons développé dans le chapitre 2 avec les modèles mixtes, il semble qu'il existe une intrication importante et complexe entre les deux phénomènes, et que l'on ne puisse pas évoquer une théorie (variabilité de l'encodage) sans faire appel à l'autre (récupération en phase d'étude).

Troisième partie : Discussion générale

Chapitre 6 : Discussion générale

Cognitive psychologists must make a greater effort to understand cognition as it occurs in the ordinary environment and in the context of natural purposeful activity. This would not mean an end to laboratory experiments, but a commitment to the study of variables that are ecologically important rather than those that are easily manageable. Neisser (1976)

6. 1. Synthèse des résultats

Les expériences présentées dans le présent travail ont permis de répliquer, suite aux travaux de Tsai (1927), Cull (2000) et Gay (1973), que l'effet de l'agencement temporel des répétitions sur plusieurs jours sur la rétention ultérieure correspond à une réalité, du moins pour le type de matériel et pour les agencements particuliers utilisés ici. Nous avons mis en évidence que l'agencement de type expansif était en général associé à de meilleures performances de rappel indicé que les agencements de type uniforme et contractant. Parfois l'effet de l'agencement n'était pas significatif, comme dans l'Expérience 3 lorsque le délai de rétention (DR) était court ; parfois il n'y avait pas de différence significative entre les agencements expansif et uniforme, comme dans l'Expérience 2. Dans les autres cas, l'agencement expansif était significativement plus efficace que les agencements uniforme et contractant. Ainsi, nos observations vont dans le sens des prédictions intuitives selon lesquelles il est efficace de faire des révisions fréquentes en début d'apprentissage puis que l'on peut ensuite se permettre d'attendre plus longtemps pour réviser (e.g., Lyon, 1914).

Le fait d'avoir observé des résultats significatifs dans ces expériences est un signe, selon nous, de la robustesse des effets observés : malgré tout le « bruit expérimental », l'avantage de l'agencement expansif a pu être montré. En effet, les expériences ont été réalisées dans des conditions relativement écologiques, puisque la nature du paradigme employé faisait en sorte que l'expérience était intégrée dans la « vie de tous les jours » des participants. Entre les sessions expérimentales, les participants étaient confrontés à un nombre très important d'informations, vivaient des événements, des périodes d'effort intellectuel et physique, des périodes de repos et de sommeil, des expériences émotionnelles variées, tout ces événements étant évidemment incontrôlables. Ainsi, bien que les sessions expérimentales elles-mêmes aient bénéficié d'un important contrôle expérimental, de nombreux éléments se produisant dans la vie des sujets a pu influencer et interférer avec les processus d'apprentissage et de rétention tout au long de la période expérimentale. Malgré tous ces effets potentiellement parasites, l'effet s'est révélé significatif ce qui est la preuve, selon nous, de la validité externe du phénomène.

Nos études utilisaient un protocole dans lequel l'agencement temporel était un facteur intra-sujets, et nous avons globalement répliqué les résultats des études de Tsai (1927) et Gay (1973) dans lesquelles le facteur agencement était un facteur inter-sujets. Le fait d'obtenir des résultats similaires en utilisant des protocoles différents renforce également l'idée de la réalité du caractère optimal de l'agencement expansif.

L'apport novateur du présent travail est d'avoir montré que l'effet de l'agencement temporel pouvait être modulé par le DR. En effet, dans l'Expérience 3, bien que les trois agencements aient été associés à des performances similaires peu après la fin de la phase d'apprentissage (i.e., J15), l'avantage de l'agencement expansif ne s'est révélé significatif que progressivement au fur et à mesure de l'allongement du DR (i.e., J19 et J26). Cette observation suggère que l'agencement expansif était associé à un taux d'oubli plus faible que les deux autres agencements. Dans la section suivante (6.2) nous nous attacherons à mettre ces résultats en lien avec les théories existantes relatives à la pratique distribuée.

Il est également important de mentionner que nous avons mis en évidence des effets de l'agencement temporel opposés à ceux observés sur une courte échelle de temps. En effet, lorsque des items sont répétés sur une échelle temporelle de quelques minutes (e.g., Balota et al., 2006 ; Karpicke & Roediger, 2007 ; Logan & Balota, 2008), il a été montré que l'agencement de type expansif est parfois le plus efficace lorsque le DR est court, mais est toujours le moins efficace lorsque le DR augmente. Nous développerons cette question dans la section 6.4.

Dans la section 6.3, nous discuterons de la notion de taux d'oubli, notion qui nous semble mériter d'être approfondie et davantage considérée. Dans la section 6.5, nous discuterons également de la question de la généralisation de nos résultats.

6. 2. Implications pour les théories de l'effet de pratique distribuée

Les deux principales théories en présence visant à expliquer l'effet de la pratique distribuée et donc l'effet de l'agencement temporel sont celles de la variabilité de l'encodage et de la récupération en phase d'étude (voir le chapitre 2). D'autres phénomènes peuvent également être considérés comme ayant un rôle dans les effets observés ; nous les discuterons successivement.

6. 2. 1. Théorie de la variabilité de l'encodage

D'une façon générale, la théorie de la variabilité de l'encodage (e.g., Glenberg, 1979) postule qu'un item doit être associé à des contextes d'encodage les plus variés possible pour bénéficier d'une trace mnésique la plus riche possible, ce qui sera avantageux lors du rappel final (voir aussi le chapitre 5 pour une discussion sur la question du contexte). Cette théorie explique pourquoi des répétitions séparées par un long intervalle inter-répétition (IIR) permettent un meilleur rappel que des répétitions séparées par un court IIR : lors du long IIR, le contexte -à la fois environnemental et interne- a le temps de fluctuer de façon plus ample que lors de l'intervalle court. Cependant, cette explication semble moins évidente lorsqu'il s'agit d'expliquer l'effet de l'agencement temporel. En effet, lorsque des agencements temporels de répétitions sont comparés, l'intervalle moyen entre deux occurrences successives est le même dans tous les agencements. Dans l'agencement uniforme, tous les IIR sont identiques ; dans l'agencement expansif, les premiers IIR sont courts et les derniers sont longs ; dans l'agencement contractant, les premiers IIR sont longs et les derniers sont courts. Ainsi, le fait d'avoir de longs IIR à la fin de l'agencement expansif (ou au début de l'agencement contractant) est *a priori* compensé par le fait qu'on a de courts IIR au début (ou à la fin pour l'agencement contractant). Ainsi, en première appréciation, la variabilité de l'encodage n'est pas un facteur évident pour expliquer les différences entre les agencements.

Par ailleurs, la théorie de la variabilité de l'encodage prédit que la durée du DR influence les effets de pratique distribuée, car les éléments contextuels présents lors du test de rétention jouent un rôle dans la récupération des souvenirs : on a tendance à récupérer les traces contenant des éléments contextuels similaires à ceux présents au moment du test. Suite à un DR court, l'agencement contractant devrait donc être favorisé puisque le contexte instancié est alors proche du contexte ayant été instancié à la fin de la phase d'apprentissage. À l'inverse, suite à un DR long, l'agencement uniforme devrait être favorisé puisque le contexte instancié est alors très différent des contextes instanciés pendant la phase d'apprentissage. La trace constituée des éléments contextuels les plus variés, c'est-à-dire celle issue de l'agencement uniforme, devrait être davantage récupérée que les autres. Or nos résultats montrent que ces prédictions n'ont pas été vérifiées.

Nous émettons de surcroît des doutes sur l'idée que le contexte fluctue progressivement avec le temps sur une longue échelle de temps. Comme discuté dans le chapitre 5, il nous semble pertinent d'affirmer que, sur une période de quelques minutes voire de quelques heures, le contexte fluctue progressivement et de façon continue. Mais prenons l'exemple de deux sessions d'apprentissage avec répétitions des items d'une session à l'autre, telles que l'IIR est dans un cas d'un jour et dans l'autre cas de quatre jours. Les contextes environnementaux (e.g., Smith et al., 1978) ainsi que les contextes internes (i.e., psychologiques ; e.g., Smith, 1995) sont-ils plus différents entre les deux sessions dans le cas de l'IIR de quatre jours que dans le cas de l'IIR d'un jour ? Nous pensons que ce n'est pas le cas, car la « coupure » qui se produit entre les sessions, pendant laquelle les sujets retournent à leurs occupations habituelles, rompt probablement la fluctuation du contexte due au temps. Ainsi, nous pensons qu'il n'y aura pas une plus grande différence de contexte entre deux sessions lorsque l'IIR augmente. Au contraire, il pourrait même exister une sorte de périodicité liée aux jours de la semaine, qui pourrait rendre les contextes davantage similaires entre eux dans certains cas. En effet, lorsqu'un IIR de sept jours sépare deux sessions, alors celles-ci ont alors lieu le même jour de la semaine, ce qui pourrait favoriser au contraire une similarité de contexte. Un étudiant, par exemple, avant d'assister à deux sessions expérimentales séparées par une semaine, a suivi des cours sur les mêmes disciplines, dans les mêmes salles et avec les mêmes enseignants. Or cela ne se produit pas lorsque les sessions sont séparées par 48 h. Pour prendre un cas extrême, si la deuxième session a lieu le samedi ou le dimanche, le contexte de l'environnement universitaire est alors particulier (e.g., pas d'étudiant sur le campus, etc) et diffère alors grandement du contexte universitaire d'un jour de semaine. Weigold (2008) défendait également cet argument.

L'Expérience 4 avait pour objectif de tester l'hypothèse de la variabilité de l'encodage en se focalisant particulièrement sur la variabilité du contexte interne du sujet au moment des occurrences de l'item. Nous avons tenté de simuler des degrés divers de fluctuation du contexte en manipulant la variété des activités réalisées pendant l'IIR. Les résultats étaient mitigés puisque la différence de performance finale, bien qu'allant dans le sens de la théorie de variabilité de l'encodage, n'était pas significativement plus élevée dans le groupe simulant une fluctuation forte que dans le groupe simulant une fluctuation faible. De plus, des effets de motivation étaient probablement en jeu dans la différence observée. Les résultats de l'Expérience 4 sont donc d'un intérêt limité, et ne se positionnent ni en faveur ni en défaveur de la théorie de la variabilité de l'encodage.

La limite principale de la théorie de la variabilité de l'encodage est, selon nous, l'absence de prise en compte de la récupération de la trace existante pour l'item lorsque celui-ci est répété, argument développé par Bellezza et ses collaborateurs (e.g., Young &

Bellezza, 1982). Or, ce problème de récupération en phase d'étude est d'autant plus aigu que les IIR sont longs. Nous avons par exemple observé qu'un certain nombre de paires n'étaient pas reconnues lorsque répétées au bout de quelques jours (Expérience 2). Ainsi, même si la variabilité de l'encodage joue un rôle dans l'explication de l'effet de l'agencement temporel, il semble raisonnable de penser que les processus de récupération en phase d'étude ont des effets autrement plus importants sur les effets observés. Comme nous l'avons vu dans le chapitre 3, des modèles mixtes ont été proposés (e.g., Verkoeijen et al., 2005) et modélisés (e.g., Mozer et al., 2009 ; Raaijmakers, 2003).

D'une façon générale, la théorie de la variabilité de l'encodage présente des aspects relativement flous et indéfinis, et qui ne sont pas explicitement décrits comme des « variantes » de l'hypothèse initiale. Comme nous l'avons discuté dans le chapitre 5, ce que signifie variabilité de l'encodage n'est pas toujours clair : s'agit-il d'encoder un item accompagné des éléments contextuels qui seraient intégrés à la trace, ou s'agit-il du fait que le contexte instancié influence sur la façon d'encoder les stimuli ? La formulation de Bower (1972) rejoindrait plutôt la seconde conception, tandis que celle de Glenberg (1979) serait à rapprocher de la première. Nous trouverions pertinent que cette théorie, qui est si souvent citée pour les effets de pratique distribuée, fasse l'objet d'une formulation explicite actualisée au vu des données empiriques récoltées depuis les années 1970.

Ainsi, en guise de conclusion tout à fait partielle sur la question de la théorie de la variabilité de l'encodage, il nous semble que, effectivement, il est pertinent de penser que les composants du contexte externe et interne sont encodés dans la trace ou, selon l'autre variante, qu'ils influencent la façon d'encoder les stimuli, et qu'ils influencent également les processus de récupération. Cependant nous ne voyons pas clairement comment la théorie de la variabilité de l'encodage pourrait expliquer simplement et à elle seule les résultats obtenus relatifs à l'effet de l'agencement temporel à long terme. Nous pensons que cette théorie néglige les phénomènes d'oubli et de récupération en phase d'étude, qui, selon nous, jouent un rôle plus important dans les effets observés.

6. 2. 2. Théorie de la récupération en phase d'étude

La théorie de la récupération en phase d'étude (e.g., Johnston & Uhl, 1976) postule que les répétitions des occurrences d'un même item, lors de la phase d'apprentissage, déclenchent la récupération des souvenirs des occurrences précédentes et que cette récupération est favorable au rappel final (voir le chapitre 2). Par exemple, la récupération du souvenir de l'épisode P1 au moment de P2 permettrait à la trace récupérée d'être enrichie par la nouvelle occurrence. Au contraire, si le souvenir de P1 ne peut pas être récupéré, une nouvelle trace est créée qui ne profite pas des occurrences antérieures, menant à des performances finales plus faibles. D'une façon générale, l'agencement expansif est celui qui devrait bénéficier du plus grand nombre de récupérations en phase d'étude successives, puisque les IIR sont d'abord courts, permettant la récupération et l'enrichissement de la trace, puis de plus en plus longs. Globalement, les résultats des Expériences 1 à 3 vont donc dans le sens de cette hypothèse.

Plus précisément, dans l'Expérience 2, nous avons mesuré les taux de reconnaissance des paires mot-pseudomot lors de la phase d'apprentissage et avons montré que les paires ayant été reconnues lors de P2 avaient une plus grande probabilité d'être également rappelées lors du test final. Or, ce sont les items de l'agencement expansif qui ont été les mieux reconnus à P2, ce qui est naturellement dû au fait que l'intervalle P1-P2 est plus court dans cet agencement que dans les autres agencements. Cette observation peut être interprétée de la façon suivante. L'agencement expansif est le plus efficace car, comme

les intervalles entre les occurrences sont initialement courts, c'est dans cet agencement que, au moment des premières répétitions, la récupération en phase d'étude est la plus probable. Les longs intervalles entre les dernières occurrences sont moins problématiques pour la récupération en phase d'étude (par exemple, au moment de P4) car les opportunités antérieures de récupération ont permis à la trace de s'enrichir, lui permettant de résister davantage à l'oubli (entre P3 et P4 par exemple). À l'inverse, l'agencement contractant devrait être le moins efficace car il est davantage probable que la seconde occurrence ne permette pas la récupération du souvenir de la première occurrence, ce qui était observé dans l'Expérience 2. Ce premier intervalle étant trop long, l'item a été oublié lorsque celui-ci se répète, et par conséquent, une nouvelle trace doit être créée. L'agencement uniforme devrait être moyennement efficace car les probabilités de récupération en phase d'étude sont intermédiaires. Bien entendu, cette explication de la récupération des traces en « tout ou rien » est schématique ; d'une façon plus modérée et réaliste, une trace résiduelle des occurrences précédentes est probablement toujours présente en mémoire. Cependant ces traces résiduelles peuvent être tellement infimes dans le cas d'un long IIR qu'on considérera que rien n'est récupéré.

La théorie de la récupération en phase d'étude telle qu'elle est présentée dans la littérature ne fait pas de prédiction concernant l'effet du DR. Cependant, on pourrait développer la théorie de la récupération en phase d'étude en postulant que la récupération des occurrences antérieures a deux conséquences : a) la trace récupérée est enrichie par de nouveaux éléments correspondant à l'épisode actuel, ce qui lui permet de bénéficier d'un plus haut niveau d'activation ; b) cette trace récupérée et enrichie devient également plus résistante dans le temps, c'est-à-dire qu'elle sera moins sensible à l'oubli au fil du temps. Par conséquent, les items appris selon un agencement expansif devraient être associés à un rappel final à la fois plus élevé et plus durable. Selon cette hypothèse, l'effet modulateur du DR observé dans l'Expérience 3 pourrait s'expliquer de la façon suivante. Peu après la fin de la phase d'apprentissage, lors de J15 par exemple, les items issus des trois agencements sont rappelés avec la même probabilité car tous ont été présentés deux jours auparavant (P4 à J13), et bénéficient donc tous d'un niveau d'activation relativement élevé. Cependant, au fur et à mesure que le temps passe avec l'allongement du DR, les souvenirs des items des agencements contractant et uniforme subissent un plus grand déclin que ceux de l'agencement expansif, ce qui mène à l'apparition progressive de la supériorité de l'agencement expansif. Le modèle de Pavlik et Anderson (2005) simule ce type de phénomène, bien qu'il n'ait pas été à même de montrer l'avantage de l'agencement expansif (Lindsey et al., 2009).

Il est intéressant de considérer que nous avons abouti à des résultats similaires aux prédictions faites par le modèle de Mozer et al. (2009), relatées dans Lindsey et al. (2009). En effet, les simulations prédisaient que l'agencement expansif serait de plus en plus optimal parmi les trois agencements au fur et à mesure de l'allongement du DR. Or, ce modèle implémentait à la fois les concepts de variabilité de l'encodage et de récupération en phase d'étude. Ainsi nos résultats sont concordants avec les prédictions d'un modèle qui implémente une certaine version des modèles mixtes.

Les modèles mixtes (e.g., Bellezza & Young, 1982) peuvent par conséquent apparaître comme une solution dans laquelle les deux théories se combinent pour donner un modèle complexe pouvant rendre compte d'un certain nombre de résultats. Par exemple, nous avons vu que pour expliquer la modulation de l'effet de l'agencement par le DR (observé dans l'Expérience 3), ni la théorie de la variabilité de l'encodage seule ni la théorie de la récupération en phase d'étude seule ne prédisaient les résultats. D'une façon générale,

il s'agirait de combiner les principes des deux théories, aboutissant aux deux principes suivants : 1) il est nécessaire que les différentes occurrences soient mises en lien dans une trace unique, par le biais de la récupération des traces antérieures lors des répétitions des items ; et 2) les contextes au moment de l'encodage et de la récupération jouent leurs rôles comme le décrit le modèle de la variabilité de l'encodage. Cette combinaison de principes pourrait expliquer pourquoi, après un DR court, l'égalité a été observée entre tous les agencements : même si l'agencement contractant n'est pas optimum concernant la récupération en phase d'étude, il est favorisé par le contexte de récupération proche du contexte d'encodage des dernières occurrences. Ensuite, avec le passage du temps lors du DR, le contexte instancié lors de la récupération est de plus en plus différent des contextes instanciés à l'encodage, ce qui favorise l'agencement expansif qui a permis aux traces correspondantes d'être plus riches, grâce aux récupérations en phase d'étude successives lors de la phase d'apprentissage. Un tel modèle peut atteindre des niveaux élevés de complexité et nécessite beaucoup de suppositions ; par conséquent, une description verbale du modèle et de ses prédictions (notamment par rapport à l'interaction entre agencement et DR) est limitée. C'est pourquoi des modélisations mathématiques telles que celles de Mozer et al. (2009) présentent un grand intérêt.

Pour revenir à la théorie de la récupération en phase d'étude, nous pensons qu'il reste des améliorations à apporter quant à sa formulation. Par exemple, la théorie de la remémoration récursive telle que formulée par Hintzman (2004 ; 2010) postule que les occurrences successives sont récupérées de façon enchâssée : à P2, on récupère la trace de P1 ; A P3, on récupère la trace correspondant au souvenir que lors de P2, on avait récupéré P1 ; A P4, on récupère la trace correspondant au souvenir qu'à P3, on avait récupéré la trace de P2 qui elle-même contenait le souvenir de la récupération de P1. Et ainsi de suite... On peut se demander si un processus aussi complexe est vraiment réaliste dans le cadre de répétitions sur plusieurs jours. D'autre part, la théorie de la récupération en phase d'étude est intimement liée à des questions telles que : Que récupère-t-on réellement lors d'une répétition ? Cette récupération est-elle associée à un état conscient particulier ? La reconnaissance est-elle vraiment le corrélat de cette récupération, comme nous l'avons postulé dans nos expériences ? Ces questions ne font pas l'objet de beaucoup de développement dans la littérature, et de nouvelles études sont nécessaires pour y répondre.

6. 2. 3. Autres théories

Les résultats observés dans nos expériences nous apportent assez peu d'éléments pour discuter des autres théories rendant compte de l'effet de pratique distribuée.

Tout d'abord, en ce qui concerne l'hypothèse de traitement déficitaire (section 2.4), il semble peu probable qu'après des IIR de 24 h ou plus, des effets d'amorçage influencent les traitements réalisés par les sujets sur les items (voir cependant Tulving, Hayman, & Macdonald, 1991, pour l'existence d'effets d'amorçage à long terme). De plus, une hypothèse basée sur le fait que les effets d'amorçage sont délétères à la rétention finale devrait être en faveur des agencements pour lesquels les IIR sont les plus longs ; or l'intervalle moyen des IIR est identique dans les différents agencements. Si on se focalise uniquement sur le premier intervalle, alors l'agencement de type contractant devrait être le moins impacté par les effets d'amorçage. En toute logique, des effets de plus grande amplitude, comme les effets de récupération en phase d'étude, devraient surpasser les conséquences d'éventuels effets d'amorçage.

Concernant les hypothèses de l'accessibilité et de la reconstruction, dans la mesure où l'apprentissage a consisté en de simples répétitions des items, et non en des tests, il semble difficile de parler de reconstruction. Cependant, il est envisageable que l'accessibilité des souvenirs, c'est-à-dire la difficulté avec laquelle s'est produit la récupération des souvenirs, ait eu une influence sur les propriétés de la trace récupérée et enrichie. Ainsi, plus le souvenir est lointain (i.e., long IIR), plus la trace serait difficile à récupérer, mais plus l'impact de cette récupération serait fort pour la rétention ultérieure de l'item. Ces considérations pourraient servir à développer la théorie de la récupération en phase d'étude, c'est d'ailleurs ce que font Benjamin et Tullis (2010) dans leur proposition de modèle de remémoration récursive.

Les hypothèses mettant en avant les processus métacognitifs (e.g., Bahrack & Hall, 2005) postulent que de longs IIR favoriseraient chez le sujet le réajustement des stratégies d'apprentissage des items, contrairement aux courts IIR qui le favoriseraient moins. De tels phénomènes volontaires d'amélioration de la stratégie d'apprentissage ont probablement eu lieu lors des sessions chez les sujets ; on peut imaginer le genre de monologue intérieur suivant : « Je me souviens avoir vu cette paire auparavant, mais je ne me souviens plus du moyen mnémotechnique utilisé, je vais donc devoir en trouver un nouveau ». Cependant, il n'est pas évident de décrire comment ceux-ci auraient pu, à eux seuls, expliquer la différence entre les agencements temporels.

Au regard des effets qu'ont pu avoir les épisodes de sommeil sur les résultats, peu d'éléments peuvent être discutés étant donné que, dans tous les cas, au moins une nuit de sommeil était intercalée entre deux répétitions successives, et ce dans tous les agencements. De plus, plusieurs nuits de sommeil se produisaient pendant le DR. Ainsi, bien que le sommeil ait eu un rôle très probable dans les processus de consolidation mnésique des items appris, il est très difficile de prédire comment le sommeil pourrait expliquer à lui seul les effets de l'agencement temporel.

6. 2. 4. Conclusions

Nous concluons que les effets de l'agencement temporel des répétitions que nous avons observés dans les Expériences 1 à 3 sont plus facilement explicables par la théorie de la récupération en phase d'étude que par la théorie de la variabilité de l'encodage, bien que la théorie de la récupération en phase d'étude nécessite alors d'être affinée pour rendre exactement compte des observations. Cependant, une explication mixte combinant ces deux phénomènes pourraient également être adoptée, mais plus d'éléments expérimentaux sont nécessaires. Nous allons maintenant discuter de la notion de taux d'oubli.

6. 3. Délai de rétention et taux d'oubli

6. 3. 1. Taux d'oubli et effet de l'agencement temporel

L'observation réellement novatrice de nos travaux est l'observation, dans l'Expérience 3, que l'agencement expansif semblait associé à un taux d'oubli plus faible que les autres agencements, et ce alors que les niveaux de performance initiale (J15) étaient similaires. Nous nous sommes interrogée sur les taux de rétention observés dans les études sur l'effet de l'agencement à long terme qui utilisaient plusieurs DR (Cull, 2000 ; Tsai, 1927). Les représentations graphiques des courbes d'oubli de ces études sont présentées dans les Figures 27 et 28, tandis que les Figures 29 et 30 présentent les courbes d'oubli observées dans les Expériences 2 et 3.

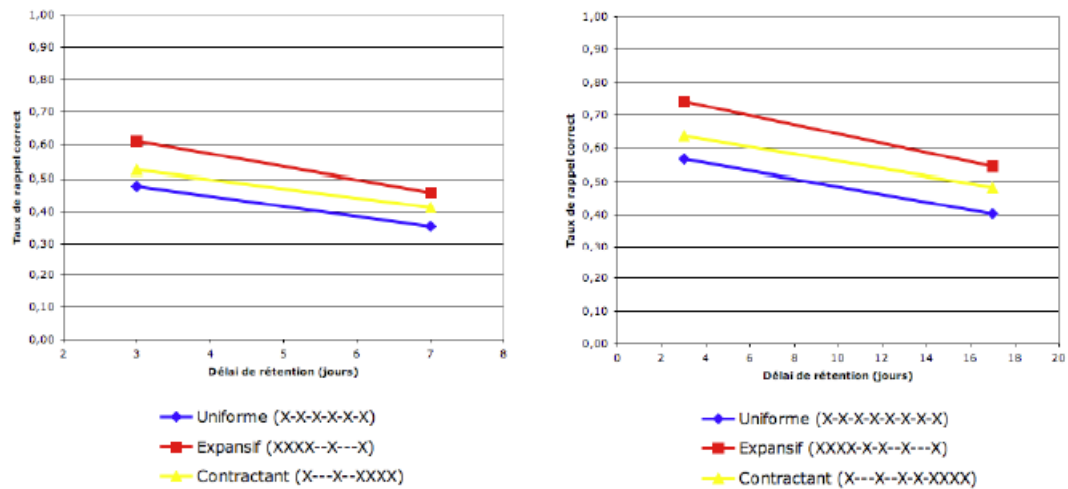


Figure 27 : Reproduction des résultats de l'étude de Tsai (1927 ; Expérience 2 à gauche et Expérience 3 à droite). Taux de rappel en fonction du délai de rétention (abscisses) et de l'agencement (couleur de la courbe).

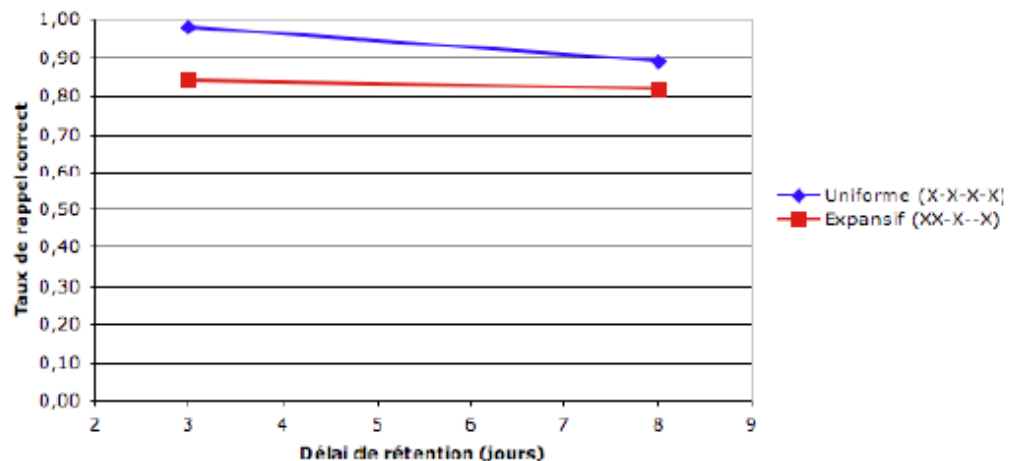


Figure 28 : Reproduction des résultats de l'étude de Cull (2000, Expériences 3 et 4). Taux de rappel en fonction du délai de rétention (abscisses) et de l'agencement (couleur de la courbe).

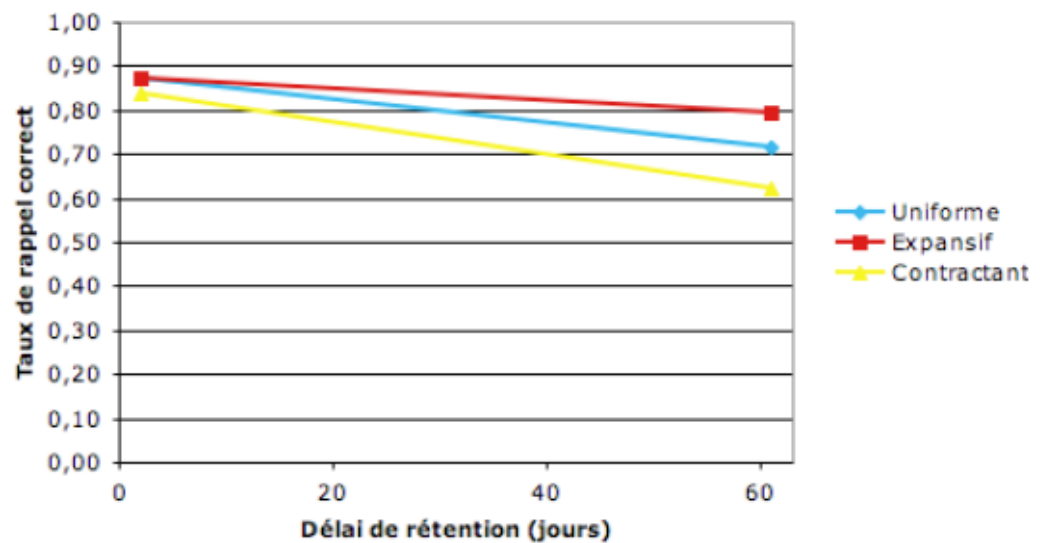


Figure 29 : Résultats de l'Expérience 2. Taux de rappel en fonction du délai de rétention (abscisses) et de l'agencement (couleur).

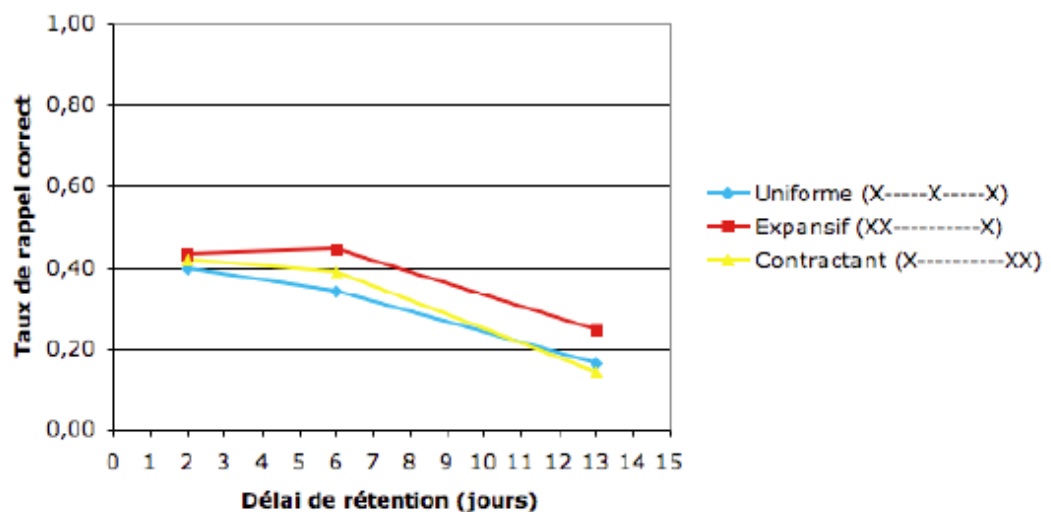


Figure 30 : Résultats de l'Expérience 3. Taux de rappel en fonction du délai de rétention (abscisses) et de l'agencement (couleur).

Avant toute discussion, il est nécessaire de rappeler qu'aucune de ces observations ne constitue une mesure convenable de l'oubli en fonction du DR, et ce pour les raisons suivantes. Premièrement, dans certains cas (i.e., Tsai, 1927 ; Expérience 2), les mêmes sujets étaient testés deux fois sur l'intégralité du matériel ; or on sait que le premier épisode de rappel modifie les processus de consolidation en cours et, par conséquent, les résultats lors du deuxième rappel étaient probablement différents de ceux qu'ils auraient été si le premier rappel n'avait pas eu lieu. Deuxièmement, dans d'autres cas (i.e., Cull, 2000 ; Expérience 3) des sujets différents étaient testés lors de deux DR, or les grandes variabilités individuelles inhérentes aux tâches de rappel rendent difficiles la comparaison

inter-groupes. Troisièmement, dans d'autres cas encore (i.e., Expériences 2), lorsque les participants étaient testés de nouveau au bout de deux mois, seul un sous-ensemble de l'échantillon initial a participé, ce qui rend les résultats moins représentatifs que si tous avaient participé.

L'observation de ces courbes d'oubli, avec toutes les réserves mentionnées, fait apparaître que l'agencement expansif était souvent associé au taux d'oubli le plus faible (Cull, 2000 ; Expérience 2 ; Expérience 3). Dans l'étude de Tsai (1927) cependant, c'est l'agencement contractant qui était associé au taux d'oubli le plus faible. On remarque cependant que dans ce dernier cas, l'agencement expansif était alors celui qui entraînait les meilleures performances à la tâche de rappel. Ainsi, dans l'ensemble de ces études, l'agencement expansif présentait soit l'avantage numérique, soit l'avantage en terme de taux d'oubli, soit les deux.

Ainsi, l'apprentissage grâce à l'agencement expansif favoriserait une rétention optimale car plus durable dans le temps que les autres agencements. Ces résultats suggèrent également, dans le cadre de la théorie de la récupération en phase d'étude, que les récupérations successives (qui sont optimisées dans l'agencement expansif) en plus de permettre un enrichissement de la trace, amélioreraient sa résistance dans le temps. Les concepts de *force* et de *résistance* de la trace mnésique (e. g., Wickelgren, voir la section 3.3.2) pourraient être mises à contribution pour expliquer ces effets. Par exemple, l'agencement expansif pourrait à la fois favoriser la force et à la résistance des traces mnésiques, ou bien uniquement leur résistance. Des modèles mathématiques qui implémentent de telles notions pourront aider à comprendre le phénomène. En tout état de cause, davantage de données sont nécessaires, et en particulier sur ce qui se produit pendant la phase d'apprentissage.

De façon intéressante, si l'on adhère à la définition de la consolidation comme un processus rendant les traces résistantes à l'oubli (e.g., Litman & Davachi, 2008), alors l'agencement expansif semble être celui qui favorise le plus la consolidation. L'importance de générer des traces mnésiques dont la résistance dans le temps est importante, c'est-à-dire des souvenirs durables, est évidente dans le domaine éducatif, puisque le but de l'enseignement est d'inculquer des connaissances durables, et non permettre de passer un examen à un moment donné.

Enfin, ces observations sont également cohérentes avec les conceptions évolutionnistes développées par Anderson et Schooler (1991 ; section 3.3.5). Si le système cognitif est capable de détecter qu'un item se répète de façon initialement fréquente, puis de moins en moins fréquemment (i.e., selon un agencement expansif), il devrait être capable de prédire que l'on aura besoin de cet item dans le futur avec des intervalles de plus en plus longs ; il est donc nécessaire de maintenir un taux de rétention élevé pour cet item. À l'inverse, un item qui se répète avec des intervalles de plus en plus courts (i.e., selon un agencement contractant) va logiquement finir par devenir omniprésent dans l'environnement ; par conséquent, seule une rétention à court terme sera nécessaire. Entre les deux, l'agencement de type uniforme devrait logiquement induire un taux de rétention moyen.

6. 3. 2. Taux d'oubli et effet de pratique distribuée

Ces observations mettent en évidence plusieurs points. Premièrement, nos résultats révèlent l'importance d'inclure plusieurs modalités du facteur DR dans les expériences, et de ne pas tirer de conclusions définitives sur l'effet d'un facteur lorsqu'on a testé la rétention après un DR unique. Ces considérations valent pour les facteurs IIR et agencement, mais

également pour d'autres comme par exemple la modalité des essais (*Etude seule* versus *Test-étude* ; e.g., Wheeler, Ewers, & Buonanno, 2003). En effet, les effets observés peuvent être différents, voire opposés, lorsqu'on fait varier la durée du DR. Cela explique peut-être pourquoi, dans la littérature, des effets parfois contradictoires ont été observés ; les DR choisis correspondaient peut-être à des DR pour lesquels il n'y avait pas de différences entre les conditions, tandis que d'autres DR auraient pu mettre en évidence des effets. Il serait par conséquent très intéressant de répliquer les Expériences 1 et 2 avec des DR différents. Selon notre interprétation que l'agencement expansif est associé à un taux d'oubli plus faible, alors l'écart entre les agencements devrait se creuser, tout comme dans l'Expérience 3. À l'inverse, il devrait se resserrer pour un DR plus court.

Deuxièmement, ces constatations nous amènent à suggérer que, dans les études sur l'effet de pratique distribuée, il serait avantageux d'adopter une nouvelle perspective. En effet, les auteurs s'intéressent classiquement à l'effet de l'IIR (c'est-à-dire quel est l'IIR optimal) en fonction du DR. A contrario, on pourrait plutôt comparer la rétention induite par chacun des IIR en fonction du DR. Ce changement de perspective serait associé à un remaniement des figures représentant les résultats. Concrètement, l'axe des abscisses représenterait le délai de rétention (et non plus les IIR) tandis que les différents IIR (et non plus les DR) seraient représentés par des courbes différentes. Cette nouvelle façon de considérer les données aurait l'avantage, selon nous, de rendre plus facile l'interprétation des courbes. Par exemple, dans l'ancienne perspective, on pouvait interpréter les courbes de la façon suivante : « Au fur et à mesure que le DR augmente, l'IIR optimal augmente également ». La même observation dans la nouvelle perspective deviendrait : « Avec l'augmentation du DR, on assiste à l'avènement progressif de la supériorité des longs IIR ». À titre d'exemple, nous avons transformé la représentation graphique des résultats de Glenberg (1976, représentés dans la Figure 2 dans la section 1.3.1) ; la nouvelle mise en forme est présentée dans la Figure 31.

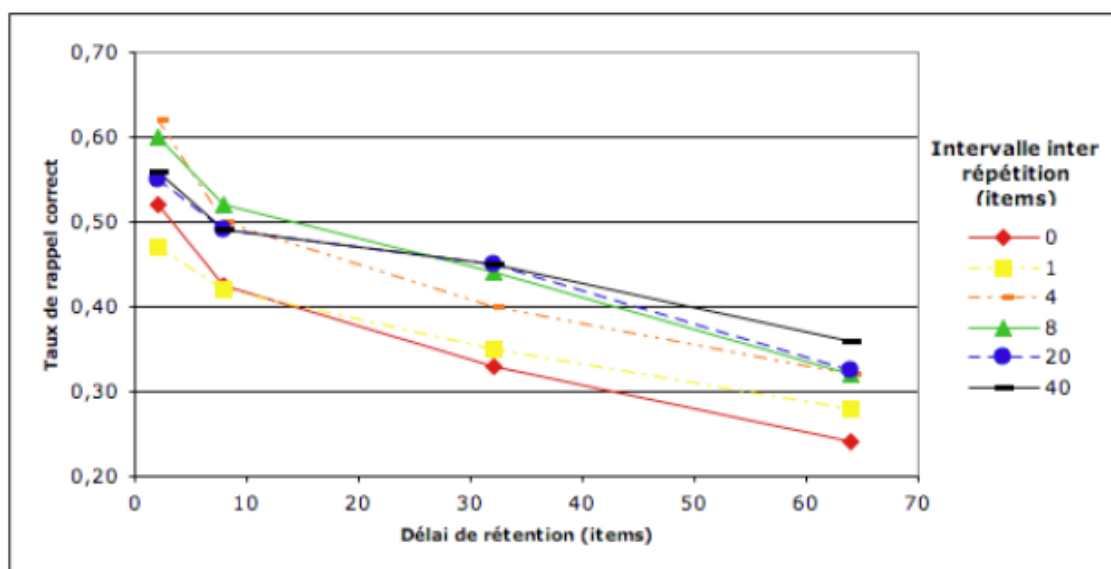


Figure 31 : Reproduction des résultats de l'Expérience 1 de Glenberg (1976). Taux de rappel en fonction du délai de rétention (abscisses) et de la durée de l'IIR (couleur).

De la même façon sont représentés dans la Figure 32 les résultats de l'étude de Cepeda et al. (2008), figure qui peut être comparée avec la Figure 4 de la section 1.3.2. On notera que, dans ces deux figures, est retrouvée la forme classique de la courbe d'oubli.

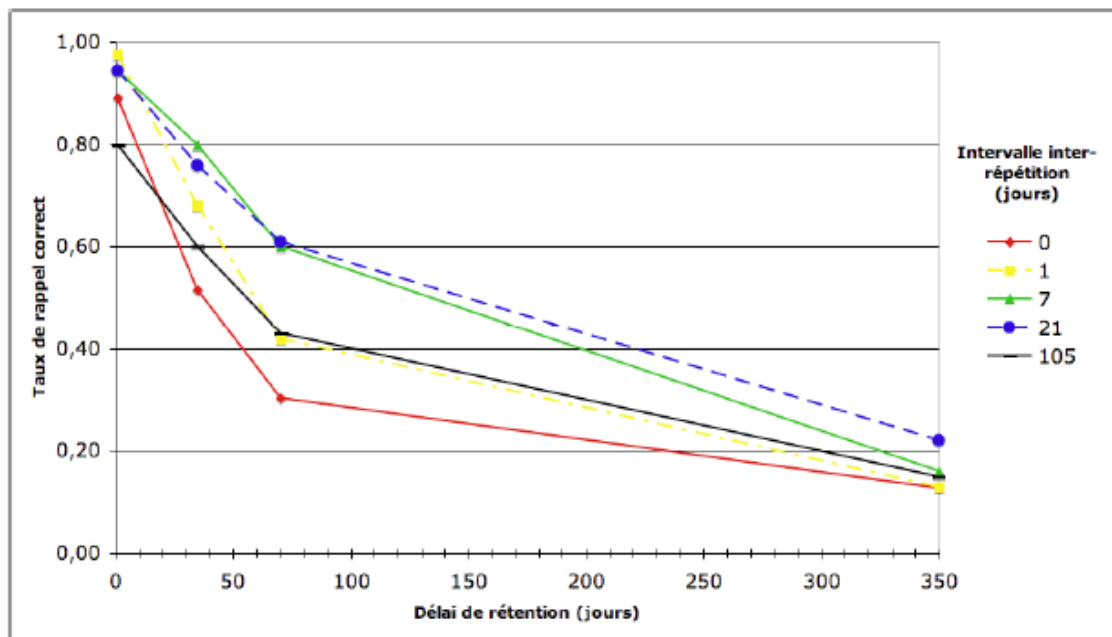


Figure 32 :Reproduction des résultats de l'étude de Cepeda et al. (2008). Taux de rappel en fonction du délai de rétention (abscisses) et de la durée de l'IIR (couleur).

L'observation de ces courbes révèle que des IIR de longue durée sont souvent associés à la fois aux performances de rétention les plus élevées et également aux taux d'oubli les plus faibles. C'est le cas par exemple des conditions où l'IIR est de 20 items dans l'étude de Glenberg (1976), et de 21 jours dans l'étude de Cepeda et al. (2008). Il nous semble qu'il serait plus simple d'expliquer ces résultats en essayant de comprendre en quoi l'IIR influence la courbe d'oubli. La courbe d'oubli suit une forme connue (e.g., Wixted & Carpenter, 2007 ; Wixted & Ebbesen, 1991), dont deux paramètres peuvent varier : d'une part, l'ordonnée à l'origine, c'est-à-dire les performances réalisées immédiatement à la suite de l'apprentissage et, d'autre part, la pente, c'est-à-dire la vitesse de déclin des souvenirs. Nous pensons que la théorie de la récupération en phase d'étude pourrait être améliorée pour pouvoir rendre compte de ces phénomènes, notamment en considérant que la récupération pourrait avoir des effets différentiels à la fois sur la force d'un souvenir et sur son taux de déclin (i.e., sa résistance).

6. 4. Effet de l'agencement selon l'échelle de temps

Nos résultats, en particulier ceux de l'Expérience 3, sont en contradiction avec les résultats des études récentes étudiant l'effet de l'agencement temporel sur une courte échelle de temps (e.g., Karpicke & Roediger 2007 ; Logan & Balota, 2008). En effet, nos résultats suggèrent que l'avantage de l'agencement expansif des répétitions sur plusieurs jours est de plus en plus important au fur et à mesure que le DR s'allonge. Au contraire, les études réalisées sur le court terme (i.e., au sein d'une seule session expérimentale) ont montré que l'avantage de l'agencement expansif parfois observé suite à un DR court (i.e., quelques minutes) est remplacé par l'avantage de l'agencement uniforme lorsque le DR est considérablement allongé (i.e., 24 h ou plus). Ainsi, l'ensemble de ces études suggère que la durée du DR module de façon différente l'effet de l'agencement temporel sur les deux échelles de temps.

Cette observation remet en question l'idée que les études à court terme constituent des simulations « en miniature » des études à long terme. Cette idée semble pourtant admise parmi les auteurs des articles récents (e.g., Cepeda et al., 2006). En effet, il n'est pas développé explicitement dans la littérature l'idée que ce qui se produit lorsque les répétitions ont lieu lors de sessions différentes (et en particulier avec des IIR de 24h ou plus) puisse refléter des phénomènes différents de ceux en jeu lorsque les répétitions et le test ont lieu au sein d'une même session. Par exemple, selon Cepeda et al. (2008), il existerait une loi générale relative aux effets joints entre l'IIR et le DR, loi valable sur toutes les échelles de temps et reflétée par l'équation décrite dans l'article.

Pourtant, nos résultats constituent des arguments expérimentaux allant à l'encontre de cette idée. Il faut toutefois admettre que nos expériences ne constituent pas une réplication exacte à long terme des études sur l'effet de l'agencement à court terme. Par exemple, dans ces expériences (e.g., Karpicke & Roediger, 2007 ; Logan & Balota, 2008), les IIR étaient de quelques items et les DR étaient de 24 ou 48h. Si on voulait respecter ce ratio à long terme dans nos expériences où les IIR sont de plusieurs jours, il faudrait tester la rétention après des DR de plusieurs années. À ce propos, Weigold (2008) a réalisé une étude pour comparer un protocole à long terme avec son équivalent, toutes proportions conservées, à court terme. Bien que ses objectifs étaient plus spécifiquement de comparer les effets du type d'essai (*Test-étude* vs *Étude seule*) en fonction de l'échelle de temps, il n'a pas observé les mêmes effets à court et à long terme. Il serait donc intéressant d'essayer, comme Weigold, de répliquer « en miniature » nos expériences, en respectant exactement les ratios entre les IIR et les DR.

Plusieurs facteurs pourraient expliquer pourquoi les résultats sont différents dans les deux échelles de temps. Premièrement, en ce qui concerne les phénomènes d'oubli (qui sont notamment en jeu dans les processus de récupération en phase d'étude), on peut penser que ceux-ci sont dépendants de la durée des IIR en valeur absolue et non en valeur relative. En effet, l'oubli n'est pas censé dépendre de ce qui va venir après ; ce n'est pas parce que la prochaine occurrence d'un item aura lieu dans 10 jours que l'on va oublier plus ou moins vite, et ce, d'autant plus quand on ne sait pas quand aura lieu la prochaine occurrence. Ainsi, les processus de récupération en phase d'étude, puisqu'ils sont étroitement liés à l'oubli, seront nécessairement différents suite à un IIR de quelques minutes que suite à un IIR de plusieurs jours. Un deuxième facteur qui différencie les deux types d'étude est le fait que, dans les expériences à long terme, il y a une réelle interruption de la situation expérimentale entre les sessions, contrairement aux études à court terme. Cela joue probablement un rôle dans les effets obtenus. Un troisième facteur est la présence d'épisodes de sommeil entre les répétitions et pendant le DR ; comme nous l'avons vu dans le chapitre 3, le sommeil a un effet de consolidation des traces mnésiques et en particulier, selon les recherches récentes un rôle de renforcement des traces contre les interférences (e.g., Ellenbogen et al., 2009). Enfin, les stratégies mises en œuvre par le sujet seront peut-être différentes si celui-ci sait que le test aura lieu plusieurs jours après ou quelques minutes après. Tous ces éléments font que, probablement, les effets de pratique distribuée dans deux échelles de temps ne diffèrent pas uniquement par l'unité de mesure du temps.

Enfin, ces questions sont importantes dans la mesure où elles indiquent qu'il faut être prudent quand il s'agit de faire des recommandations pratiques sur les stratégies d'espacement à long terme alors qu'on se base sur les résultats d'études à court terme.

6. 5. Généralisation des résultats

Dans une optique d'application de nos résultats en dehors du laboratoire, il est nécessaire de se demander si les effets observés, c'est-à-dire la supériorité de l'agencement expansif pour la rétention à long terme, sont de nature aussi robuste et ubiquitaire que les effets de pratique distribuée. Dans quelle mesure peut-on considérer que ces résultats sont généralisables ?

Avant tout autre considération, il faut préciser que nous avons évoqué jusqu'à présent les notions générales d'agencement expansif, uniforme et contractant. Cependant, seule la détermination de l'agencement uniforme est totalement contrainte ; en effet, pour une durée fixe entre la première et la dernière présentation et pour un nombre donné d'occurrences de l'item, alors la durée des IIR est déterminée strictement par la durée totale divisée par le nombre de répétitions moins un. Au contraire, les agencements expansif et contractant peuvent être de plusieurs types. Par exemple, dans l'Expérience 3, plusieurs agencements expansifs auraient pu être utilisés, comme par exemple : X-X-----X, X--X-----X, ou encore X---X-----X, et ainsi de suite. (L'agencement contractant aurait été le symétrique de l'agencement expansif.) Il ne faut pas perdre de vue que, selon la théorie de la récupération en phase d'étude, c'est la durée absolue entre les présentations qui est importante pour les phénomènes de récupération pendant l'apprentissage. Ainsi, par exemple, un agencement expansif du type X---X-----X sera peut-être moins efficace qu'un agencement contractant du type X---X-X du fait de la durée de l'intervalle entre P1 et P2. Il faut également prendre en compte la durée totale de la période d'apprentissage, en postulant que globalement, plus elle est longue plus elle est bénéfique à la rétention. On ne peut donc évidemment pas conclure que *tous* les agencements expansifs sont plus efficaces que *tous* les agencements contractants. D'autre part, la comparaison de plusieurs agencements expansifs entre eux pourrait être très intéressante également, et aucune donnée n'est disponible sur cette question à notre connaissance. Ces considérations mettent en évidence le grand nombre de combinaisons possibles associées au choix des agencements puisqu'il faut considérer le nombre de répétitions, la durée de la phase d'apprentissage, la durée des IIR en valeur absolue, le rapport entre les IIR successifs, la durée du DR, etc. Une étude exhaustive de tous ces facteurs et de leur interaction est impossible, et pour cela, des outils de simulation mathématique peuvent trouver toute leur utilité, ce qui était l'objectif de l'étude de Lindsey et al. (2009 ; voir la section 3.2.4).

L'étude de l'effet de pratique distribuée soulève d'autres points de discussion. Premièrement, on peut s'interroger sur les différences inter-individuelles dans l'effet de l'agencement. Pourquoi l'agencement optimal ne serait-il pas différent selon les personnes et, en particulier, selon les capacités mnésiques de chacun ? En effet, si on se place dans le cadre de la récupération en phase d'étude en postulant que la récupération est nécessaire, il est légitime de penser que les personnes ayant des capacités mnésiques « faibles » auront besoin d'IIR initiaux plus courts que ceux ayant des capacités mnésiques « élevées ». Quelques études se sont intéressées à cette question : Weigold (2008), par exemple, a tenté de savoir si l'effet de pratique distribuée était modulé par les aptitudes académiques des étudiants, mais n'a pas trouvé d'éléments en faveur d'une telle modulation. Madsen (1963) a mis en évidence que l'effet d'espacement n'apparaissait que chez les sujets à faible QI mais pas chez les autres. Cull et al. (1996, Expérience 4) ont formé deux groupes de niveau selon la performance de rappel d'items présentés une fois ; il s'est avéré que la rétention était plus élevée pour l'agencement expansif (1-5-9) que pour l'agencement uniforme (5-5-5) chez le groupe le plus faible, mais sans différence significative chez le groupe au score le plus élevé. Ainsi, il faudrait idéalement pouvoir ajuster les agencements en fonction des capacités de chacun. Cependant, il nous semble raisonnable de penser que, si on doit choisir entre des agencements pour lesquels l'IIR moyen est identique, l'agencement expansif sera au mieux

le plus efficace, au pire de même efficacité que les autres. Nous ne voyons pas de raison de penser qu'un agencement uniforme, et encore moins contractant, puisse être le plus efficace pour une personne donnée.

Deuxièmement se pose la question de la généralisation à d'autres types de matériel et également aux autres modalités de présentation des stimuli. La littérature de l'effet d'espacement semble suggérer que celui-ci s'observe pour un grand ensemble de stimuli verbaux. En effet, la méta-analyse de Cepeda et al. (2006), par exemple, englobait les études ayant utilisé des présentations de mots, de paires verbales, de faits, de phrases, de paragraphes de texte, d'objets, de paires visage-nom, d'images, etc. Il n'y a pas de raison évidente de prédire que la supériorité de l'agencement expansif serait remise en question lorsqu'un matériel verbal autre que des paires de vocabulaire est utilisé. Il faut toutefois prendre en compte précisément ce qui doit être appris lorsque du matériel verbal est présenté. Par exemple, dans nos études utilisant des paires mot-pseudomot, les sujets devaient apprendre d'une part le pseudomot qui était pour eux un stimulus nouveau et, d'autre part, l'association entre ce nouveau stimulus et le mot qui, lui, possède déjà une représentation en mémoire. Or, des travaux récents en imagerie cérébrale fonctionnelle suggèrent que les régions cérébrales sous-tendant d'une part l'apprentissage d'association d'items et d'autre part l'apprentissage d'items nouveaux sont différentes (e.g., Litman & Davachi, 2008). On ne peut donc pas exclure que les effets de pratique distribuée peuvent se manifester différemment pour ces deux types d'apprentissage. Des travaux complémentaires sont nécessaires pour en savoir plus sur cette question.

Il a été observé que l'effet de pratique distribuée (en particulier l'effet d'intervalle à court terme) peut varier en fonction de la tâche mnésique employée pour évaluer la rétention, notamment entre des tâches de rappel libre, de rappel indicé, et de reconnaissance (e.g., Greene, 1989). Par exemple, l'IIR optimal n'est pas toujours le même dans ces trois tâches bien qu'elles mesurent toutes la rétention explicite. Cela s'expliquerait par le fait que les influences des indices de récupération instanciés lors de la récupération (par l'intermédiaire du contexte environnemental et des indices fournis par la tâche elle-même) seraient différentes dans chacune des tâches. Par exemple, en tâche de reconnaissance, les stimuli sont présentés en tant que tel, tandis qu'en rappel libre, aucun indice explicite n'est présenté et le sujet doit alors utiliser d'autres stratégies pour récupérer les stimuli. Dans nos expériences, nous nous sommes concentrée sur une tâche de rappel indicé, et il serait intéressant d'explorer si les mêmes effets d'agencement sont observés pour d'autres tâches de rétention.

Toujours relativement à la question de l'extension des effets à d'autres types de matériel, il se trouve que peu d'études ont été menées sur l'effet de pratique distribué avec du matériel non verbal. Peu de données sont donc disponibles concernant ces effets avec, par exemple, des formes nouvelles, des images non verbalisables, des sons non verbaux comme la musique, et des stimuli présentés dans des modalités inhabituelles comme le goût et l'odorat. Cependant, d'une façon générale, si l'on adhère à la théorie de la récupération en phase d'étude, celle-ci devrait également imposer le principe selon lequel les IIR successifs doivent permettre de récupérer à chaque occurrence les traces mnésiques créées par les occurrences précédentes. Il y a cependant au moins un domaine pour lequel l'effet d'espacement n'a pas pu être mis en évidence : l'induction (Kornell & Bjork, 2008). Par conséquent, tout laisse à penser que les agencements n'auraient pas d'effet non plus dans ce domaine. Enfin, il existe également une littérature de l'effet de pratique distribuée pour les souvenirs procéduraux comme les mouvements séquentiels (pour revue, voir Lee & Genovese, 1988 ; Donovan & Radosovich, 1999), et pour lesquels les effets de pratique

distribuée classiques semblent être observés. Toutefois aucune étude à notre connaissance n'a testé l'effet de différents agencements sur la rétention de procédures à court terme ou à long terme. Il s'agirait d'une piste intéressante à explorer étant donné les applications potentielles, notamment dans la sphère professionnelle ou sportive.

Enfin, on peut se demander si l'avantage de l'agencement expansif serait observé également dans le cadre d'un apprentissage non volontaire de la part des participants. Aucune étude à notre connaissance ne s'est intéressée à l'effet de pratique distribué sur le long terme dans un contexte d'apprentissage incident. Cependant, dans le cadre de la théorie de la récupération en phase d'étude, le fait de ne pas donner la consigne d'apprendre n'empêche pas la récupération des traces des occurrences précédentes et ainsi, il semble cohérent que l'agencement expansif soit toujours le plus efficace. En revanche, le fait de ne pas apprendre volontairement devrait rendre les traces mnésiques moins fortes et donc la fenêtre temporelle permettant la récupération des traces lors des occurrences ultérieures devrait probablement être plus réduite, ce qui, en toute logique, favoriserait toujours l'avantage de l'agencement expansif sur les autres agencements.

6. 6. Discussion des concepts utilisés

Notre problématique nous a amenée à faire appel à certaines notions qui mériteraient elles-mêmes d'être développées davantage, mais que la limitation de place et de temps nous empêche de développer plus avant. Nous aborderons successivement la question du terme de *trace* mnésique, des concepts d'encodage/stockage/récupération, du lien entre les échelles d'étude de niveau cellulaire et psychologique, et enfin la question des modèles théoriques de la mémoire à long terme.

En particulier, nous avons beaucoup utilisé le terme de trace mnésique, particulièrement la notion de récupération de trace, sans réellement définir ou discuter de ce terme. Une trace est, par définition, l'empreinte laissée par un événement passé. Ainsi, en psychologie, le terme de trace est utile dans le sens où il permet de faire référence de façon commode aux phénomènes induits - dans l'esprit et le cerveau - par la perception de stimuli. Cette simplicité a par ailleurs l'inconvénient de masquer la complexité de l'ensemble des processus cognitifs et cérébraux qui induisent la mémoire à long terme. En psychologie, les traces mnésiques sont impalpables et immatérielles mais se reflètent indirectement par le résultat de la récupération dans une tâche de mémoire. On considère que les traces sont dynamiques ; elles évoluent avec le temps et peuvent décliner, se transformer, ou s'enrichir. L'imagerie fonctionnelle permet d'obtenir des informations d'un autre niveau sur les traces mnésiques ; lorsqu'on rappelle une information, les zones spécifiquement activées sont, logiquement, celles où sont "stockées" les traces (dans la mesure où on peut réellement parler de stockage, étant donné les hypothèses selon lesquelles on reconstruit le souvenir plutôt que de le récupérer). Des années de recherche sont encore nécessaires pour mieux comprendre ces phénomènes, et pour s'accorder sur une définition précise, commune aux champs de la psychologie et de la physiologie.

De la même façon, les théories de la mémoire font appel aux notions d'encodage, de stockage et de récupération. Nous pensons également que ces termes sont des outils commodes pour se représenter de façon intelligible la complexité des phénomènes cognitifs et physiologiques en jeu dans le champ de la mémoire. En tout état de cause, chacune de ces grandes étapes a un fonctionnement plus ou moins distinct, bien qu'il soit en pratique difficile de les séparer. La distribution des répétitions dans le temps joue probablement un rôle sur ces trois processus.

Par ailleurs, nous avons fait référence à des découvertes relatives à l'effet de pratique distribuée à une échelle microscopique (section 3.3.1). Bien entendu, il y existe un écart immense, sur les échelles à la fois temporelle et spatiale, entre les phénomènes neurophysiologiques de niveau cellulaire chez la drosophile à un extrême de l'échelle, et les effets comportementaux de la pratique distribuée sur plusieurs jours chez l'humain à l'autre extrême de l'échelle. Nous ne prétendons pas réduire les processus psychologiques observés aux mécanismes moléculaires que nous décrivons. Nous pensons plutôt qu'il y a des interactions mutuelles entre ces deux échelles extrêmes : les processus psychologiques d'apprentissage des stimuli influencent l'anatomie et le fonctionnement des neurones, qui eux-mêmes ont une influence sur le comportement mnésique (lors du rappel par exemple). En outre, un ensemble de niveaux intermédiaires sont intercalés entre ces deux échelles extrêmes : les connexions synaptiques, les regroupements de neurones, les aires cérébrales, etc... Des années de recherche sont encore nécessaires pour faire le lien entre ces échelles de temps et d'espace et pour faire se rapprocher les traditions neurophysiologiques et psychologiques.

Enfin, nous ne nous sommes pas positionnée dans une conception particulière de la mémoire, par exemple en mettant en exergue l'opposition entre les conceptions structuraliste et fonctionnaliste de la mémoire à long terme. Une des particularités de l'étude des effets de pratique distribuée est que la démarche ne suit pas la logique des recherches visant à démontrer la validité d'une conception théorique. Par exemple, si l'on souhaite tester la validité d'un modèle de mémoire, on réalise des expériences qui mettent en évidence des effets de telle ou telle nature, qui, en retour, appuient ou non l'hypothèse initiale. Dans le domaine de la pratique distribuée, il s'agit d'expliquer un phénomène mnésique robuste et connu depuis longtemps, et pour cela on émet des hypothèses qui s'avèrent validées ou non ; en quelque sorte on part des effets pour arriver aux causes, contrairement à la démarche habituelle. Cependant, il est possible que les découvertes faites dans un domaine puissent servir d'éléments pour avancer dans l'autre domaine.

6. 7. Perspectives de recherche

En premier lieu, il s'avère indispensable de conduire de nouvelles études dans le but de répliquer l'effet de modulation de l'effet de l'agencement par le délai de rétention (comme observé dans l'Expérience 3). En particulier, il serait très intéressant de reproduire les Expériences 1 et 2 avec un délai de rétention plus long, par exemple deux semaines. Si notre hypothèse que l'agencement expansif est associé au taux d'oubli le plus faible est correcte, alors l'écart entre les résultats de l'agencement expansif et ceux des autres agencements devrait être amplifié dans l'Expérience 1 et devrait apparaître dans l'Expérience 2.

Par ailleurs, il est également nécessaire de continuer à explorer l'hypothèse de la récupération en phase d'étude. L'idéal serait de ne pas avoir à modifier la tâche d'apprentissage (comme nous l'avons fait dans l'Expérience 2 en y ajoutant une tâche de reconnaissance), ce qui peut possiblement interagir avec les effets d'agencement. Toutefois, l'apport des techniques d'imagerie pourrait peut-être apporter de nouveaux éléments. Une étape préalable sera alors d'étudier la situation, plus simple, où seulement deux occurrences surviennent sur plusieurs jours, séparées par un IIR court (e.g., 1 jour) ou par un IIR long (e.g., 7 jours) ; on pourrait alors faire réaliser les sessions correspondant à P1, à P2, et au test dans le scanner IRM et comparer les activations cérébrales induites par ces deux conditions. Ces résultats apporteraient peut-être des informations sur les processus de récupération en phase d'étude et sur les effets d'amorçage à long terme. Par

exemple, on pourrait examiner si les effets d'une répétition après 24 h sont similaires à ceux d'une répétition après 7 jours. On peut également se demander si les activations observées lors du rappel sont différentes selon la durée de l'IIR à l'apprentissage. Il faut cependant s'attendre à des difficultés techniques liées par exemple au fait que de nombreux essais doivent être réalisés pour que les résultats soient exploitables. En tout état de cause, un champ entier d'étude (initié par Wagner et al., 2000) reste à explorer.

Il existe une autre possibilité pour tester l'hypothèse de la récupération en phase d'étude sans modifier la tâche. Dans un agencement contractant à long terme, on peut considérer que l'épisode P2 survient trop tard dans le temps et que, par conséquent, cet agencement pâtirait du fait que les traces de P1 ne peuvent pas être récupérées lors de P2. On pourrait explorer cette hypothèse en comparant les performances mnésiques finales associées à un agencement contractant donné (par exemple X-----XX) avec celles d'un agencement identique mais ne comportant pas l'épisode P1 (par exemple -----XX). Si l'hypothèse est correcte, les deux conditions devraient mener à des performances identiques.

Enfin, une question intrigante est celle de l'efficacité relative de plusieurs agencements expansifs ; en effet, pour un même temps total d'étude et un même nombre de présentations, différents agencements expansifs peuvent être conçus, en fonction notamment de la durée du premier IIR. On pourrait par exemple comparer un agencement de type X-X-----X avec un agencement de type X---X-----X. En effet, il a été suggéré dans la littérature de l'agencement à court terme que ce serait principalement le premier intervalle qui détermine l'efficacité de l'agencement (e.g., Karpicke & Roediger, 2007). Cette hypothèse, en lien avec la théorie de la récupération en phase d'étude, pourrait alors être explorée sur une longue échelle de temps.

Malheureusement, l'exploration de ces effets est limitée par les contraintes techniques dues à la nécessité pour les participants de se rendre au laboratoire à plusieurs reprises. Une solution à cette contrainte peut être alors de faire réaliser les sessions via Internet (comme dans l'Expérience 1), ce qui permet au sujet de s'affranchir de déplacements contraignants.

6. 8. Applications potentielles

Les applications potentielles des effets observés sont assez évidentes. S'il s'avère que la supériorité de l'agencement expansif sur les autres types d'agencement sur le long terme est un effet robuste, alors la recommandation pourrait être faite aux pédagogues et aux apprenants de suivre un plan expansif pour les apprentissages. Les bénéfices observés par rapport aux autres agencements sont de l'ordre d'une dizaine de pour cent, alors que le temps total passé à l'étude est similaire. Il faut toutefois garder à l'esprit que l'effet qui été montré s'applique pour un apprentissage de type « par cœur » ; il est donc infondé de recommander de suivre un agencement expansif quand il s'agit par exemple d'assimiler un cours complexe nécessitant des efforts de compréhension, de synthèse, etc (e.g., Kornell & Bjork, 2008). Par ailleurs, adopter un plan d'apprentissage basé sur un agencement expansif des sessions de travail nécessite une certaine organisation, puisqu'il faut planifier à l'avance les jours où devront avoir lieu les sessions d'apprentissage. Il n'est cependant pas complètement indispensable que les sessions aient lieu exactement les jours prévus dans le planning. Il peut y avoir une certaine souplesse dans la programmation : on devrait pouvoir sans problème avancer ou reculer une session d'un jour, l'agencement restant tout de même expansif. L'agencement expansif présente également un intérêt certain lorsqu'on ne sait pas quand aura lieu le test de récupération. Ainsi, l'agencement expansif peut être poursuivi

à l'infini, et permet de réactiver les souvenirs après un intervalle toujours plus long, ce qui est intéressant si le besoin de récupération se manifeste de façon imprévisible. De plus, l'agencement expansif permet de nécessiter de moins en moins de sessions au fil du temps, ce qui pourrait permettre d'éviter une certaine lassitude due à des répétitions fréquentes, ainsi que de diminuer progressivement la durée de travail consacrée à un matériel donné. Cela n'est évidemment pas le cas dans la situation d'agencement contractant, pour laquelle il est nécessaire de savoir quand aura lieu le moment de la récupération. Ce n'est pas non plus le cas de la situation d'agencement uniforme car si la récupération se produit très tardivement, on est contraint de continuer les sessions d'apprentissage sur le même rythme pendant une durée indéfinie, aboutissant à un grand nombre de sessions. Enfin, adopter l'agencement expansif permet, au fur et à mesure du passage du temps et de la diminution de la fréquence des sessions, de laisser la place à l'introduction d'un nouveau matériel à apprendre, qui lui-même suivra un agencement expansif, permettant à son tour d'introduire un troisième type de matériel, et ainsi de suite.

Il est intéressant de constater que des logiciels sont commercialisés qui se basent sur le principe de la pratique distribuée à long terme, et qui par ailleurs se basent également sur les bénéfices associés à l'effet de test (Roediger & Karpicke, 2006). Citons par exemple l'application SuperMemo (<http://www.supermemo.com>) dans laquelle l'utilisateur peut introduire son propre matériel à apprendre, item par item, et n'a qu'à se laisser guider par le programme puisque celui-ci implémente un algorithme calculant le moment optimal auquel présenter ce matériel. La contrainte est, par ailleurs, de faire fonctionner l'application tous les jours. Le matériel est donc présenté lors des jours les plus adéquats, et ce sous forme de test avec feedback. Selon la réponse de l'utilisateur, le logiciel calcule la durée du prochain intervalle optimal avant la prochaine occurrence de l'item : en cas d'échec, la prochaine occurrence aura lieu rapidement, et en cas de succès, la prochaine occurrence est reportée à un délai plus long que le précédent intervalle. Par conséquent, l'agencement des présentations devrait être expansif dans la mesure où le succès de récupération devrait être garanti par le bon choix des intervalles.

Conclusion générale

En guise de conclusion, nous souhaitons mettre l'accent sur trois points. En premier lieu, nous pensons que la recherche a pour vocation finale de bénéficier à la société dite civile. En l'occurrence, les questions d'apprentissage et de mémoire concernent l'ensemble de la population, du jeune enfant à la personne âgée. Encore beaucoup de recherches sont nécessaires pour connaître davantage les rouages de la mémoire et donc pour savoir comment optimiser nos apprentissages. Nos résultats apportent une contribution infime à cette vaste question, en montrant que la mémorisation par cœur peut être améliorée sans coût, simplement en mettant en place une certaine organisation temporelle des sessions de travail. Or, il apparaît un cruel manque de transfert des connaissances issues de la recherche vers le monde éducatif ; par exemple, l'effet de pratique distribuée -qui n'est pourtant pas une découverte récente- n'est pas véritablement mis en pratique dans les classes, les livres scolaires, les sessions de formation, etc. D'autre part, il nous semble crucial de mettre davantage en œuvre des expérimentations de type écologique, qui bien que ne permettant pas un contrôle aussi strict des conditions expérimentales que les recherches en laboratoire, possèdent en contrepartie une validité externe et donc une applicabilité plus importante. Pour ce qui concerne l'étude de la mémoire, il nous semble essentiel d'inclure davantage la dimension « long terme » dans les expérimentations.

Bien entendu, l'optimisation est une chose, mais la compréhension des processus cognitifs à l'œuvre en est une autre, tout aussi essentielle. De façon générale, nos travaux nous ont conduit à penser qu'il est nécessaire de prendre en compte - et de favoriser - les phénomènes de récupération ayant lieu pendant l'apprentissage lui-même. Cela confirme l'idée qu'il est trompeur de considérer que la phase d'apprentissage consisterait uniquement en des processus d'encodage tandis que la phase de test consisterait uniquement en des processus de récupération. D'autre part, il nous semble indispensable d'explorer la dynamique des souvenirs sur le long terme, et notamment l'effet des répétitions sur cette dynamique. En particulier, l'exploration des effets de la durée du délai de rétention ainsi que l'étude de l'oubli nous semblent des pistes majeures et prometteuses.

Nous pensons enfin que relativement à l'étude de la mémoire, des rapprochements fertiles pourraient être davantage engagés entre les recherches en psychologie et en physiologie ; il apparaît en effet que chaque champ disciplinaire intègre difficilement les découvertes de l'autre champ. Nous en voulons pour preuve que certaines notions ont parfois une définition quelque peu différente en fonction des champs d'étude (e.g., la consolidation). Nous ne pouvons qu'espérer que l'avenir verra le développement d'une démarche fondamentalement pluridisciplinaire dans l'étude de la mémoire et, plus généralement, de la cognition.

Références

- Alberini, C. M. (2005). Mechanisms of memory stabilization: Are consolidation and reconsolidation similar or distinct processes? *Trends in Neurosciences*, 28(1), 51-56.
- Alvarez, P., Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences of the United States of America*, 91, 7041-7045.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036-1060.
- Anderson, J. R. & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96(4), 703-719.
- Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, 2(6), 396-408.
- Anderson, R. B., Tweney, R. D., Rivardo, M., & Duncan, S. (1997). Need probability affects retention: A direct demonstration. *Memory & Cognition*, 25(6), 867-872.
- Ausubel, D. P. (1966). Early versus delayed review in meaningful learning. *Psychology in the Schools*, 3(3), 195-198.
- Baird, H. P. (1979). Maintenance of knowledge: Questions about memory we forgot to ask. *Journal of Experimental Psychology: General*, 108(3), 296-308.
- Baird, H. P., Baird, L. E., Baird, A. S., & Baird, P. E. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science*, 4(5), 316-321.
- Baird, H. P., & Hall, L. H. (2005). The importance of retrieval failures to long term retention : A metacognitive explanation of the spacing effect. *Journal of Memory and Language*, 52, 566-577.
- Baird, H. P., & Phelps, E. (1987). Retention of Spanish vocabulary over eight years. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13, 344-349.
- Balota, D. A., Duchek, J. M., Sargent-Marshall, S. D., & Roediger, H. L. III. (2006). Does expanded retrieval produce benefits over equal-interval spacing ? Explorations of spacing effects in healthy aging and early stage Alzheimer's disease. *Psychology and Aging*, 21(1), 19-31.
- Bellezza, F. S., Winkler, H. B., & Andrasik, F. (1975). Encoding processes and the spacing effect. *Memory and Cognition*, 3(4), 451-457.
- Bellezza, F. S., & Young, D. R. (1989). Chunking of repeated events in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(5), 990-997.
- Benjamin, A. S. & Ross, B. H. (in press). The causes and consequences of reminding. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A Festschrift in honor of Robert A. Bjork*. New York, NY: Psychology Press.

- Benjamin, A. S., & Tullis, J. (2010). What makes distributed practice effective ? *Cognitive Psychology*, 61(3), 228-247.
- Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (Vol. 1, pp. 396 - 401). New York: Wiley.
- Bloom, K. C., & Shuell, T. J. (1981). Effects of massed and distributed practice on the learning and retention of a second-language vocabulary. *Journal of Educational Research*, 74(4), 245-248.
- Bonin, P., Méot, A., Aubert, L., Malardier, N., Niedenthal, P., & Capelle-Toczek, M.-C. (2003). Normes de concrétude, de valeur d'imagerie, de fréquence subjective et de valence émotionnelle pour 866 mots. *L'année Psychologique*, 104, 655-694.
- Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C: Winston.
- Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval, and time in working memory: Evidence from once-presented words. *Memory*, 6, 37-65.
- Brewer, J. B., Zhao, Z., Desmond, J. E., Glover, G. H., Gabrieli, J.D.E. (1998). Making memories: brain activity that predicts how well visual experience will be remembered. *Science*, 281, 1185-1187.
- Brown, M. W., & Xiang, J. Z. (1998). Recognition memory : Neuronal substrates of the judgement of prior occurrence. *Progress in Neurobiology*, 55, 149-189.
- Brozinsky, C., Yonelinas, A., Kroll, N., & Ranganath, C. (2005). Lag-sensitive repetition suppression effects in the anterior parahippocampal gyrus. *Hippocampus*, 15, 557–561.
- Bunch, M. E. (1941). The measurement of retention by the relearning method. *Psychological Review*, 48, 450-456.
- Burt, H. E., & Dobell, E. M. (1925). The curve of forgetting for advertising material. *Journal of Applied Psychology*, 9, 5-21.
- Callan, D. E., & Schweighofer, N. (2010). Neural correlates of the spacing effect in explicit verbal semantic encoding support the deficient-processing theory. *Human Brain Mapping* 31, 645-659.
- Carew, T. J. (1996). Molecular enhancement of memory formation. *Neuron*, 16, 5-8.
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name learning. *Applied Cognitive Psychology*, 19, 619-636.
- Cepeda, N. J., Coburn, N., Rohrer, D., Wixted, J. T., Mozer, M. C., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology*, 56(4), 236-246.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132(3), 354–380.
- Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridge of optimal retention. *Psychological Science*, 19(11), 1095-1102.

- Challis, B. H. (1993). Spacing effects on cued-memory tests depend on level of processing. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 19(2), 389-396.
- Cohen, J., McWhinney, B., Flatt, M., & Provost, J. (1993). Psyscope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioral Research Methods, Instruments and Computers*, 25, 257-271.
- Crowder, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- Cuddy, L. J., & Jacoby, L. L. (1982). When forgetting helps memory: An analysis of repetition effects. *Journal of Verbal Learning and Verbal Behavior*, 21, 451-467.
- Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, 14, 215-235.
- Cull, W. L., Shaughnessy, J. J., & Zechmeister, E. B. (1996). Expanding understanding of the expanding-pattern-of-retrieval mnemonic : Toward confidence in applicability. *Journal of Experimental Psychology: Applied*, 2, 365-378.
- De Bonis, M. (1968). La loi de Yerkes-Dobson : Problèmes méthodologiques liés à sa vérification. *L'année Psychologique*, 68(1), 121-141.
- Dempster, F. N. (1987a). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology*, 79, 162-170.
- Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, Vol. 11(2), 114-126.
- Donovan, J. J., & Radosovich, D. J. (1999). A meta-analytic review of the distribution of practice effect. *Journal of Applied Psychology*, 84, 795-805.
- Drosopoulos, S., Wagner, U., & Born, J. (2005). Sleep enhances explicit recollection in recognition memory. *Learning & Memory*, 12, 44-51.
- Ebbinghaus, H. (1885). *Memory: A Contribution to Experimental Psychology*, (translated by H.A. Ruger and C.E. Bussenius, 1913). New York: Teachers College, Columbia University.
- Ellenbogen, J. M., Hulbert, J. C., Jiang, Y., & Stickgold, R. (2009). The sleeping brain's influence on verbal memory: Boosting resistance to interference. *PLoS One*, 4(1).
- Ellenbogen, J. M., Hulbert, J. C., Stickgold, R., Dinges, D. F., Thompson-Schill, S. L. (2006). Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. *Current Biology*, 16, 1290-1294.
- Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review*, 62(5), 369-377.
- Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory and Cognition*, 13(4), 333-345.
- Forcato, C., Argibay, P. F., Pedreira, M. E., & Maldonado, H. (2009). Human reconsolidation does not always occur when a memory is retrieved: The relevance of the reminder structure. *Neurobiology of Learning and Memory*, 91, 50-57.
- Forcato, C., Burgos, V. L., Argibay, P. F., Molina, V. A., Pedreira, M. E., & Maldonado, H. (2007). Reconsolidation of declarative memory in humans. *Learning and Memory*, 14, 295-303.

- Forcato, C., Rodríguez, M. L. C., Pedreira, M. E., & Maldonado, H. (2010). Reconsolidation in humans opens up declarative memory to the entrance of new information. *Neurobiology of Learning and Memory* 93, 77-84.
- Frankland, P. W., & Bontempi, B. (2005). The organization of recent and remote memories. *Nature Review Neuroscience*, 6, 119-130.
- Gais, S., Lucas, B., Born J. (2006). Sleep after learning aids memory recall. *Learning and Memory*, 13, 259-262.
- Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. E. Tulving, E. Fergus, & I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). New York: Oxford University Press.
- Gay, L. R. (1973). Temporal position of reviews and its effect on the retention of mathematical rules. *Journal of Educational Psychology*, 64(2), 171-182.
- Genoux, D., Haditsch, U., Knobloch, M., Michalon, A., Storm, D., & Mansuy, I. M. (2002). Protein phosphatase 1 is a molecular constraint on learning and memory. *Nature*, 418, 970-975.
- Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, 15, 1–16.
- Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory and Cognition*, 7(2), 95-112.
- Glenberg, A. M., & Lehman, T. S. (1980). Spacing repetitions over one week. *Memory and Cognition*, 8(6), 528-538.
- Glenberg, A. M., & Smith, S. M. (1981). Spacing repetitions and solving problems are not the same. *Journal of Verbal Learning and Verbal Behavior*, 20, 110-119.
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66(3), 325-331.
- Greene, R. L. (1989). Spacing effects in memory: Evidence for a two-process account. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, 15(3), 371-377.
- Greene, R. L., & Stillwell, A. M. (1995). Effects of encoding variability and spacing on frequency discrimination. *Journal of Memory and Language*, 34, 468–476.
- Hardt, O., Orn Einarsson, E., & Nader, K. (2010). A bridge over troubled water: Reconsolidation as a link between cognitive and neuroscientific memory research traditions. *Annual Review of Psychology*, 61, 141-167.
- Henson, R. N. A. (2003). Neuroimaging studies of priming. *Progress in Neurobiology*, 70, 53-81.
- Henson, R. N., Rylands, A., Ross, E., Vuilleumier, P., & Rugg, M. D. (2004). The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *NeuroImage*, 21, 1674-1689.
- Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, 287, 1269-1272.

- Hintzman, D. L. (1974). Theoretical implications of the spacing effect. In R. L. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium* (pp. 77-99). Hillsdale, NJ: Erlbaum.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95(4), 528-551.
- Hintzman, D. L. (2010). How does repetition affect memory? Evidence from judgments of recency. *Memory & Cognition*, 38(1), 102-115.
- Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 88(3).
- Hintzman, D. L., & Block, R. A. (1973). Memory for the spacing of repetitions. *Journal of Experimental Psychology*, 99(1), 70-74.
- Hintzman, D. L., Block, R. A., & Summers, J. J. (1973). Modality tags and memory for repetitions: Locus of the spacing effect. *Journal of Verbal Learning and Verbal Behavior*, 12(3), 229-238.
- Hintzman, D. L., Summers, J. J., & Block, R. A. (1975). Spacing judgment as an index of study-phase retrieval. *Journal of Experimental Psychology: Human Learning and Memory*, 104(1), 31-40.
- Hintzman, D. L., Summers, J. J., Eki, N. T., & Moore, M. D. (1975). Voluntary attention and the spacing effect. *Memory & Cognition*, 3, 576-580.
- Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology* 46, 269-299.
- Howell, D. C. (2006). Méthodes statistiques en sciences humaines. *Éditions De Boeck Université*.
- Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4), 923-941.
- Hupbach, A., Hardt, O., Gomez, R., & Nadel, L. (2008). The dynamics of memory context-dependent updating. *Learning & Memory*, 15, 574-579.
- Hupbach, A., Gomez, R., Hardt, O., & Nadel, L. (2007). Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learning and Memory*, 14, 47-53.
- Hupbach, A., Gomez, R., & Nadel, L. (2009). Episodic memory reconsolidation: Updating or source confusion? *Memory*, 17(2), 502-510.
- Isarida, T., & Isarida, T. K. (2010). Effects of simple- and complex-place contexts in the multiple-context paradigm. *The Quarterly Journal Of Experimental Psychology*, 63(12), 2399 -2412.
- Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17(6), 649-667.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, Vol 110(3), 306-340.

- Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research*, 30, 138–149.
- Ji, D., & Wilson, M. A. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience* 10, 100-107.
- Johnston, W. A., Uhl, C. N. (1976). The contributions of encoding effort and variability to the spacing effect on free recall. *Journal of Experimental Psychology: Human Learning and Memory*, 2(2), 153-160.
- Karpicke, J. D., & Roediger, H. L. III. (2007). Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention. *Journal of Experimental Psychology: Learning Memory and Cognition*, 33(4), 704-719.
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, 19(6), 585-592.
- Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P.E. Morris and R. N. Sykes (Eds), *Practical aspects of memory* (pp. 625-632). London: Academic Press.
- Landauer, T. K. & Ross, B. H. (1977). Can simple instructions to use spaced practice improve ability to remember a fact? An experimental test using telephone numbers. *Bulletin of the Psychonomic Society*, 10, 215-218.
- Lattal, K. M., & Abel, T. (2004). Behavioral impairments caused by injections of the protein synthesis inhibitor anisomycin after contextual retrieval reverse with time. *Proceedings of the National Academy of Sciences of the United States of America*, 101(13), 4667- 4672.
- Lindsey, R., Mozer, M. C., Cepeda, N. J., & Pashler, H. (2009). Optimizing memory retention with cognitive models. In A. Howes, D. Peebles, & R. Cooper (Eds.), *Proceedings of the Ninth International Conference on Cognitive Modeling (ICCM)*. Manchester, UK.
- Litman, L., & Davachi, L. (2008). Distributed learning enhances relational memory consolidation. *Learning & Memory*, 15(9), 711-716.
- Logan, J. M., & Balota, D. A. (2008). Expanded vs. equal interval spaced retrieval practice: exploring different schedules of spacing and retention interval in younger and older adults. *Aging, Neuropsychology and Cognition*, 15 (3), 257-280.
- Lyon, D. O. (1914). The relation of length of material to time taken for learning and the optimum distribution of time. Part III. *Journal of Educational Psychology*, 5(3), 155-163.
- Mammarella, N., Avons, S., & Russo, R. (2004). A short-term perceptual priming account of spacing effects in explicit cued-memory tasks for unfamiliar stimuli. *European Journal of Cognitive Psychology*, 16, 387-402.
- Mammarella, N., Russo, R., & Avons, S. (2002). Spacing effects in cued-memory tasks for unfamiliar faces and nonwords. *Memory & Cognition*, 30, 1238-1251.
- Maquet, P., Laureys, S., Peigneux, P., Fuchs, S., Petiau, C., Phillips, C., Aerts, J., Del Fiore, G., Degueldre, C, Meulemans, T., Luxen, A., Franck, G., Van Der Linden,

- M., Smith, C., & Cleeremans, A. (2000). Experience-dependent changes in cerebral activation during human REM sleep. *Nature Neuroscience*, 3(8), 831-836.
- Martin, E. (1967). Relation between stimulus recognition and paired-associate learning. *Journal of Experimental Psychology*, 74(4), 500-505.
- Martin, E. (1968). Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review*, 75(5), 421-441.
- Maskarinec, A. S., & Thompson, C. P. (1976). The within-list distributed practice effect: Tests of varied context and varied encoding hypotheses. *Memory & Cognition*, 4, 741-747.
- Maviel, T., Durkin, T. P., Menzaghi, F., & Bontempi, B. (2004). Sites of neocortical reorganization critical for remote spatial memory. *Science*, 305, 96-99.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419-457.
- Melton, A. W. (1967). Repetition and retrieval from memory. *Science*, 158(3800), 532.
- Melton, A. W. (1970). The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior*, 9(5), 596-606.
- Morris, C. D., Bransford, J. D., Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, 16(5), 519-533.
- Mozer, M. C., Pashler, H., Cepeda, N., Lindsey, R., & Vul, E. (2009). Predicting the optimal spacing of study: A multiscale context model of memory. In Y. Bengio, D. Schuurmans, J. Lafferty, C.K.I. Williams and A. Culotta (Eds), *Advances in neural information processing systems* (Vol. 22, pp. 1321-1329). NIPS Foundation, La Jolla, CA.
- Nader, K., Schafe, G. E., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, 406, 722-726.
- Pagani, M. R., Oishi, K., Gelb, B. D., & Zhong, Y. (2009). The phosphatase SHP2 regulates the spacing effect for Long-Term Memory induction. *Cell*, 139, 186-198.
- Pashler, H., Zarow, G., & Triplett, B. (2003). Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1051-1057.
- Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science: A Multidisciplinary Journal*, 29(4), 559-586.
- Peigneux, P., Laureys, S., Fuchs, S., Collette, F., Perrin, F., Reggers, J., Phillips, C., Degueldre, C., Del Fio, G., Aerts, J., Luxen, A., & Maquet, P. (2004). Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron*, 44, 535-545.
- Peterson, H. A., Ellis, M., Toohill, N., & Kloess, P. (1935). Some measurement of the effects of reviews. *Journal of Educational Psychology*, 26, 65-72.

- Postman, L., & Knecht, K. (1983). Encoding variability and retention. *Journal of Verbal Learning and Verbal Behavior*, 22, 133-152.
- Raaijmakers, J. G. W. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science*, 27, 431-452.
- Reynolds, J. H., Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology*, 55(5), 297-308.
- Robinson, E. S. (1921). The relative efficiencies of distributed and concentrated study in memorizing. *Journal of Experimental Psychology*, 4, 327-343.
- Roediger, H. L. III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181-210.
- Rohrer, D., Taylor, K. (2006). The effects of overlearning and distributed practise on the retention of mathematics knowledge. *Applied Cognitive Psychology*, 20, 1209-1224-374.
- Ross, B. H., & Landauer, T. K. (1978). Memory for at least one of two items: Test and failure of several theories of spacing effects. *Journal of Verbal Learning and Verbal Behavior*, 17, 669-676.
- Russo, R., Mammarella, N., & Avons, S. (2002). Toward a unified account of spacing effects in explicit cued-memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 819-829.
- Russo, R., Parkin, A. J., Taylor, S. R., & Wilks, J. (1998). Revising current two-process accounts of spacing effects in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 161-172.
- Sara S. J. (2000). Retrieval and Reconsolidation: Toward a Neurobiology of Remembering. *Learning and Memory*, 7, 73-84.
- Scharf, M. T., Woo, N. H., Lattal, K. M., Young, J. Z., Nguyen, P. V., & Abel, T. (2002). Protein synthesis is required for the enhancement of long-term potentiation and long-term memory by spaced training. *Journal of Neurophysiology*, 87, 2770-2777.
- Seabrook, R., Brown, G. D. A., & Solity, J. E. (2005). Distributed and massed practise: From laboratory to classroom. *Applied Cognitive Psychology*, 19, 107-122.
- Shaughnessy, J. J. (1976). Persistence of the spacing effect in free recall under varying incidental learning conditions. *Memory & Cognition*, 4(4), 369-377.
- Sisti, H. M., Glass, A. L., & Shors, T. J. (2007). Neurogenesis and the spacing effect: Learning over time enhances memory and the survival of new neurons. *Learning and Memory*, 14, 368-375.
- Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, 5(5), 460-471.
- Smith, S. M. (1995). Mood Is a Component of Mental Context: Comment on Eich (1995). *Journal of Experimental Psychology: General*, 124(3), 309-310.
- Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition*, 6(4), 342-353.

- Sones, A. M., & Stroud, J. B. (1940). Review, with special reference to temporal position. *Journal of Educational Psychology*, 31, 665-676.
- Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, 8(2), 203-220.
- Squire, L. R. & Kandel, E. R. (2002). La mémoire : De l'esprit aux molécules. *DeBoeck Université*.
- Staddon, J. E. R., Chelaru, I. M., & Higa, J. J. (2002). Habituation, memory and the brain: The dynamics of interval timing. *Behavioural Processes*, 57, 71-88.
- Starch, D. (1912). Periods of work in learning. *Journal of Educational Psychology*, 3(4), 209-213.
- Stark, S. M., Gordon, B., & Stark, C. E. L. (2008). Does the presence of priming hinder subsequent recognition or recall performance? *Memory*, 16(2), 157-173.
- Strong, E. K. (1916). The factors affecting a permanent impression developed through repetition. *Journal of Experimental Psychology*, 1, 319-338.
- Takashima, A., Nieuwenhuis, I. L. C., Rijpkema, M., Petersson, K. M., Jensen, O. & Fernández, G. (2007). Memory trace stabilization leads to large-scale changes in the retrieval network: A functional MRI study on associative memory. *Learning & Memory*, 14, 472-479.
- Thios, S. J., D'Agostino, P. R. (1976). Effects of repetition as a function of study-phase retrieval. *Journal of Verbal Learning and Verbal Behavior*, 15(5), 529-536.
- Tononi, G., & Cirelli C. (2006). Sleep function and synaptic homeostasis. *Sleep Medicine Review*, 10(1), 49-62.
- Toppino, T. C. (1991). The spacing effect in young children's free recall: Support for automatic-process explanations. *Memory & Cognition*, 19(2), 159-167.
- Toppino, T. C., & Bloom, L. C. (2002). The spacing effect, free recall, and two-process theory: A closer look. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 437-444.
- Toppino, T. C., Fearnow-Kenney, M. D., Kiepert, M. H., & Teremula, A. C. (2009). The spacing effect in intentional and incidental free recall by children and adults: Limits on the automaticity hypothesis. *Memory & Cognition*, 37(3), 316-325.
- Tsai, L.-S. (1927). The relation of retention to the distribution of relearning. *Journal of Experimental Psychology*, 10, 30-39.
- Tulving, E., Hayman, C. A., & Macdonald, C. (1991). A Long-lasting perceptual priming and semantic learning in amnesia: A case experiment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4), 595-617.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352-373.
- Tzeng, O. J. L., & Cotton, B. (1980). A study-phase retrieval model of temporal coding. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 705-718.
- Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Schmidt, H. G. (2004). Detrimental influence of contextual change on spacing effects in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(4), 796-800.

- Verkoeijen, P. P. J. L., Rikers, R. M. J. P., & Schmidt, H. G. (2005). Limitations to the spacing effect: Demonstration of an inverted u-shaped relationship between interrepetition spacing and free recall. *Experimental Psychology*, 52(4), 257-263.
- Wagner, A., Dale, A. M., Rosen, B. R., & Buckner, R. L. (1998). Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281, 1188-1191.
- Wagner, A. D., Maril, A., & Schacter, D. L. (2000). Interactions between forms of memory: When priming hinders new episodic learning. *Journal of Cognitive Neuroscience*, 12:Supplement 2, 52-60.
- Weigold, A. (2008). *The Relationship Between Restudying and Testing in the Short and Long Term*. Unpublished doctoral dissertation, Texas Tech University.
- Wheeler, M. A., Ewers, M., Buonanno, J. F. (2003). Different rates of forgetting following study versus test trials. *Memory*, 11(6), 571-580
- Whittlesea, B. W. A., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. *Journal of Memory and Language*, 29, 716-732.
- Whittlesea, B. W. A., & Leboe, J. P. (2003). Two fluency heuristics (and how to tell them apart). *Journal of Memory and Language*, 49, 62-79.
- Wickelgren, W. A. (1973). Trace resistance and the decay of long-term memory. *Journal of Mathematical Psychology*, 10, 418-455.
- Winograd, E., & Soloway, R. M. (1985). Reminders as a basis for temporal judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 262-271.
- Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review Psychology*, 55, 235-269.
- Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science*, 18(2), 133-134.
- Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, 2(6), 409-415.
- Wu, G. Y., Deisseroth, K., & Tsien, R. W. (2001). Spaced stimuli stabilize MAPK pathway activation and its effects on dendritic morphology. *Nature Neuroscience*, 4(2), 151-157.
- Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R. A., & Dong, Q. (2010a). Facilitating memory for novel characters by reducing neural repetition suppression in the left fusiform cortex. *PLoS ONE*, 5(10), e13204.
- Xue, G., Mei, L., Chen, C., Lu, Z.-L., Poldrack, R. A., & Dong, Q. (2010b). Spaced learning enhances subsequent recognition memory by reducing neural repetition suppression. *Journal of Cognitive Neuroscience*, 23(7), 1624-1633.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, 46, 441-517.
- Young, D. R., & Bellezza, F. S. (1982). Encoding variability memory organization, and the repetition effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 545-559.

Notes bibliographiques

- Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R.L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Hillsdale, NJ: Erlbaum.
- Chee, M., Westphal, C., Goh, J., Graham, S., & Song, A. (2003). Word frequency and subsequent memory effects studied using event-related fMRI. *Neuroimage*, 20, 1042-1051.
- Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science*, 294, 1030-1038.
- Lee, T. D., & Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport*, 59, 277–287.
- Madsen, M. C. (1963), Distribution of practice and level of intelligence. *Psychological Reports*, 13, 39-42.
- Otten, L., Henson, R., & Rugg, M. (2001). Depth of processing effects on neural correlates of memory encoding: Relationship between findings from across- and within-task comparisons. *Brain*, 124, 399.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, pp. 207–262). New York: Academic Press.

Annexes

Annexe n°1 : Caractéristiques des mots utilisés au sein des paires, selon Bonin et al. (2003).

	Concrétude	Imagerie	Fréquence subjective	Valence émotionnelle
Vache	4,86	4,80	3,64	3,44
Souris	4,77	4,76	3,16	2,88
Cheval	4,82	4,76	3,52	3,72
Chien	4,77	4,92	4,52	3,96
Chat	4,91	4,96	4,24	3,96
Oiseau	4,77	4,56	3,96	3,68
Bras	4,73	4,76	4,44	3,40
Main	4,82	4,72	4,60	4,28
Pied	4,73	4,68	4,48	3,40
Nez	4,82	4,80	4,28	2,80
Jambe	4,73	4,60	4,36	3,28
Oeil	4,68	4,92	4,44	3,72
Stylo	4,91	4,84	4,68	3,60
Lampe	4,81	4,44	4,04	3,48
Réfrigérateur	4,77	4,56	3,08	3,44
Chaise	4,91	4,72	4,60	3,12
Téléviseur	4,64	4,92	4,88	3,76
Table	4,95	4,68	4,64	3,56
Fraise	4,91	4,92	3,72	4,16
Raisin	4,86	4,88	3,24	3,88
Poire	4,91	4,76	3,24	3,64
Banane	4,91	4,92	3,48	3,80
Abricot	4,82	4,76	3,08	4,00
Pomme	4,86	4,92	4,52	3,96
Piano	4,91	4,76	3,00	3,88
Violon	5,00	4,84	2,28	3,48
Guitare	4,91	5,00	3,00	3,92
Trompette	4,86	4,76	2,20	3,04
Tambour	4,91	4,12	2,48	2,92
Accordéon	4,86	4,68	2,04	2,20

Les mots issus des listes 1, 2, et 3 sont représentés sur des fonds de couleur bleue, verte, et jaune, respectivement. Les 5 catégories sémantiques (animal, partie du corps, objet, fruit, et instrument de musique), sont présentées successivement.

Annexe n°2 : Liste des paires distractrices utilisées dans les Expériences 1 et 2

Disciplines		Légumes		Animaux	
MATHEMATIQUES	DEGARA	TOMATE	LOCISE	OURS	SEVULE
CHIMIE	PEBERI	HARICOT	DOCOVA	LOUP	SAROGA
GEOGRAPHIE	BINIRU	COURGE	JISOPE	ARAIGNEE	VARIFA
HISTOIRE	VISOVA	COMCOMBRE	GITINA	SERPENT	PIDUBE
ECONOMIE	FUGINE	POIREAU	DAFITE	LAPIN	PEVAFE
LITTERATURE	CIDOVU	SALADE	GEJOVI	POISSON	BERICE
Qualités		Éléments		Fruits	
JEUNESSE	VUGUPO	SOLEIL	GUTOCO	KIWI	TUSUMO
VIEILLESSE	NEJABU	LUNE	FEJUBO	ANANAS	TALOCI
BEAUTE	COTINI	NUAGE	GUMOPE	MANGUE	LAMETE
SAGESSE	FEPIPO	ETOILE	CINIGA	CERISE	BELEJA
PATIENCE	GAPAGI	PLUIE	JAVICI	PRUNE	ROPULO
OPTIMISME	SEMOTE	NEIGE	MUVOSO	FIGUE	LONEPU
Principes		Instruments		Objets	
JUSTICE	GOGIDE	VIOLON	CEGEFI	TELEPHONE	BILONA
LIBERTE	NICARU	PIANO	LAVODI	REVEIL	TOLIRO
EGALITE	PIBOGE	TROMPETTE	BUBISU	BUREAU	NOGAFU
FRATERNITE	LIVEBI	GITARE	LADIFI	PLACARD	BITOPE
DIGNITE	FOPILU	ACCORDEON	NUMARE	TABOURET	JAJAGU
HONNEUR	DITALE	HARPE	MATIRO	VASE	ROPAMA
Sentiments		Bâtiments		Corps	
PEUR	MAPOTI	MAISON	LEGEGI	DOIGT	CASICE
JOIE	BOMARI	IMMEUBLE	TEMASI	CHEVEU	SINOLI
TRISTESSE	PUJAGI	MUSEE	VIPATU	POIGNET	GERILA
SURPRISE	FATAMI	CHATEAU	JOSEBA	GENOU	POMAVU
COLERE	MACETU	FERME	CAPIRO	COU	CAPEGA
DOULEUR	CULIGE	GARAGE	NAFORE	OREILLE	RETAFA

Annexe n°3 : Feuille de route de l'Expérience 1.

Feuille de route – Groupe A

Vous allez devoir vous connecter au site du laboratoire EMC tous les jours pendant 7 jours consécutifs. Il vous faudra donc planifier vos séances sur la semaine afin qu'elles aient lieu à la même heure du jour dans la mesure du possible, toujours sur le même ordinateur (avec une connexion haut débit) et au même endroit. Il est très important que vous **ne ratiez aucune séance**, si c'est le cas contactez-moi.

Pendant les 15 minutes de séance quotidienne, assurez-vous que vous serez **dans le calme**, que vous pourrez vous concentrer sans interruption (coupez votre téléphone, prévenez votre entourage...). Munissez-vous de vos lunettes si vous en portez. Installez-vous confortablement à une distance de votre écran d'ordinateur qui vous permet de le lire (50 cm).

Entre chaque séance quotidienne, **ne repensez pas à ce que vous avez appris**, ne révisez pas, n'en parlez pas à votre entourage. Si des souvenirs vous viennent, essayez de penser à autre chose. C'est très important pour la recherche que nous menons, ça peut vraiment rendre vos résultats inutilisables. Le but de recherche est d'explorer comment fonctionne la mémoire, faites du mieux que vous pouvez pour apprendre les mots et répondre au test mais ne trichez pas, ça ne sert à rien car vous ne serez pas « noté ».

L'adresse Internet du site : <http://unpc.univ-lyon2.fr/>

Cliquer sur l'icône **Accueil** (en rouge)

Descendez tout en bas de la page vers le paragraphe « Participation à une expérience sur l'apprentissage d'un vocabulaire imaginaire »

Cliquer sur la case **Groupe A**.

Entrez le login suivant : **xxxxxx**

Et le mot de passe : **xxxxx**

Vous pouvez relire les consignes.

Et vous commencez l'apprentissage en cliquant dans la case correspondant à la session du jour. (Les séances doivent être **toutes** suivies, et **dans l'ordre**).

A la fin de la séance, notez dans le tableau suivant le jour, l'heure de fin de session et le **code** apparaissant à la fin de la session ainsi que vos commentaires éventuels.

	Date	Heure prévue	Heure réelle	Code de fin de session	Commentaires?
Session 1					
Session 2					
Session 3					
Session 4					
Session 5					
Session 6					
Session 7					
Test au laboratoire					

Consignes :

- Trouver du sens dans le pseudomot, à rapprocher du mot associé
- Fabriquer une image mentale ou fabriquer une phrase de cette association (Souvenez-vous de l'exemple de TABULORI – CAFETIERE)

Coordonnées de l'expérimentatrice (n'hésitez pas à me téléphoner en cas de problème) :

Emilie Gerbier

Master 2 Sciences cognitives – Lyon 2

Emilie.Gerbier@univ-lyon2.fr

06 XX XX XX XX

Annexe n°4 : Items présentés aux participants du groupe A dans l'Expérience 1

Session							
	1	2	3	4	5	6	7
1	LAMPE	CONCOMBRE	SURPRISE	KIWI	SAGESSE	PLUIE	POMME
2	PIED	IMMEUBLE	MAIN	POIGNET	TELEVISEUR	JEUNESSE	VACHE
3	OISEAU	CHEVAL	POIREAU	REVEIL	CHIMIE	ABRICOT	TABLE
4	CHAISE	SOLEIL	BRAS	OPTIMISME	SOURIS	OURS	MAIN
5	JAMBE	REFRIGERATEUR	REFRIGERATEUR	ACCORDEON	POMME	TABLE	OISEAU
6	STYLO	DOULEUR	LITTERATURE	TABOURET	VACHE	ECONOMIE	CHAISE
7	ABRICOT	NEZ	SOURIS	LOUP	JAMBE	OISEAU	BRAS
8	CHEVAL	PIANO	BANANE	HISTOIRE	ANANAS	VIEILLESSE	POIRE
9	MAIN	POIRE	GUITARE	JOIE	LAMPE	POMME	LAMPE
10	REFRIGERATEUR	NEIGE	CHAISE	TROMPETTE	OEIL	GENOU	ABRICOT
11	OEIL	FERME	FRAISE	COURGE	BEAUTE	POISSON	SOURIS
12	SOURIS	ARaignee	PIED	OREILLE	MAIN	TELEVISEUR	NEZ
13	BANANE	MANGUE	VACHE	GARAGE	CHAT	DIGNITE	CHAT
14	TABLE	CHAISE	JUSTICE	GEOGRAPHIE	RAISIN	CHÂTEAU	RAISIN
15	RAISIN	PEUR	CHIEN	BUREAU	MATHEMATIQUES	ETOILE	PIED
16	NEZ	DOIGT	STYLO	HONNEUR	ABRICOT	FIGUE	TELEVISEUR
17	FRAISE	VIOLOU	CHEVAL	HARICOT	BRAS	CHAT	CHEVAL
18	TELEVISEUR	CHIEN	LAMPE	COLERE	HARPE	PLACARD	OEIL
19	CHIEN	VASE	POIRE	SERPENT	TABLE	JAMBE	STYLO
20	POMME	PIED	COU	CHEVEU	FRAISE	CERISE	BANANE
21	VACHE	FRATERNITE	RAISIN	PATIENCE	OISEAU	EGALITE	JAMBE
22	POIRE	LAPIN	NUAGE	LIBERTE	MAISON	MUSEE	CHIEN
23	BRAS	BANANE	NEZ	SALADE	STYLO	OEIL	FRAISE
24	CHAT	TRISTESSE	PRUNE	TELEPHONE	TOMATE	LUNE	REFRIGERATEUR

Les items de l'agencement uniforme sont écrits sur fond orange, ceux de l'expansif sur fond bleu, ceux du contractant sur fond jaune. Seuls les mots de chaque paire sont représentés. La position sérielle est indiquée dans la première colonne.

Annexe n°5 : Items présentés aux participants du groupe A dans l'Expérience 2

	Session						
	1	2	3	4	5	6	7
1	NEZ	CONCOMBRE	SURPRISE	KWI	SAGESSE	PLUIE	POMME
2	SOURIS	IMMEUBLE	MAIN	POIGNET	CHAT	JEUNESSE	CHEVAL
3	ABRICOT	CHAISE	REFRIGERATEUR	REVEIL	CHIMIE	OISEAU	LAMPE
4	CHIEN	SOLEIL	POIREAU	OPTIMISME	RAISIN	ECONOMIE	JAMBE
5	LAMPE	REFRIGERATEUR	PIED	ACCORDEON	TELEVISEUR	TELEVISEUR	REFRIGERATEUR
6	POMME	DOULEUR	LAMPE	TABCURET	ANANAS	OURS	OISEAU
7	STYLO	PIANO	POIRE	LOUP	STYLO	VIEILLESSE	BANANE
8	CHEVAL	NEIGE	LITTERATURE	HISTOIRE	ABRICOT	TABLE	TABLE
9	OEIL	NEZ	SOURIS	JOIE	SOURIS	GENOU	MAIN
10	CHAISE	MANGUE	BANANE	TROMPETTE	JAMBE	POISSON	POIRE
11	MAIN	PEUR	BRAS	COURGE	BEAUTE	JAMBE	BRAS
12	BANANE	CHIEN	CHEVAL	OREILLE	MAIN	DIGNITE	ABRICOT
13	OISEAU	DOIGT	GUIARE	GARAGE	MATHEMATIQUES	POMME	NEZ
14	FRAISE	POIRE	RAISIN	GEOGRAPHIE	VACHE	CHATEAU	VACHE
15	CHAT	VIOLOW	CHAISE	BUREAU	POMME	CHAT	OEIL
16	RAISIN	PIED	JUSTICE	HONNEUR	LAMPE	CERISE	FRAISE
17	PIED	FERME	FRAISE	HARICOT	OEIL	EGALITE	CHIEN
18	TABLE	ARaignEE	NEZ	COLERE	HARPE	ABRICOT	TELEVISEUR
19	BRAS	VASE	VACHE	SERPENT	TOMATE	ETOILE	SOURIS
20	REFRIGERATEUR	CHEVAL	COU	CHEVEU	OISEAU	FIGUE	CHAISE
21	JAMBE	FRATERNITE	STYLO	PATIENCE	BRAS	PLACARD	CHAT
22	POIRE	BANANE	NUAGE	LIBERTE	TABLE	OEIL	STYLO
23	VACHE	LAPIN	CHIEN	SALADE	FRAISE	MUSEE	PIED
24	TELEVISEUR	TRISTESSE	PRUNE	TELEPHONE	MAISON	LUNE	RAISIN

Les items de l'agencement uniforme sont écrits sur fond orange, ceux de l'expansif sur fond bleu, ceux du contractant sur fond jaune. Seuls les mots de chaque paire sont représentés. La position sérielle est indiquée dans la première colonne.

Annexe n°6 : Items présentés aux participants du groupe A dans l'Expérience 3

	Jour				
	1	2	7	12	13
1	SOLEIL	TABOURET	HAUTOBOIS	ARAIGNEE	GENOU
2	ANANAS	SURPRISE	FRATERNITE	NUAGE	CONCOMBRE
3	OISEAU	IMMEUBLE	MANGUE	CYMBALE	NEZ
4	PIANO	GUIARE	VACHE	OEIL	LAMPE
5	REFRIGERATEUR	JUSTICE	HONNETETE	FERME	TAMBOUR
6	TAMBOUR	PIED	SOURIS	CHAT	MAIN
7	BRAS	POIREAU	COU	GEOGRAPHIE	BANANE
8	CHIEN	ETOILE	VIOLON	SALADE	SOURIS
9	TELEVISEUR	POIRE	PEUR	POISSON	JAMBE
10	FRAISE	VASE	STYLO	TABLE	CHEVAL
11	NEZ	CHIEN	HARICOT	ECONOMIE	FRAISE
12	VIOLON	LITTERATURE	REVEIL	ACCORDEON	OEIL
13	BANANE	OURS	BEAUTE	OPTIMISME	REFRIGERATEUR
14	ACCORDEON	CLAVECIN	PIANO	DOIGT	ABRICOT
15	STYLO	NEZ	JOIE	CHATEAU	VIOLON
16	POIRE	PRUNE	BRAS	TAMBOUR	TABLE
17	TABLE	MUSEE	SERPENT	TOMATE	CHIEN
18	SOURIS	CHAISE	HISTOIRE	OISEAU	POMME
19	JAMBE	BASSON	FRAISE	HARPE	VACHE
20	LAMPE	TELEPHONE	LUNE	PLACARD	TELEVISEUR
21	PIED	LIBERTE	GARAGE	POMME	TROMPETTE
22	CHAT	CHEVAL	FIGUE	POIGNET	BRAS
23	GUIARE	SAGESSE	DOULEUR	BUREAU	CHAISE
24	OEIL	NEIGE	LAPIN	COLERE	CHAT
25	VACHE	BANANE	RAISIN	JAMBE	PIANO
26	POMME	DIGNITE	HONNEUR	EGALITE	PIED
27	CHEVAL	MAISON	MAIN	ABRICOT	OISEAU
28	RAISIN	TROMPETTE	JEUNESSE	CHEVEU	POIRE
29	TROMPETTE	PATIENCE	COURGE	CHIMIE	STYLO
30	ABRICOT	REFRIGERATEUR	LAMPE	CERISE	GUIARE
31	CHAISE	FLUTE	KIWI	TELEVISEUR	RAISIN
32	MAIN	VIEILLESSE	OREILLE	PLUIE	ACCORDEON

Les items de l'agencement uniforme sont écrits sur fond orange, ceux de l'expansif sur fond bleu, ceux du contractant sur fond jaune. Seuls les mots de chaque paire sont représentés. La position sérielle est indiquée dans la première colonne.

Annexe n°7 : Items présentés dans l'Expérience 4

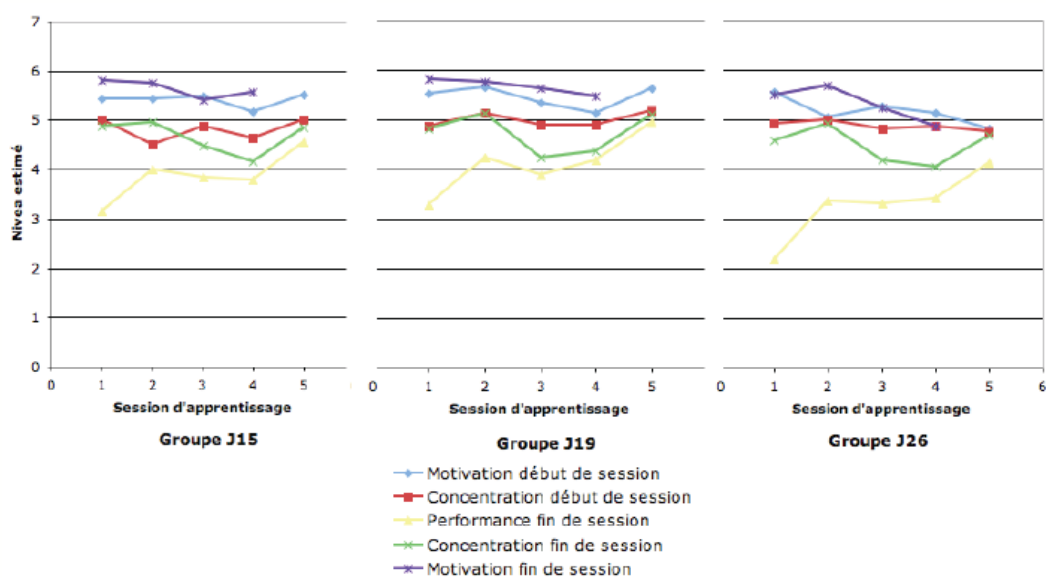
	P1			P2		Test
1	BILONA	TOMATE		MAPOTI	OREILLE	JAMIDO
2	NEJABU	TELEPHONE		FEJUBO	POISSON	MECUSI
3	COBEMI	STYLO		VUMOGI	FRAISE	BAFISO
4	POCUGA	CHIEN		FIRECO	BRAS	TAGOBU
5	TODEGI	JAMBE		CAJESI	VACHE	FASOLU
6	GEFIPA	RAISIN		DERUPA	OEIL	TODEGI
7	FASOLU	CHEVAL		MECUSI	LAMPE	SONIJA
8	DAPIGU	TABLE		POCUGA	CHIEN	DAPIGU
9	JAMIDO	SOURIS		COBEMI	STYLO	DERUPA
10	LATOMU	ABRICOT		GEFIPA	RAISIN	CAJESI
11	VAMELO	TELEVISEUR		TODEGI	JAMBE	LATOMU
12	LUTISA	PIED		FASOLU	CHEVAL	PUSIRO
13	VUMOGI	FRAISE		SONIJA	POMME	GEFIPA
14	CAJESI	VACHE		PUSIRO	CHASSE	VAMELO
15	FIRECO	BRAS		TAGOBU	MAIN	FIRECO
16	MECUSI	LAMPE		RUTOLA	OISEAU	RUTOLA
17	DERUPA	OEIL		BAFISO	POIRE	VUMOGI
18	BAFISO	POIRE		VAMELO	TELEVISEUR	COBEMI
19	RUTOLA	OISEAU		LATOMU	ABRICOT	LUTISA
20	TAGOBU	MAIN		JAMIDO	SOURIS	POCUGA
21	PUSIRO	CHASSE		DAPIGU	TABLE	
22	SONIJA	POMME		LUTISA	PIED	
23	FEJUBO	POISSON		NEJABU	TELEPHONE	
24	MAPOTI	OREILLE		BILONA	TOMATE	

Les items correspondant à des catégories sémantiques différentes sont écrits sur des fonds de couleur différente. Les items inscrits sur un fond gris sont les items de contrôle des effets de primauté et de récence.

Annexe n°8 : Formulaire utilisé dans l'Expérience 2

SESSION n° 3								
NOM et Prénom _____								
Estimez votre niveau de motivation pour l'apprentissage au début de la session :								
Pas du tout motivé(e)	1	2	3	4	5	6	7	Extrêmement motivé(e)
Jugement de reconnaissance des paires mot-pseudomot (entourer la réponse) :								
	Vu	Pas vu				PS	MS	S
Paire n°1 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°2 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°3 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°4 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°5 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°6 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°7 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°8 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°9 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°10 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°11 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°12 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°13 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°14 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°15 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°16 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°17 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°18 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°19 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°20 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°21 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°22 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°23 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paire n°24 :	<input type="checkbox"/>	<input type="checkbox"/>	Degré de confiance :	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Remarques éventuelles sur la session :								
.....								
.....								
Estimez votre motivation pour poursuivre l'expérience les jours prochains:								
Pas du tout motivé(e)	1	2	3	4	5	6	7	Extrêmement motivé(e)

Annexe n°9 : Scores de motivation / concentration / performance en fonction des groupes de DR dans l'Expérience 3



Annexe n°10 : Questionnaire de débriefing de l'Expérience 4.

NOM

Avez-vous trouvé cette expérience plutôt facile ou difficile ? Pourquoi ?

Avez-vous utilisé des stratégies d'apprentissage particulières ? OUI NON
Si oui, lesquelles ?

Pensez-vous avoir utilisé (au moins) une stratégie d'apprentissage pour :

☐ 100 % des paires
☐ 80 % des paires
☐ 60 % des paires
☐ 40 % des paires
☐ 20 % des paires
☐ 0 % des paires

Les stratégies ont-elles été utilisées à la première et à la deuxième phase d'apprentissage ? Ou uniquement lors d'une phase ? Laquelle ?

Avez-vous tenté de réviser les paires pendant la période entre les deux phases d'apprentissage (phase 2) et/ou pendant la période entre la deuxième phase d'apprentissage et l'évaluation (phase 4) ?
OUI NON

Des éléments que vous avez appris sont-ils revenus à votre esprit pendant la période entre les deux phases d'apprentissage (phase 2) ou pendant la période entre la deuxième phase d'apprentissage et l'évaluation (phase 4) ?
OUI NON

Si oui, détaillez votre réponse :

Pensez-vous que vos activités durant la période entre les deux phases d'apprentissage (phase 2) ont pu influencer vos performances finales ?
OUI NON

Si oui, expliquer comment :

Vous êtes vous ennuyé(e) à un moment donné de l'expérience ?

Avez-vous apprécié de participer à cette expérience ?
☐ Oui, beaucoup ☐ Oui, un peu ☐ Non, pas trop ☐ Non, pas du tout

Avez-vous des remarques particulières ?

Annexe n°11 : Détails sur les activités réalisées dans l'Expérience 4

Salle C (Fluctuation Faible) :

Les participants étaient assis dans la salle C autour d'une table carrée. Ils devaient réaliser des sudokus comme ils le souhaitaient pendant environ 30 minutes. Quatre grilles de sudoku, de niveau enfant, facile, intermédiaire et difficile, étaient disponibles sur la même feuille, avec leurs corrections sur la deuxième page. Ils ne devaient pas parler entre eux. J.F. était responsable de ce groupe.

Salle B (Fluctuation Forte) :

Les participants étaient réunis dans la salle B et réalisaient un ensemble d'activités dirigées par les expérimentatrices C.S. et E.G. Chaque activité durait environ 3 minutes et les activités réalisées de façon assise et debout étaient alternées.

- La première activité consistait en un ensemble de mouvements sur place, de la tête et des bras essentiellement, réalisés debout, sous la direction de C.S.
- La deuxième activité consistait en un jeu collectif par groupe de 2, 3 ou 4 participants consistant à relier des points sur une grille à tour de rôle dans le but de fermer un carré. Cette activité était réalisée assis.
- La troisième activité consistait à résoudre collectivement un puzzle par groupe de 2, 3 ou 4 participants. Cette activité était réalisée debout.
- La quatrième activité consistait à calculer individuellement des opérations de multiplication à deux nombres de 4, 5 ou 6 chiffres. Six opérations étaient proposées sur une feuille.
- La cinquième activité consistait pour les participants à suivre et imiter l'expérimentatrice C.S. qui sortait dans les couloirs et réalisait un parcours dans lequel il fallait bouger les bras, monter les escaliers à l'envers, sauter à pieds joints, etc.
- La sixième activité consistait à dessiner un trait continu sur une feuille blanche en écoutant de la musique et en laissant libre cours à son imagination. La consigne était de ne pas lever la mine du crayon de la feuille pendant toute la durée de l'activité. Cette activité était réalisée assis et de façon individuelle.
- La septième activité consistait en un jeu collectif de mime. Des noms de métiers étaient tirés au sort par les participants à tour de rôle qui devaient mimer le métier aux autres participants pour le leur faire deviner. Cette activité était réalisée debout.
- La huitième activité consistait en un pliage individuel d'un bateau en papier, sous les instructions de l'expérimentatrice E.G. Cette activité était réalisée assis.

Vidéo diffusée pendant le délai de rétention :

Le film diffusé pendant le délai de rétention avait une durée de 20 minutes et comportait une alternance de séquences représentant des paysages, sans représenter d'êtres humains, d'animaux, ni d'objets.

- Les 3 premières minutes consistaient en une vue de montagnes à partir d'un avion en vol, qui était accompagnée d'une musique calme.
- Les 3 minutes suivantes représentaient un désert rouge et un volcan. Des effets sonores accompagnaient la séquence, de type percussions. Il s'agissait d'un extrait du film *Baraka*, 1992.
- Les 3 minutes suivantes représentaient une vue d'une plage, accompagnée des sons du paysage correspondant. Il s'agissait d'un extrait de la chaîne de télévision MyZenTV.
- Les 3 minutes suivantes représentaient de nouveau une vue de montagne à partir d'un avion en vol, qui était accompagnée d'une musique plus dynamique de type électro.
- Les 3 minutes suivantes représentaient un désert filmé la nuit, ainsi qu'un ciel étoilé. De nouveau des sons de type percussions et une musique électro accompagnaient la séquence (*Baraka*, 1992).
- Les 5 dernières minutes représentaient des plages filmées de jour ou au coucher de soleil, accompagnées des sons réels de ce paysage (MyZenTV).

Des captures d'écran de cette vidéo sont présentées dans les images suivantes.



Image 1 : Montagnes vues d'avion



Image 2 : Volcan vu d'avion

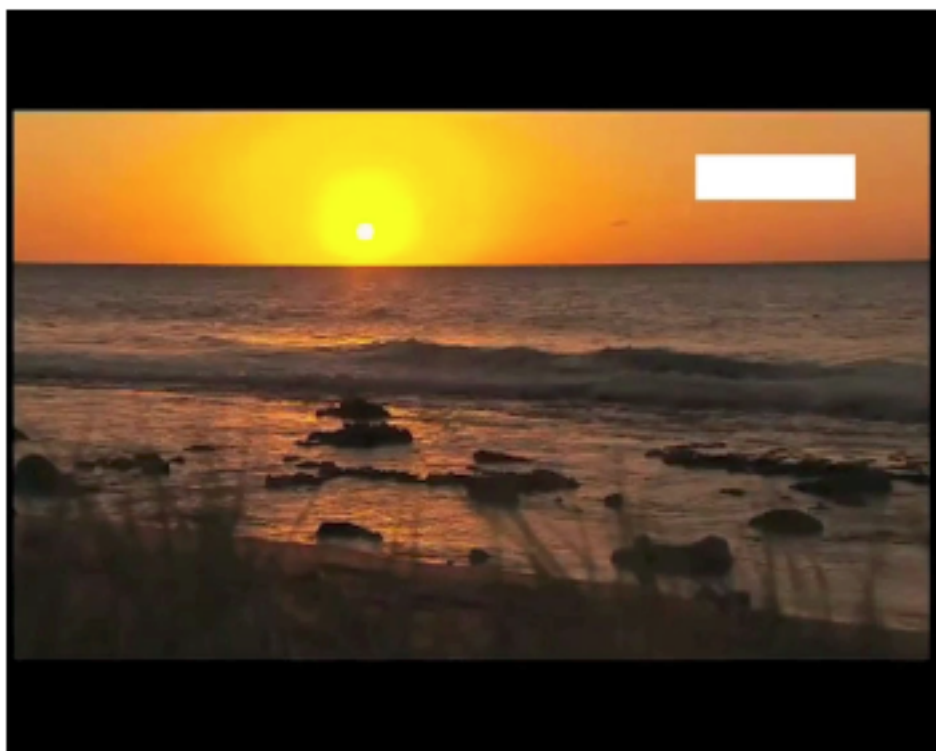


Image 3 : Plage