



THÈSE

En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par
Danaé REMON

Le 21 octobre 2019

**Stratégies d'apprentissage et mémoire à long terme
d'associations mot-objet chez le jeune enfant et le chien**

Ecole doctorale : **CLESCO - Comportement, Langage, Education, Socialisation,
Cognition**

Spécialité : **Neurosciences, comportement et cognition**

Unité de recherche :

CERCO - Centre de Recherche Cerveau et Cognition

Thèse dirigée par
Simon THORPE

Jury

Mme Juliane Kaminski, Rapporteure

M. Edouard Gentaz, Rapporteur

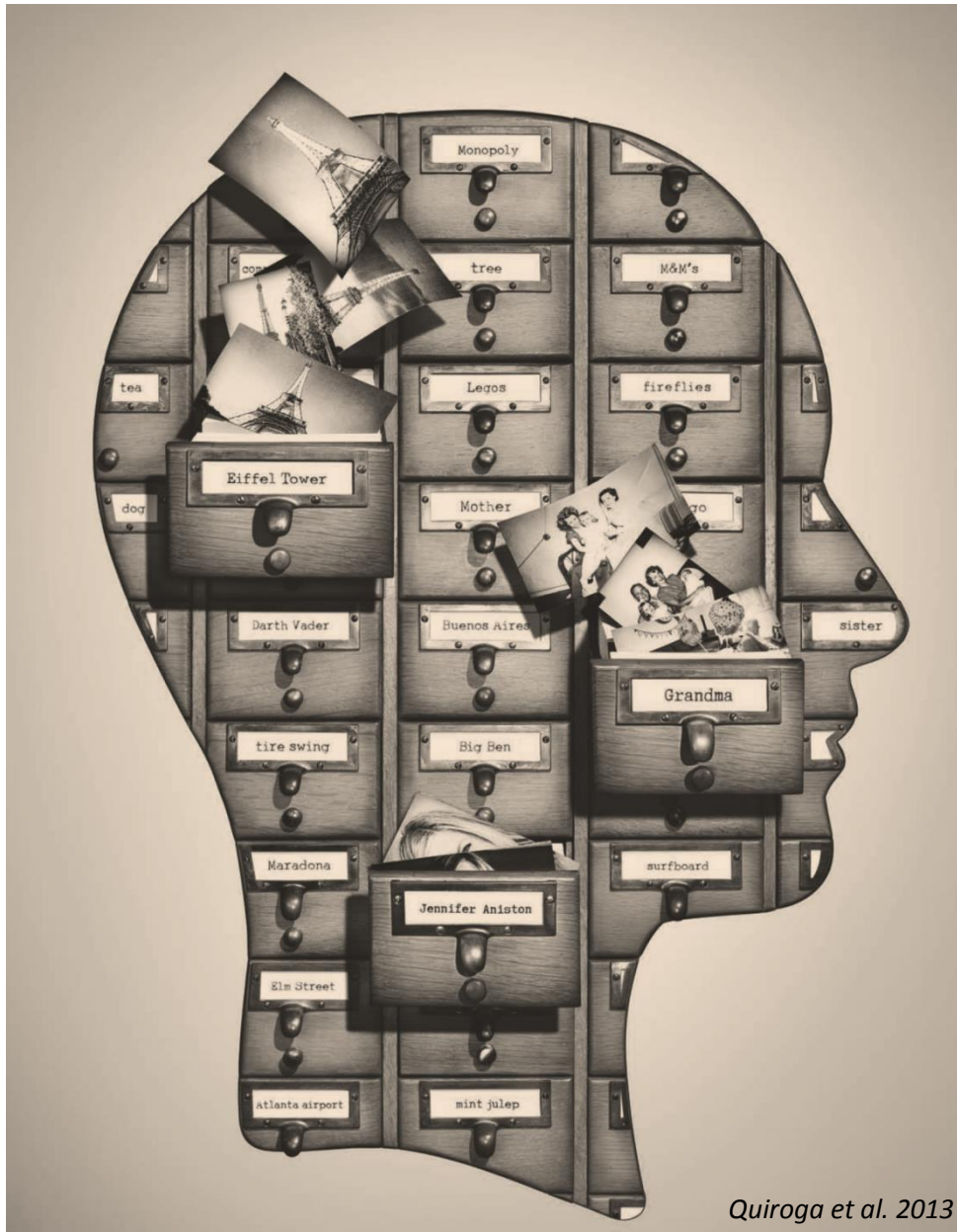
Mme Magali Batty, Examinatrice

Mme Virginie Durier, Examinatrice

M. Michel Fayol, Examineur

M. Simon THORPE, Directeur de thèse

Learning strategies and long-term memory of word-object pairs in young children and dogs. Danaé Rémon



Quiroga et al. 2013

Abstract

All of us occasionally remember information or personal events that occurred a long time ago. Many of us sometimes wonder how our brain encodes and stores such sensory information for a long time period. Whilst a myriad of researches enlightened some of the mechanisms of long-term memory formation in human adults, little is known about the mechanisms implemented in young children and distantly related animals. Do young children and animals also form memories that can last in time? And if so, does memory formation rely on the same principles than those established in human adults? Although these questions may seem trivial, they are far from being resolved. As evidenced, our inability to recall personal events that occurred during the first years of our lives – a phenomenon known as infantile amnesia – is counterintuitive given the extraordinary ability of infants and young children to acquire novel information on a daily basis. It might be that memory requires repetitions to be maintained durably, and this, especially for early-developing brain systems. On the basis of an abstract concept, the fact that words can refer to objects, I showed in this thesis that the ability of children to remember object names was positively correlated to the number of times they encountered these objects during the learning phase, whatever their age (18 months, 24 months, 4 years of age). Surprisingly, I also found that from 2 years of age, a single learning trial was sufficient to promote the retention of the name of objects after a 30min delay. It is the first time that such a demonstration has been possible. Finally, I highlighted the efficacy of two different learning strategies (“*ostensive labeling*”: i.e. passive learning implemented by an unambiguous naming of the objects; “*fast-mapping*”: i.e. active learning requiring the use of a logical deductive reasoning) in memorizing object names at these different ages.

During my thesis, I also intended to know whether some of these rules also apply to genetically distant animals. I chose the domestic dog as animal model. This species, largely unknown within the scientific community less than 20 years ago, is currently a privileged animal model given his incredible skills to communicate and cooperate with humans. Moreover, recent studies revealed the remarkable ability of some dogs in comprehending hundreds, even a thousand object names. Thanks to a partnership with the National Veterinary School of Toulouse and a professional dog trainer, I recruited and trained 40 dogs to acquire the word-referent concept. I demonstrated that acquiring such a complex task was age-dependent in dogs, just like the ability to recall object names after an intervening period of several months. A series of exploratory experiments conducted with the most expert dog of the cohort enabled me to show (i) that this dog required multiple trials to reliably learn the name of a novel object; (ii) that “ostensive labeling” was not an adequate learning method for him; (iii) that he was able to fast-map a novel word to its referent by exclusion despite his low “lexical repertoire”; and (iv) that learning from his own errors by a process of positive reinforcement was the most robust way to acquire and store durably novel information. Despite these notable learning divergences with humans, the accuracy of this dog to remember object names was highly significant even after several months of delay, suggesting that once information is learned, it can be stored in memory and remain “dormant” for months, maybe even years without decaying, as demonstrated in humans.

Key words: long-term memory; associative learning; word-referent concept; learning strategies; development; young children; dogs

Résumé

Nous faisons tous l'expérience de nous souvenirs d'informations ou d'épisodes de vie vécus il y a fort longtemps. Nombreux sommes-nous à nous demander comment nous encodons et stockons de telles informations sensorielles de façon durable dans le temps. A l'heure où les mécanismes de formation de mémoires à long terme trouvent de plus en plus d'éclaircissements chez l'Homme adulte, qu'en est-il pour nos jeunes enfants et pour les espèces animales génétiquement éloignées de nous ? Les enfants et les animaux forment-ils également des souvenirs qui peuvent demeurer intacts durant de longues périodes ? Et si tel est le cas, la formation de leurs souvenirs dépend-elle des mêmes règles que celles actuellement établies chez l'Homme adulte ? Ces questions, aussi triviales puissent-elles paraître, sont loin d'avoir été résolues. Pour preuve, l'incapacité à nous remémorer des événements vécus durant nos premières années de vie – phénomène connu sous le terme d'amnésie infantile – est contre intuitive avec l'incroyable capacité qu'ont les nourrissons et jeunes enfants à acquérir quotidiennement une multitude d'information. Il se pourrait qu'une mémoire ait besoin de répétitions pour s'établir de façon pérenne, et ce notamment pour un organisme en développement. En me basant sur un concept abstrait, celui que des mots peuvent désigner des objets, j'ai pu montrer que la capacité d'un enfant à se souvenir du nom d'un objet était en effet corrélée au nombre de fois que cet objet lui avait été présenté au cours d'une phase d'apprentissage, et ce quel que soit son âge (18 mois, 24 mois, 4 ans). Chose surprenante, j'ai également mis en évidence qu'un apprentissage en un seul essai était suffisant pour qu'un enfant âgé de seulement 2 ans se souvienne d'associations mot-objet apprises 30 minutes auparavant. C'est la première fois qu'une telle démonstration a pu être faite. Enfin, j'ai pu établir l'efficacité de deux stratégies d'apprentissage (« *ostensive labeling* »: apprentissage passif via une dénomination non ambiguë des objets ; « *fast-mapping* »: apprentissage actif via l'utilisation d'un raisonnement de déduction logique) sur la mémorisation de noms d'objets à ces différents âges.

Au cours de ma thèse, je ne suis également demandée si certaines de ces règles étaient transposables à des animaux génétiquement distants de nous. J'ai choisi pour modèle d'étude le chien domestique. Cette espèce, encore méconnue de la communauté scientifique il y a seulement 20ans de cela, est aujourd'hui un modèle de choix en cognition animale grâce à ses incroyables aptitudes à communiquer et coopérer avec l'Homme. De plus, des études récentes ont révélé la capacité remarquable de certains chiens à comprendre des centaines voire un millier de noms d'objets. En créant un partenariat avec l'Ecole Nationale Vétérinaire de Toulouse et une éducatrice canine professionnelle, j'ai entraîné 40 chiens à acquérir le concept mot-objet. J'ai pu démontrer que l'acquisition d'une telle tâche complexe est âge dépendante chez le chien, tout comme la capacité à se remémorer des noms d'objets après une latence de plusieurs mois. A l'issue d'une série d'études exploratoires menées avec le chien le plus expert de la cohorte, j'ai pu montrer (i) que l'apprentissage d'un nouveau nom d'objet nécessitait un grand nombre d'essais chez ce chien ; (ii) que « l'*ostensive labeling* » n'était pas une méthode adéquate pour lui; (iii) que ce dernier pouvait inférer un nouveau nom à l'objet référent par exclusion (« *fast-mapping* ») malgré son faible « répertoire lexical »; et (iv) qu'apprendre de ses erreurs par renforcement positif était la manière la plus robuste pour acquérir et retenir durablement des informations. Malgré ces divergences d'apprentissage notable avec l'humain, la performance de ce chien à retenir des noms d'objets était très significative même après plusieurs mois de latence, ce qui suggère, qu'une fois l'information acquise, elle serait stockée et pourrait rester en dormance des mois, voire années, sans être altérée, comme démontré chez l'Homme.

Mots clefs : mémoire à long terme ; apprentissage associatif ; concept mot-objet ; stratégies d'apprentissage ; développement ; jeunes enfants ; chiens.

Acknowledgements

Je me permets d'écrire cette section plus personnelle dans ma langue maternelle.

Une thèse, c'est aussi et peut-être avant tout une expérience humaine. J'ai eu la chance de côtoyer, connaître, vivre et travailler avec des personnes formidables que j'aimerai chaleureusement remercier ici.

Je remercie en tout premier lieu Simon Thorpe, sans qui cette aventure n'aurait probablement jamais vue le jour. Merci pour tes grandes qualités humaines, pour ton savoir et tes connaissances qui m'ont apporté de nombreux éclaircissements tout au long de ma thèse, pour ton incroyable ouverture d'esprit, pour l'autonomie que tu m'as donnée et la confiance que tu m'as portée tout au long de ces 4 années. Merci de m'avoir donné cette chance.

Un très grand merci à Lucie Lacroux-Negro, qui a tout de suite accepté de travailler avec moi sur le projet chien et qui a largement contribué à la réalisation de cette étude. Grâce à ton expertise, ton œil avisé et ton esprit critique, nous avons pu mener du mieux que nous le pouvions cette étude ensemble. Nous avons dû nous remettre de nombreuses fois en question, surmonter des situations que nous étions loin d'avoir imaginées, développer nous-même des méthodologies ; en bref je pense qu'au travers de cette étude les chiens nous auront aussi beaucoup appris. Grâce à toi, plus que jamais je sais qu'il faut travailler « avant tout le fond et non la forme ».

Merci à toi Gabriel. Je n'aurai toujours pas compris ton secret : être investi dans une multitude d'activités personnelles et professionnelles - que tu mènes toutes d'une manière remarquable -, être toujours de bonne humeur, curieux et disponible, être engagé pour des causes justes et en plus de tout cela, consacrer 2 de tes précieuses années d'études vétérinaires à mes expérimentations de thèse avec ton chien Moka. Je n'oublierai jamais votre complicité, ce duo parfait que vous formez, l'intelligence de ton chien, cette confiance que vous vous portez l'un l'autre. Merci d'avoir donné autant de ton temps, merci pour ta patience, pour ta générosité. Merci à toi Moka.

Merci à tous les autres couples maîtres-chiens qui ont accepté eux aussi de donner de leur précieux temps, qui m'ont accordé leur confiance, et qui pour beaucoup, ont tenu bon malgré les obstacles pour aller au bout de l'entraînement avec leur chien. J'ai beaucoup aimé partager cette étude avec vous. J'ai été impressionnée par la relation privilégiée que vous entretenez avec vos compagnons à quatre pattes. Vous avez été pour beaucoup une source d'inspiration pour moi. Merci donc à : Emeline & Jeika ; Amélie & Yuki et Alba ; Béatrice & Harley et Happy ; Emilie & Glee ; Roxane & Kinlo ; Eleonore & Tanga ; Camille & Kinai ; Diane & Nausicäa ; Laura &

Eïwa ; Ciska & Baïka ; Hugues & Jazz ; Margalide & Balto ; Johanna & Loki ; Sophie & Hyuri ; Lisa & Vénus ; Aurélia & Sunshine ; Damien & Elis ; Blandine & Harper ; Laura & Izotz ; Aurélien & Kupo ; Théo & Alaska ; Noémie & Oréo ; Violette & Nayla ; Clémence & Miko ; Léa & Fenrir ; Etienne & Neurone ; Angèle & Nova ; Claire & Mango ; Emilie & Aléombre ; Léa & Nami ; Elodie & Nova ; Patrick & Night ; Marie & Mao ; Michel & Maïka ; Nicole & Nyx ; Christine & Nash.

Merci à Nathalie Priymenko et Elisabeth Jeunesse pour cette collaboration. Grâce à vous j'ai pu recruter ces étudiants assidus et mener mes expérimentations au sein de l'Ecole Vétérinaire de Toulouse. Merci pour les réflexions que nous avons pu avoir sur ce projet.

Un grand merci à toi Olivier Pascalis. Je te remercie de m'avoir si souvent et toujours si bien accueillie à Grenoble. Je m'y suis sentie comme à la maison grâce à toi. Merci d'avoir priorisé la disponibilité de ton Babylab pour mes manipes lors de mes visites et de m'avoir permise de recruter autant d'enfants durant ma thèse. Merci pour tes conseils et discussions scientifiques, mais aussi pour tes nombreuses anecdotes et histoires, pour la bonne ambiance que tu génères au LPNC. Et bien sûr merci pour ce voyage à Philadelphie que nous avons partagé à plusieurs lors du congrès ICIS de 2018. J'en garderai de supers souvenirs, au-delà des conférences...

De Grenoble, je tiens aussi à remercier Hélène Loevenbruck, qui m'a beaucoup aidé lors de la réflexion et mise en place des manipes enfants. Ton expertise en psychologie de l'enfant et développement du langage m'a été très précieux. Merci pour la liste de pseudo-mots. Merci à Christophe Savariaux et Marion Dohen pour l'aide aux enregistrements audio et vidéos.

Toujours de Grenoble, merci à Marie Sarremejeanne, pour le recrutement des enfants, pour l'aide à la passation lorsque nécessaire, pour ta disponibilité et tous simplement pour tous les bons moments partagés. Hélène Mottier, un petit clin d'œil au passage aussi.

Du CerCo, je ne peux que chaleureusement remercier Karine Bouyer. Tu as été ma « petite maman » et plus encore pendant la quasi intégralité de ma thèse, à veiller que tout aille bien, à m'assouplir tout le côté administratif, à me relire lorsque nécessaire, à me remonter le moral lorsque j'étais moins bien, à m'écouter lorsque j'avais besoin de me confier. Que ce fut précieux de t'avoir à mes côtés (au sens premier comme figuré).

Je remercie aussi les 3 stagiaires de Master que j'ai eu la chance d'encadrer durant ma thèse : Djama, Ines et Océane. Vous avez été de belles rencontres humaines. Vous avez contribué au bon déroulement des manipes et votre aide fut précieuse. J'espère vous avoir aussi apporté en retour, j'espère que vos chemins vous mèneront là où vous le désirez, vous le méritez.

Merci à mes supers collègues du 2nd étage, sans vous ma thèse n'aurait pas eu la même saveur. Il y a eu tout d'abord Christelle, mon petit poisson, j'espère qu'on continuera à se voir ; Jaya et sa cool attitude ; Martin, nous on n'est pas sans se revoir et un IMMENSE merci pour tous ce que tu as fait en programmation pour moi ; Ludo, le maître des statistiques et des élastiques ; Elodie et ses piments en formes de zizis ; Jimmy et ses bougonneries mais heureusement aussi ses drôleries et nos discussions sport ; Sophie, merci pour ton aide au recrutement ; Pierre-Yves pour ton initiative du super trip dans la Death Valley et pour toutes nos discussions passionnantes ; Anne une très belle personne avec qui je n'aurai malheureusement pas passé assez de temps ; Anna...

A special thank to you Pascale Cook, pour avoir relu une bonne partie de mon manuscrit de thèse. Grâce à toi, je ne devrai normalement plus jamais écrire « such as for example », ni « consist in » !! Les distinctions entre « these » et « those » ou « among » et « amongst » n'ont désormais plus de secrets pour moi.

Merci à tous les membres de l'équipe MAOS, ce fut un plaisir de faire partie de cette équipe dynamique, où convergent de nombreuses disciplines et compétences. J'ai appris beaucoup.

Merci à Maroufa !! Ton sourire, tes bons petits plats, tes petits coucous à 19-20h le soir quand tu venais nettoyer mon bureau et me dire qu'il était temps de rentrer !!

Et puis bien sûr merci aussi aux 278 familles qui ont accepté de venir au CerCo ou au Babylab du LPNC pour faire participer leur enfant à l'une de mes manipes ! Merci donc aussi à vous les petits loulous, d'avoir effectué mes « petits jeux sur télé ». C'est grâce à vous que j'ai eu tous ces résultats. Et merci pour vos dessins.

Merci à tous les membres du jury pour avoir évalué mon travail de thèse.

Merci à ma famille, à mes amis.

Merci à Alexis. Avoir tout laissé derrière toi pour me suivre dans cette aventure, ce n'était quand même pas rien.

Merci à Matis. De loin, la meilleure chose que j'aurai accomplie durant ma thèse...

Financial support

I had the privilege of receiving funding during my entire thesis. This funding was provided by the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n°323711-Project M4). I want to acknowledge this funding agency for its substantial financial contribution.



Preamble

Recently, my little boy of almost 3 years of age asked me while looking at his dish:

He: « *Mum, what is that?* »

Me: « Mushrooms! »

He: « *The mushrooms we picked yesterday?* »

Me: « I don't know what you are talking about? We haven't been picking mushrooms yesterday! »

He: « *Yes, we did! The mushrooms we found in the forest with Flo and Alexandra...* »

I was speechless. This anecdotal event happened once, 8 months ago and we never talked about it since.

I bet all parents have experienced such type of episode with their child.

Paradoxically, there is very few evidence in the scientific literature that children under 3-4 years of age form memories that can last in time. Paradoxically too, my son will normally not be able to recall this event as an adult. Turned the other way round, adults do generally not remember any personal event they experienced early in life, a phenomenon defined as infantile (and juvenile) amnesia. So makes the difference between an early-developing brain system and an adult mature brain in the ability to sustain a memory trace after a substantial delay?

Table of contents

ABSTRACT.....	2
RESUME.....	3
ACKNOWLEDGMENTS.....	4
FINANCIAL SUPPORT.....	7
PREAMBLE.....	8
GENERAL INTRODUCTION.....	15
1 How is information processed and stored?	16
1.1 A little bit of history	16
1.2 From reductionist approaches to computational models.....	17
1.2.1 Hebbian theory	17
1.2.2 Aplysia californica: a great model for the study of neuronal plasticity	18
1.2.3 Spike-Time Dependent Plasticity: a refined form of Hebb's learning rule.....	21
1.2.4 Theory of neuronal selectivity	22
1.3 Information storage	24
1.3.1 Sparse coding theory and "grandmother" cells	24
1.3.2 Distributed coding theory.....	26
2 The M4 project.....	27
2.1 Ten "provocative claims"	27
2.2 From neuron to behavior	28
2.2.1 "Bringing very Long-Term Memories back to life"	28
2.2.2 Long-term memory for noise.....	29
2.2.3 Memory for repeated images via RSVP	29
3 Presentation of the thesis	31
3.1 Objectives of this thesis.....	31
3.2 Models of research of this thesis.....	32
3.3 Our experimental paradigm	33
PART 1: LEARNING STRATEGIES AND LONG-TERM MEMORY OF WORD-OBJECT PAIRS IN YOUNG CHILDREN.....	37
<u>CHAPTER I. Neural substrate of the human mnemonic system during development.....</u>	<u>38</u>
1 Brain development from gestation to early adulthood	38
1.1 Prenatal brain development	38
1.2 Postnatal brain development.....	40
1.2.1 Synaptogenesis	40
1.2.2 Myelination	41
1.2.3 Brain metabolism: glucose uptake	42
1.2.4 Regressive events.....	42

2	The mnesic system and its development	44
2.1	Procedural (implicit) memory.....	47
2.1.1	The neural substrate of procedural memory.....	47
2.1.2	Development of the neural bases of procedural memory.....	47
2.2	Declarative (or explicit) memory	48
2.2.1	The neural substrate of declarative memory	48
2.2.2	Development of the neural bases of declarative memory.....	54
<u>CHAPTER II. Relation between brain and long-term memory development: a focus on declarative memory.....</u>		<u>60</u>
1	Mobile conjugate reinforcement task: an operant conditioning task?	61
2	Recognition memory: Visual Paired Comparison (VPC) tasks	63
3	Basic relational memory functions.....	66
3.1	Spatial relational memory.....	66
3.2	Deferred imitation	67
4	Beyond long-term memory development: do infants and young children generalize?	69
<u>CHAPTER III. Relation between brain and semantic memory development: a focus on word learning.....</u>		<u>74</u>
1	Generalities about language acquisition	74
1.1	Word reception	75
1.2	Word production	76
1.3	Theoretical and computational models to explain early word learning	78
2	The word-referent concept.....	82
2.1	Word learning strategies.....	83
2.1.1	Cross-statistical word learning (or cross-situational word learning).....	84
2.1.2	Ostensive labeling (“ostensive naming”, or “unambiguous naming”)	86
2.1.3	Referent selection (or “disambiguation ability”)	88
2.2	Long-term storage of word-referent pairs in young children	92
2.2.1	Ostensive labeling: long-term retention of ostensively labeled objects.....	92
2.2.2	“Fast-mapping”: referent retention.....	93
2.3	Which neural bases support the formation of a memory for ostensively-named and fast-mapped words?.....	95
<u>CHAPTER IV. Experimental projects.....</u>		<u>101</u>
1	Presentation of the project	101
2	Overall materials and methods.....	102
2.1	Collaborations.....	102
2.2	Participants	103
2.3	Ethics.....	103
2.4	Stimuli.....	103
2.4.1	Learning stimuli.....	104

2.4.2	Testing stimuli	105
2.5	Apparatus.....	105
2.6	Procedure.....	105
2.7	Analyses	108
2.7.1	Accuracy during testing.....	108
2.7.2	Differences between groups	108
2.7.3	Effect of the language level	109
2.7.4	Effect of attentiveness	109
3	FIRST EXPERIMENT. <i>Ostensive labeling: effect of the number of presentations on word retention</i>	110
3.1	Introduction	110
3.2	Methods.....	112
3.2.1	Participants.....	112
3.2.2	Procedure	112
3.3	Results	112
3.3.1	Main results.....	112
3.3.2	Complementary analyses.....	118
3.4	Discussion	122
4	SECOND EXPERIMENT. <i>24-month-olds and above remember novel object names after a single learning event</i>	127
4.1	Introduction	127
4.2	Methods.....	129
4.2.1	Participants.....	129
4.2.2	Procedure	129
4.3	Results	129
4.3.1	Mains results	129
4.3.2	Complementary analyses.....	133
4.4	Discussion	135
4.5	Conclusion.....	140
5	THIRD EXPERIMENT. <i>“fast-mapping”: an influent strategy to promote word retention in 4-year-olds</i>	140
5.1	Introduction	140
5.2	Methods.....	142
5.2.1	Participants.....	142
5.2.2	Stimuli.....	142
5.2.3	Procedure	142
5.3	Results	144
5.3.1	Main results.....	144
5.3.2	Complementary analyses.....	150
5.4	Discussion	154

5.5	Conclusion.....	160
CHAPTER V. Ongoing experiments.....		161
1	FOURTH EXPERIMENT. <i>Ostensive labeling versus “fast-mapping”: which strategy is the most efficient to promote word retention during childhood?</i>	161
1.1	Introduction	161
1.2	Methods.....	164
1.2.1	Participants.....	164
1.2.2	Stimuli.....	164
1.2.3	Procedure	165
1.3	Results	166
1.4	Discussion	169
2	FIFTH EXPERIMENT. <i>Immediate versus delayed memory in children</i>	173
2.1	Introduction	173
2.2	Methods.....	174
2.2.1	Participants.....	174
2.2.2	Procedure	174
2.3	Results	175
2.4	Discussion	176
3	SIXTH EXPERIMENT. <i>Long-term memory of word-object pairs in children</i>	178
3.1	Introduction	178
3.2	Methods.....	179
3.2.1	Participants.....	179
3.2.2	Procedure	180
3.2.3	Analysis	180
3.3	Results	181
3.4	Discussion	184
4	GENERAL DISCUSSION	187
PART 2: LEARNING STRATEGIES AND LONG-TERM MEMORY OF WORD-OBJECT PAIRS IN DOGS.....		190
CHAPTER VI. Canine cognition research: state-of-the-art.....		192
1	Origins and evolutionary social skills of the “man’s best friend”	192
2	The ontogeny of lab research on canine cognition.....	195
2.1	Canine research in few figures	195
2.2	Brain correlates of canine cognition.....	196
3	Word comprehension by dogs	199
3.1	Word learning within the animal kingdom	199
3.2	Word learning by dogs	202
3.2.1	The word-referent concept: dogs retrieve objects by name.....	202
3.2.2	Multiple-items “sentence” comprehension by dogs.....	204

3.2.3	Word generalization abilities by dogs	205
3.2.4	“Fast-mapping” abilities by dogs	208
4	Memory in dogs	211
4.1	Behavioral research on memory in dogs.....	211
4.1.1	Working memory.....	211
4.1.2	Episodic-like memory	213
4.1.3	Long-term memory in dogs	216
4.2	Canine brain anatomy of the mnesic system.....	218
5	Conclusion of the chapter and presentation of the experimental project.....	219
<u>CHAPTER VII. Empirical evidence of long-term memory of word-object pairs in dogs.....</u>		<u>222</u>
1	FIRST STUDY. <i>Age effect on the acquisition and retention of a high-order cognitive task in dogs: the word-object pair paradigm</i>	222
1.1	Introduction	222
1.2	Materials and methods.....	224
1.2.1	Collaborations	224
1.2.2	Ethics Statements.....	225
1.2.3	Recruitment.....	225
1.2.4	Subjects.....	225
1.2.5	Stimuli.....	227
1.2.6	Procedure	228
1.2.7	Analyses, modelling and statistics.....	234
1.3	Results	236
1.3.1	Comprehension of the word-referent concept by dogs	236
1.3.2	Long-term retention of object names by dogs	238
1.4	Discussion	240
<u>CHAPTER VIII. Moka, a case study to understand the learning principles underlying long-term memory formation of word-object pairs in dos.....</u>		<u>250</u>
2	SECOND STUDY. <i>Memory maintenance across time</i>	250
2.1	Introduction	250
2.2	Materials and methods.....	252
2.2.1	Subject	252
2.2.2	Stimuli.....	252
2.2.3	Procedure	252
2.2.4	Analyses.....	253
2.3	Results	253
2.4	Discussion	255
3	THIRD STUDY. <i>Efficiency of various learning strategies in learning and memorizing novel word-object pairs</i>	256
	General Introduction.....	256
3.1	Does ostensive labeling induce word learning in a dog?	258

3.1.1	Introduction	258
3.1.2	Material and methods.....	258
3.1.3	Results and discussion.....	261
3.1.4	Conclusion	265
3.2	Is memory formation related to the number of exposures in dogs?	266
3.2.1	Introduction	266
3.2.2	Material and methods.....	268
3.2.3	Results and discussion.....	268
3.3	Which learning strategy is the most appropriate to induce a long lasting memory of an object name? 273	
3.3.1	Introduction	273
3.3.2	Material and methods.....	274
3.3.3	Results and discussion.....	276
4	FOURTH STUDY. <i>Generalization of Moka's knowledge to novel stimuli</i>	279
4.1	Introduction	279
4.2	Materials and methods.....	281
4.3	Results	282
4.4	Discussion	284
5	GENERAL DISCUSSION	287
	CONCLUSION, LIMITS AND PERSPECTIVES	296
	REFERENCES.....	301
	RESUME SUBSTANTIEL.....	321

GENERAL INTRODUCTION

Life is everywhere. From the simplest unicellular organism to the most complex creature, life is present. But life is also complex. To live, organisms must interact and evolve in an extraneous and potentially dangerous environment. Via the process of natural selection, some species will arise and develop while others will decay. To survive, compete and reproduce, a large part of the living creatures are provided with a fascinating capability of adaptation. Behavioral adaptation, which can be implemented consciously or reflexively, requires animals to “*learn*” from their environment. Animals continuously process multiple sensory inputs and when endowed with this capability, integrate that information for later use. Being equipped with the ability to store information in memory gives the living beings an extraordinary advantage to survive in the wild. But how do diverse and distantly related species learn and potentially memorize information? Are there some standard principles and/or mechanisms that govern learning? And how can we, human beings, explore these rules and mechanisms in a broad range of species to better comprehend how learning and memory functions? It’s been decades or even centuries since these fundamental questions have captivated researchers. However, understanding how information is physically processed and stored is challenging. Researchers are attempting to explore these issues at multiple levels of organization, from the molecular, cellular, tissue, behavioral and recently even computational levels. They also examine these questions from very simple to very complex systems and try to figure out whether they share common features or whether evolution shaped heterogeneous learning and memorizing modalities. On one hand, some researchers are interested in investigating animal models that best fit the human model, especially for clinical purposes (e.g. developing a convincing animal model to study Alzheimer’s disease). On the other hand, other researchers aim to better comprehend whether the principles and mechanisms that guide learning and memory formation in humans, are unique to humans or conversely largely widespread within the animal kingdom. In this second case, from an evolutionary perspective, it may help understanding how these mechanisms have evolved and perhaps became more complex during evolution. In this thesis, we were interested in examining if some of the principles that govern the establishment of long-term memories in adult humans would also apply to an unrelated animal species and to an early human developing brain system. The domestic dog appeared to us as an interesting species since a myriad of recent researches have highlighted its impressive communicative skills with humans, suggesting that despite strong anatomical, morphological and functional differences with us, this species has developed some surprising cognitive traits that are perhaps not so remote from our own. Furthermore, how human beings learn and store information early in life remains poorly elucidated. Do babies and young children already learn and memorize

sensory inputs in a similar way than adults? Or conversely, are learning and memory formation cognitive skills that slowly improve during development and that depend on specific and mature brain correlates?

By exposing young children and domestic dogs to similar complex cross-modal sensory inputs, we have tried to shed light on the expansionist aspect of some principles that underlie learning and memory formation. This thesis does not aim to simply compare the performance of an animal model to that of a young child on a given task, but instead to understand whether they share some basic principles with adult humans.

These principles originate from the M4 project (Memory Mechanisms in Man and Machine) initiated by Simon Thorpe in 2012 and will be mainly investigated at a behavioral level. However, in this thesis we will try to relate our findings to the general knowledge provided by the literature about the underlying brain mechanisms and putative neural correlates. Thus, this thesis is at the edge between behavior, cognition and neurosciences.

1 How is information processed and stored?

1.1 A little bit of history

It was originally believed that learning required the formation of new neurons. In 1894, **S. Ramon y Cajal**, a Spanish neuroanatomist, was the first to suggest that learning novel information might rather be caused by a process of strengthening connections between existing neurons to improve the effectiveness of their communication at the synapse. His idea was theorized and further developed by **Donald Hebb** in 1949. At that time, neuroscientists were not equipped with neurophysiological techniques needed to determine if their theory applies at a biological level. The first empirical evidence of this postulate came from electrophysiological recordings within isolated neurons using a reductionist approach. This research was conducted by **Eric Kandel** in the years 1960-70s on *Aplysia*, a gastropod mollusk. At the same time, a second empirical demonstration arose from the work of **Terje Lomo** and **Tim Bliss** who were at the origin of the elucidation of the phenomena of **long-term potentiation** and **long-term depression**, subsequently to investigations in the hippocampus of rabbits. Nowadays, advanced techniques allow researchers to better characterize the processes underlying learning and memory formation in more complex organisms like primates (including humans). Computational models also help refining the original theories.

1.2 From reductionist approaches to computational models...

1.2.1 *Hebbian theory*

The Hebbian theory was introduced by Donald Hebb in 1949 and is also sometimes called **Hebb's rule** or **Hebb's postulate**. This theory postulates that a repeated stimulation of a presynaptic cell on a postsynaptic cell would result in an increased synaptic efficacy between these cells (Hebb, 1949).

“Let us assume that the persistence or repetition of a reverberatory activity (or “trace”) tends to induce lasting cellular changes that add to its stability. [...] When an axon of cell *A* is near enough to excite a cell *B* and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such as *A*'s efficiency, as one of the cells firing *B*, is increased” (Hebb, 1949).

This postulate is often summarized as “*cells that fire together wire together*”, which is a misleading restatement of Hebb's theory. Indeed, Hebb pointed out that the presynaptic neuron (cell *A*) needs to “take part in firing” the postsynaptic neuron (cell *B*). In other words, he implied the notion of causality and temporal precedence: the fact that cell *B* can only fire if cell *A* fired before, not at the same time.

According to Hebb, this mechanism would allow the **increase of the synaptic strength between recruited neurons**. He also assumed that this strengthening process helps form highly connected assemblies of neurons that could be the neuronal support of *engrams*¹.

“Any two cells or systems of cells that are repeatedly active at the same time will tend to become “associated” so that activity in one facilitates activity in the other. [...] When one cell repeatedly assists in firing another, the axon of the first cell develops synaptic knobs (or enlarges them if they already exist) in contact with the soma of the second cell” (Hebb, 1949).

¹**Engram:** *a theorized biophysical or biochemical change in the brain in response to external stimuli as a means of storing memories.*

Unfortunately, at that time, Donald Hebb could not verify if his theory was truthfully the one implemented in a living organism. About 20 years later, the pioneering laboratory work of Eric Kandel provided for the first time strong evidence of the involvement of the Hebbian theory as a learning mechanism in a simple-organized animal model.

1.2.2 *Aplysia californica: a great model for the study of neuronal plasticity*

Investigating the Hebbian theory and more generally how sensory information is processed in a billion of neurons brain organism is challenging. Moreover, highlighting the brain mechanisms underlying a given behavior in a complex mammalian brain is an extreme difficult task. Eric Kandel, recipient of the Nobel Prize in physiology or medicine in 2000, used a reductionist approach to tackle these challenging questions. He used an invertebrate animal model, the *Aplysia californica*, which is a marine gastropod mollusk. This organism is an advantageous model for neurobiologists, because it has only about 20,000 neurons. Moreover, the **gill and siphon withdrawal reflex** is an involuntary, defensive reflex that causes the retraction of the siphon and gill each time the animal is disturbed. This reflex undergoes classical conditioning, in a way that is comparable to classical conditioning described in other animals like mammals (i.e. stimulus-response, temporal specificity, effect of context, etc.). Furthermore, this reflex is mediated for one-third by monosynaptic connections between sensory neurons and motor neurons. Consequently, using a preparation of an isolated abdominal ganglion of the *Aplysia* connected to a piece of skin from the tactile receptive field of the reflex, Kandel and coworkers were able to relatively easily investigate the cellular bases of this behavior by recording the intracellular activity of identified sensory and motor neurons, simultaneously with behavior.

During conditioning, a weak tactile stimulation administrated to the siphon automatically elicits the retraction of the gill [Figure 1]. After repeated exposures, the gill progressively stops retracting. This progressive decrease in response to a particular (and weak) stimulus is called **habituation**. Now, if the animal is presented to another novel stimulus or to a stimulation on another part of the body, the automated behavior (gill retraction) is partially or completely restored (i.e. **dishabituation**). Finally, if the animal is presented with a novel strong and noxious stimulus (e.g. an electrical choc), an increased response occurs (i.e. **sensitization**).

At a neuronal level, the authors found that the synapses that connect the sensory neurons to the motor neurons underwent plasticity that was activity-dependent (Castellucci, Pinsker, Kupfermann, & Kandel, 1970). Specifically, they found that the excitatory post-synaptic potential (EPSP) exhibited a **low-frequency depression when repeatedly stimulated** [Figure 1], and a **facilitation after application of a novel stimulus** to another pathway. Consequently, the authors concluded that both habituation and dishabituation can be explained by changes in the efficacy of specific excitatory synapses. Such **plastic changes in the functional effectiveness of synapses** have been suggested as the neuronal mechanism of a behavioral modification at short-term and is in agreement with Hebb's postulate.

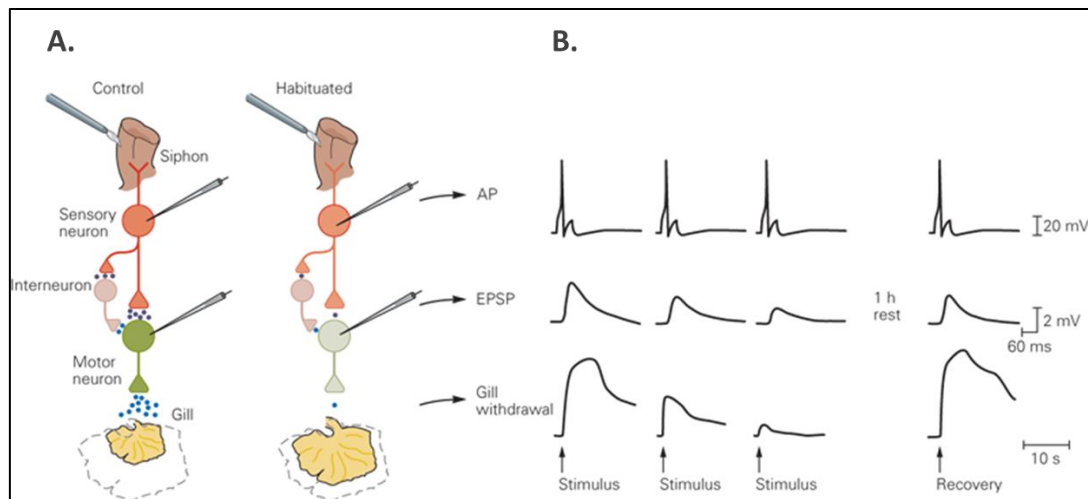


Figure 1 – A. Illustration of the gill and siphon withdrawal reflex before (left) and after conditioning (right). When habituated to an external stimulus, the system results in a depressed efficacy of the synapses between the sensory and motor neurons. **B.** Illustration of the electrophysiological recordings within the sensory neuron and the motor neuron, as well as the amplitude of retraction of the gill during a repeated stimulation of the siphon. *From Neurosurgery; cellular mechanisms of implicit memory storage and the biological basis of individuality, 2017.*

Later on, Kandel and coworkers identified proteins that, when synthesized, mainly contribute to convert short-term memories into long-lasting memories (namely **CREB: cAMP Response Element Binding Protein**). The activation of this protein results in **an increase in the number of synaptic connections**.

Taken together, their work led to the discovery that short-term memory is primarily shaped by functional changes in the effectiveness of existing synapses (increased or reduced activity in response to stimulus repetition), while long-term memory requires a restructuration of the synaptic connections such as a change in the number of synaptic connections. Their work was the first empirical evidence of Donald Hebb's theory in a living animal. Since that time, evidence of synaptic plasticity - either transiently or durably – and of an adaptation of the spiking activity in response to repeated external stimulations has been reported in more complex organisms including monkeys (e.g. De Baene & Vogels, 2010), and humans (e.g. Pedreira et al., 2010), thanks to the development of neuroimaging techniques for instance.

The pioneering findings about the learning mechanisms gave rise to investigations at a molecular level and probably contributed to the elucidation of the now well-defined **long-term potentiation²** and **long-term depression²** mechanisms that mainly occur in the medial temporal lobe (Lomo, 1971, 2003).

²**Long-term potentiation (LTP):** *persistent increase in synaptic strength (or potentiation) following high-frequency stimulations and caused by an enhancement in signal transmission between two neurons. Unlike other processes of synaptic plasticity, LTP lasts a (very) long time (from minutes to months, even years). This durable modification at the synapses contributes to the learning process and to the maintenance of a memory at long-term.*

An early form of LTP (E-LTP) depends on a transient activation of protein kinases (CaMKII and PKC) that are autonomously able to carry out the phosphorylation events that underlie this potentiation (note: phosphorylation consists of adding a phosphate to a molecule in order to change its activity). This E-LTP transiently increases the effectiveness of the connection at the synapse. A late form of LTP (L-LTP) requires gene expression and/or the synthesis of proteins in the postsynaptic cell, which is caused by a persistent activation of protein kinases during E-LTP. It has also been shown that the induction of L-LTP coincides with the activation of CREB (a transcription factor) responsible for the long-term synthesis of proteins at the synapse (i.e. maintenance of the structural changes operated at the synapse). This L-LTP is required for the persistence of a memory at (very) long-term.

LTP is often studied in the hippocampus of several animals, including humans, but is also observed in other neural structures, including the cerebral cortex, cerebellum and amygdala.

Long-term depression (LTD): *activity-dependent reduction in the efficacy of neuronal synapses following a prolonged presentation of a stimulus. It selectively weakens specific synapses. It is assumed that LTD prevents synapses from reaching a ceiling level of efficacy, which would prohibit the encoding of new information (i.e. the formation of novel memories).*

At the same time too, other groundbreaking experiments highlighted other effects of repetitive sensory processing on changes at a neuronal level. Indeed, being repeatedly exposed to a given stimulus not only strengthens or weakens the connections between the neurons; it also leads to **neuronal selectivity**. In the 50s, **David Hubel and Torsten Wiesel** recorded the activity of neurons thanks to microelectrodes implanted in the primary visual cortex of anesthetized cats while they projected patterns of light and dark on a screen in front of the cat (Hubel & Wiesel, 1959). They found that some neurons fired rapidly and consistently when presented with distinct line orientations, while other neurons discharged only when presented with another angle of light (their work introduced the notion of *receptive field*). They also established that “complex” neurons responded preferentially to moving lines of light displayed in a certain angle, suggesting that there might be a **hierarchy in the integration of sensory information** and that the **brain**

constructs complex representations on the basis of simple sensory information (Rose & Blakemore, 1974). Their work introduced the notion of **experience dependent selectivity**, a point that will be further developed below. Note that these two neuropsychologists were also awarded in 1981 (Nobel Prize in Physiology or Medicine) for their discovery of the ocular dominance columns in kittens as a consequence of visual deprivation in one eye early in life.

1.2.3 *Spike-Time Dependent Plasticity: a refined form of Hebb's learning rule*

Donald Hebb's original assumption about synaptic plasticity, followed by several empirical demonstrations of its involvement in simple or complex living organisms, is still the foundation of current theory about learning and memory formation. Hebb's postulate was later refined and gave rise to the ideas that (1) **synaptic weights are only modified when the postsynaptic neuron fires a spike** and that (2) **the induction of synaptic modification depends on the temporal order in the pre- and postsynaptic spiking** (Bi & Poo, 2001; Markram, Lubke, Frotscher, & Sakmann, 1997). These ideas were recapitulated in the Spike-Time Dependent Plasticity (STDP) theory (Markram et al., 1997). In its classical version, the STDP theory postulates that the **synapse of an afferent neuron that fires before a given neuron will be strengthened** (in this case, LTP is observed). In contrast, if **the inputs of a neuron arrive after the given neuron fired, the synapse gets weakened** (in this case, LTD is observed) [Figure 2].

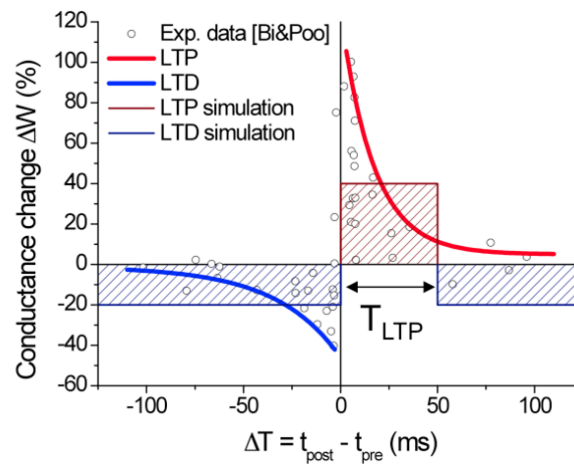


Figure 2 – Graph illustrating different versions of the STDP learning rule. Classically, synapses activated before the spiking activity of a given neuron are potentiated (LTP, solid red line) as a function of the time difference between input and output spikes (curving shape of the solid red line). In contrast, synapses activated after the spiking activity of a given neuron are depressed (LTD, blue solid line) as a function of the time difference between input and output spikes (curving shape of the solid blue line). These curves were based on the experimental data of Bi and Poo, 2001 (circles). In another STDP version, synapses are potentiated provided that the spiking activity of the incoming neuron is very close in time to the spiking activity of the given neuron (red hatched region); in any other cases synapses are depressed (blue hatched region). From Bichler et al. 2012.

In a revised STDP version, some researchers emphasize the importance of **temporal contiguity** suggesting that synapses get systematically weakened (depressed) except if the afferent neuron

fires in a narrowed time-window before the postsynaptic spike (Bichler et al., 2012; Guyonneau, VanRullen, & Thorpe, 2005). It stresses the importance of temporal coding constraints for the strengthening process to occur (i.e. a “**critical time window**”) [Figure 2].

In sum, repeated exposures to an external sensory stimulus (e.g. image, sound, etc.) will result in a **strengthened neural network** involving only the neurons that were the first to fire in response to this specific stimulus and whose firing activities were temporally close to each other. Such mechanism would prevent all the cells from learning the same pattern. Consequently, neurons equipped with such STDP-learning rules will naturally become **selective to a given repeated external pattern** (Masquelier & Thorpe, 2007) as mentioned above.

1.2.4 *Theory of neuronal selectivity*

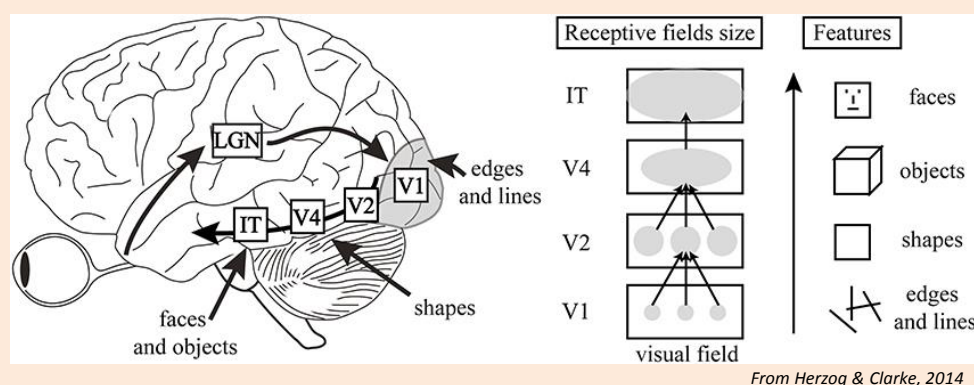
By using a feedforward computational network implementing an unsupervised³ STDP-based learning algorithm, Masquelier et al. found that neurons (that mimicked neurons in higher order visual areas) gradually became selective to frequently occurring features (stimuli consisted of pictures from the Caltech Dataset). Less than one hundred presentations were enough for neurons in the network to start becoming selective to given patterns. Moreover, the responses of the neurons became also more and more rapid (i.e. a given image was more rapidly processed over the course of presentation of the images). Even more intriguing, by using the simplified STDP learning rule described above, computational studies found that **neurons became selective to any repeating stimuli and only after a ten of repetitions** (Bichler et al., 2012).

³*Unsupervised learning: No instructions are given to the network about what should be learned*

These findings established at a computational level are supported by research in humans using a Stereotaxic Electro Encephalography method (intracranial single unit recording) implemented in epileptic patients suffering from intractable epilepsy, and Electro Encephalography (EEG) methods in healthy participants. Humans seem indeed able to form selective neurons for arbitrary patterns (at least from the visual and auditory domains), simply as the consequence of repeated exposures (Agus, Thorpe, & Pressnitzer, 2010). Moreover, evidence for rapid and sharp selectivity has been found in participants who incidentally learnt acoustic patterns (Andrillon, Kouider, Agus, & Pressnitzer, 2015). In their supplements, the authors showed that only five presentations sufficed to observe neuronal selectivity. **Neurons selective for complex and high-order patterns have essentially been found in the medial temporal lobe of humans** - a brain region that only receives *highly integrated information*⁴ and that is known to be involved in memory processes (see Chapter I) – as well as in **neocortical regions**, thought to be the final

repositories of long-term memories (e.g. Kreiman, Koch, & Fried, 2000). Some authors proposed that the selectivity of neurons within the medial temporal lobe is only transient while long-lasting selectivity would be specific to neocortical neurons (see Chapter I).

⁴How is information processed in the brain? To briefly summarize: The sensory organs first transform physical external stimuli (e.g. chemical molecules, sound waves, photons, etc.) into electrical signals: action potential. The spikes first reach the primary sensory cortices (e.g. V1 for the visual domain) where only basic unimodal features are processed (e.g. cells only code for lines orientation or edges, etc.). Bottom-up (i.e. feedforward) processes then enable the integration of multiple low-level features to create coherent and more complex representations (i.e. inputs from the different sensory cortices are then sent to other brain regions along the hierarchy of processing where higher-level processing occur – e.g. cells from V2, V4, IT for the visual domain are respectively selective to shapes, full objects and distinct faces). Ultimately, inputs reach association cortices and finally the medial temporal lobes where highly integrated polysensory information are related together to form concepts and whole representations (e.g. the name of an object). Note that top-down (i.e. feedback) processes are also often implemented and seed-up the recognition of some features.



Selective neurons can be seen as **highly local representations of specific features/stimuli/information**. Specifically, it has been shown that a selective neuron only fires when the stimulus for which it is selective is present (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005). For example, researchers reported the cases of neurons that were specifically selective to faces of Jennifer Anniston, Halle Berry or Oprah Winfrey (Quiroga et al., 2005). Similar findings were reported for particular places such as the Sydney Opera or the Taj Mahal (Quiroga et al., 2005) [Figure 3]. Interestingly, the same authors showed that a given neuron not only responded to images of its favorite percept but also to stimuli from other modalities linked to this percept (e.g. text (the word form of the person or place), auditory (the person's voice),

etc.)(Quiroga, Kraskov, Koch, & Fried, 2009). Hence, **single neurons in the human medial temporal lobe respond selectively to representations of the same percept across different sensory modalities**. This finding gave rise to the idea of **highly-integrated “concept neurons”**.

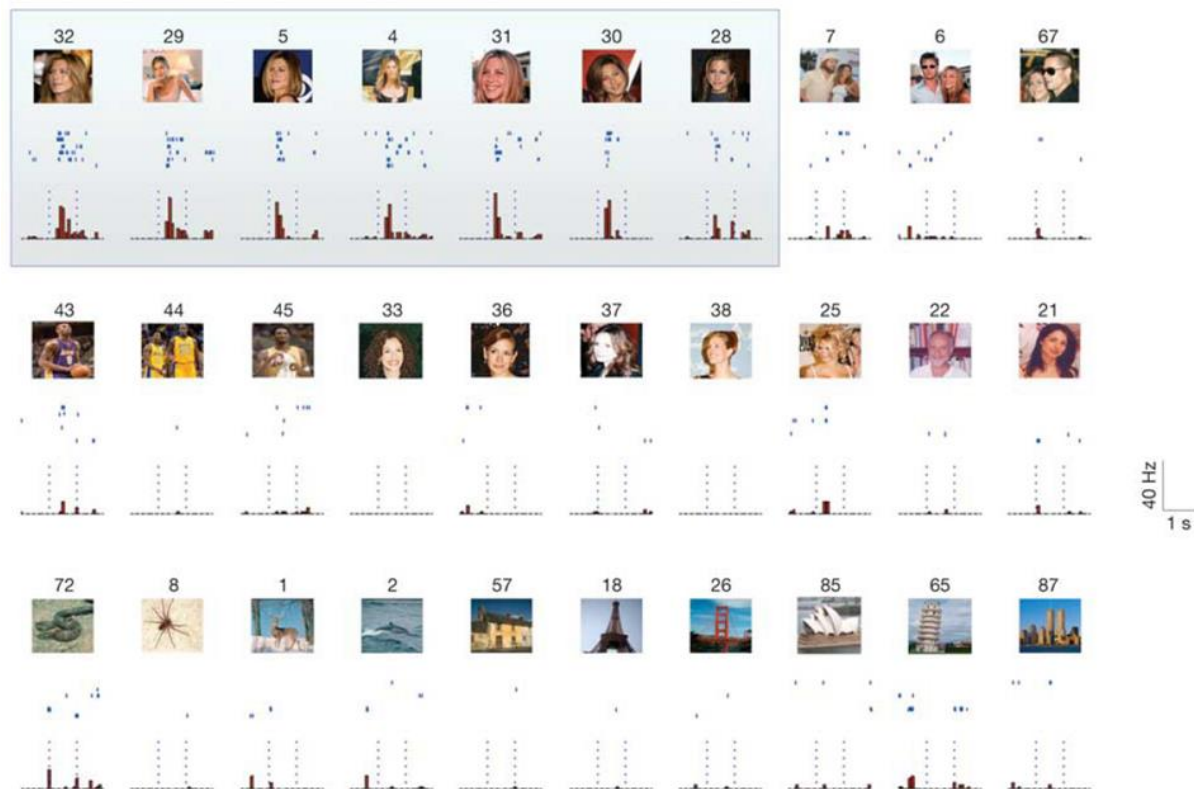


Figure 3 – A single unit in the left posterior hippocampus activated exclusively by different views of the actress Jennifer Anniston. Responses to 30 of the 87 images are shown. For each picture, the corresponding raster plots and post-stimulus time histograms are given. Vertical dashed lines indicate image onset and offset. *From Quiroga et al. 2005.*

1.3 Information storage

1.3.1 Sparse coding theory and “grandmother” cells

These groundbreaking findings led to the assumption that such highly selective neurons should play a critical role in memory functions. On one hand, some researchers claimed that relatively few neurons – thousands or less - that store information in such a highly specific manner may be the repositories for any particular concept (e.g. a person, a thing, etc.). According to this **sparse coding theory**, the information in its complete form is supported by (relatively) few specialized neurons locally settled (i.e. a “node”) [Figure 4]. If so, reactivating these neurons would be sufficient to trigger recall.

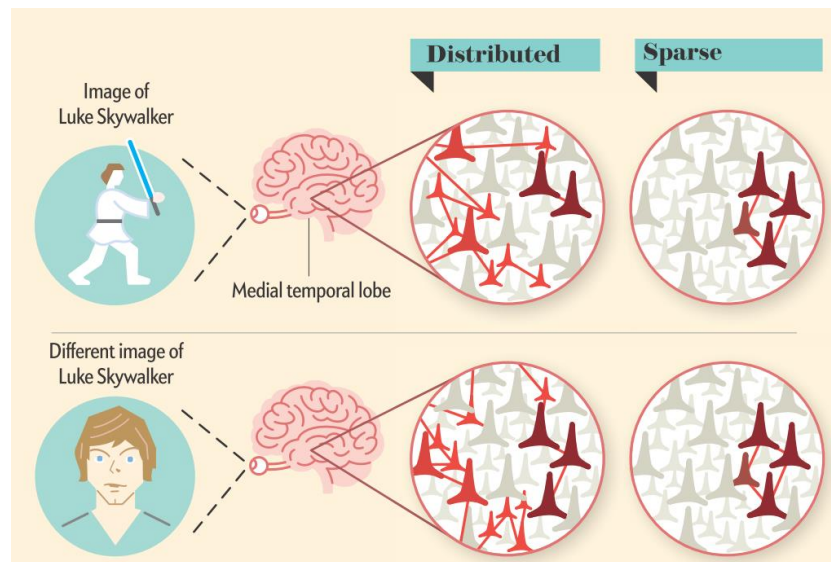


Figure 4 – Illustration of two alternative theories about how memories are stored in the brain. The distributed coding theory claims that information/concepts are stored as bits and pieces distributed across millions or billions of neurons. The sparse coding theory argues that a (relatively) few neurons (thousands or even less) constitute a “sparse” representation of an information/concept. *From Quiroga et al 2013.*

This theory gave weight to the controversial concept of “grandmother cells” claiming that highly selective neurons in the neocortex are the neuronal support for (very) long-term memories. The postulate is the following: given that a selective neuron only fires when the stimulus/concept/information for which it is selective is present, in the absence of this stimulus/concept/information, the neuron should not fire at all. If so, the selective neuron may preserve its pattern of selectivity and connectivity within the network for the entire period of absence of the stimulus. It was proposed that the **neuron may remain “dormant” months, years or even decades later, ready to fire again if the stimulus appears again** (Thorpe, 2011b, 2011a). In a radical way of considering the “grandmother cell” hypothesis, only one neuron would respond to one specific stimulus or concept. In other words, each stimulus or concept would be coded by a different neuron. According to some authors, such a mechanism would not be biologically relevant, because it would imply that if the single neuron responsible for a given concept dies or becomes damaged, the whole trace of this concept would vanish from memory (Quiroga, Fried, & Koch, 2013). Therefore, it may be more likely that more than a solitary neuron is specialized in any concept. It might also be that neurons are specialized in more than a single concept (Quiroga, Fried, & Koch, 2013). Thus, there is a debate about whether these neurons should be called “grandmother cells” or “concept neurons” (Quiroga, Fried, & Koch, 2013). Finally, note that there is a substantial difference in the theorized localization of “grandmother” neurons and “concept” neurons. Indeed, “grandmother” cells are believed to be situated in cortical regions (neocortex), whereas “concept cells” as defined by Q. Quiroga were essentially found in the medial temporal lobe of humans. As mentioned above and as it will be larger reviewed later on in this thesis, it is now well-established that the shift from mildly long to very long-term

memory is characterized by a transfer of the information from the MTL regions to cortical regions (see Chapter 1). Therefore, it was proposed that a neuron from the hippocampus (i.e. one component of the MTL, see Chapter 1) that for example selectively responds to “Halle Berry” at t_0 , might respond to something completely different few years later (i.e. transient selectivity). One way to reconcile these divergences is to imagine that the specialization property of neurons within the MTL is different from that of neurons from the neocortex. Hippocampal neurons may be selective for more than a unique concept, while cortical neurons, which are much more numerous, may be effectively specialized for unique information/concepts.

1.3.2 *Distributed coding theory*

On the other hand, according to an alternative viewpoint, storing any specific concept or random stimulus can only be permitted by the collective activity of millions or even billions of neurons distributed widely throughout the brain. According to this **distributed coding theory**, no few neurons support the entire information but multiple neurons that store a small independent part of the information [Figure 4]. A concept/information (e.g. “Halle Berry”) would then be the sum of individual neurons. Recall would be the result of activating all these distributed neurons. The major argument of proponents of this theory against a sparse coding hypothesis is that we may not have enough neurons to represent all possible concepts and external stimuli. However, it is argued that a typical person remembers about 10,000 concepts, not more, which is far below the number of neurons available.

To summarize

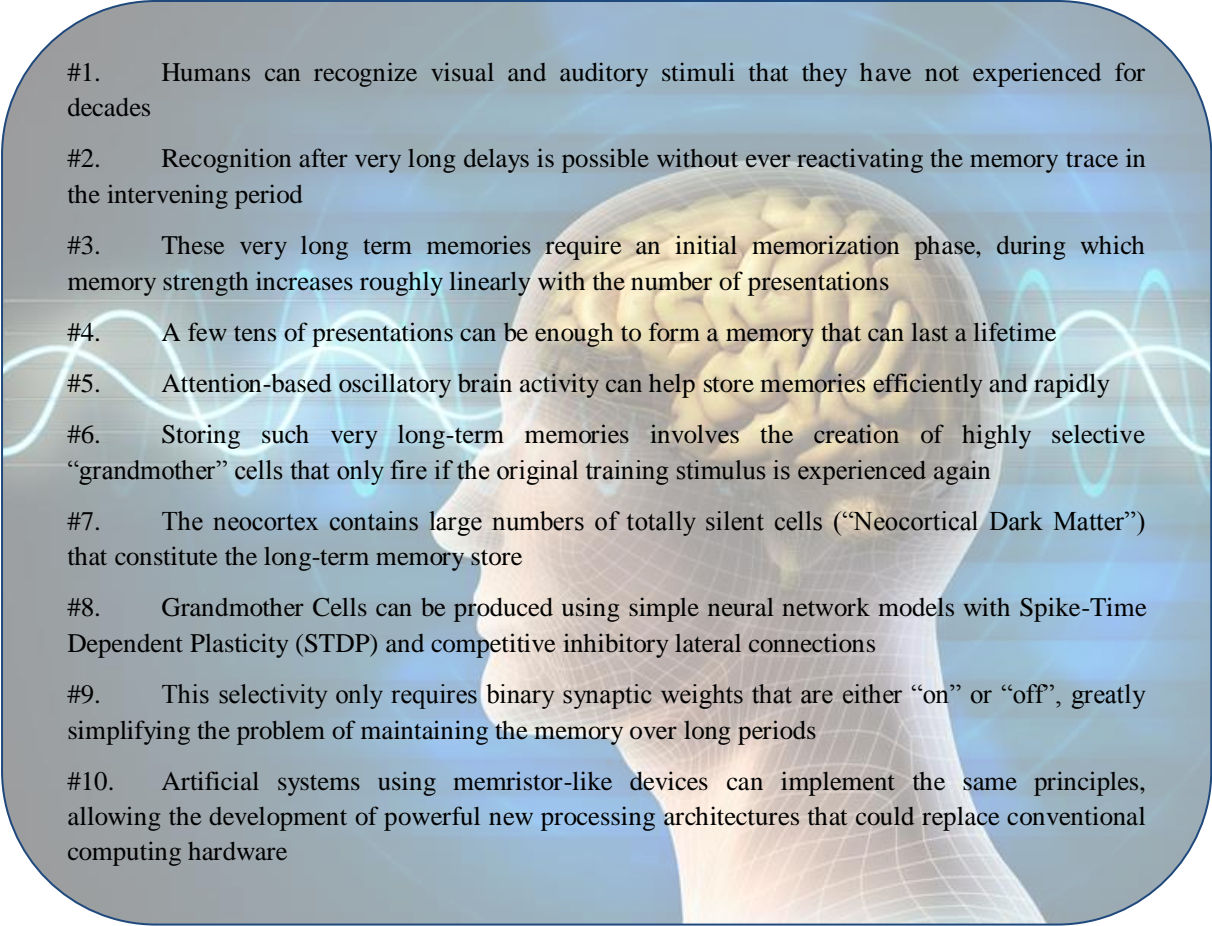
We started this introduction from a theory that gave rise to a simple behavioral observation – the fact that a gastropod that reflexively retracts his gill in response to an external stimulation progressively reduces this pattern of response if the harmless stimulation is repeated – and ended with the idea that repeated exposures to external stimuli or concepts induce the emergence of highly selective cells following processes of synaptic plasticity. We discussed that these changes at the synapses can last in time and consequently that a given neuron can maintain its selectivity for a long time period even if the stimulus is not presented again (though there is a transfer of the information from the MTL to cortical regions). Altogether, this highlights the powerful role of repetition in the formation of memories. **Repetition would be necessary for a neuron to become selective. Repetition would be necessary for the information to be maintained in the brain for a long time period.** In other words, repetition would be necessary for the information not to be “lost” (or at least inaccessible).

But what is repetition? How do we define it? Should repetition necessarily be close in time for a neuron to become selective or can it also be distributed in time? In a later discussion, we will discuss the fact that *repetition* is not as easy to define and that its temporal boundaries (especially in the visual domain) are often vague. Moreover, as it will also be discussed later on, repetition can be either directly or indirectly provided. In the case of personal events, although we usually experience them only once, there are several ways to indirectly rehearse them: dreams, mental rehearsal, narration to a tierce person, photos, videos, etc. Each of these rehearsals may contribute to the maintenance of the memory trace at very long term. But **how many times should a human being be either directly or indirectly exposed to an information/stimulus/concept to keep a trace of it during a long period of time?** This is a challenging question, especially since “indirect” repetitions are impossible to control.

2 The M4 project

2.1 Ten “provocative claims”

The questions addressed above along with what has already been reviewed above, are at the heart of a project initiated by Simon Thorpe who received a European Grant to test these issues. This project, the M4 project (Memory Mechanisms in Man and Machine, <http://m4.ups-tlse.fr/>), is articulated around 10 “*provocative claims*” regarding how information is encoded and stored. These claims are as follows:

- 
- #1. Humans can recognize visual and auditory stimuli that they have not experienced for decades
 - #2. Recognition after very long delays is possible without ever reactivating the memory trace in the intervening period
 - #3. These very long term memories require an initial memorization phase, during which memory strength increases roughly linearly with the number of presentations
 - #4. A few tens of presentations can be enough to form a memory that can last a lifetime
 - #5. Attention-based oscillatory brain activity can help store memories efficiently and rapidly
 - #6. Storing such very long-term memories involves the creation of highly selective “grandmother” cells that only fire if the original training stimulus is experienced again
 - #7. The neocortex contains large numbers of totally silent cells (“Neocortical Dark Matter”) that constitute the long-term memory store
 - #8. Grandmother Cells can be produced using simple neural network models with Spike-Time Dependent Plasticity (STDP) and competitive inhibitory lateral connections
 - #9. This selectivity only requires binary synaptic weights that are either “on” or “off”, greatly simplifying the problem of maintaining the memory over long periods
 - #10. Artificial systems using memristor-like devices can implement the same principles, allowing the development of powerful new processing architectures that could replace conventional computing hardware

To further appraise the role of repetition in the formation and maintenance of sensory memories, below I wanted to review the research conducted by colleagues of the M4 team who investigated some of the 10 claims at a behavioral level, since this level of investigation is the one that interests us in this thesis. Thus, we would like to leave the neuronal theories for a moment to come back to behavior.

2.2 From neuron to behavior

2.2.1 “Bringing very Long-Term Memories back to life”

During her thesis, my colleague Christelle Larzabal, investigated the retrieval of memories that were kept inactive since many decades. In one study, adult participants watched short cuts of TV programs that were broadcasted on average 44 years ago (Larzabal, Bacon-macé, Muratot, & Thorpe, 2017). In this experiment, she rigorously controlled that participants could not have re-experienced the material during this long time period by selecting TV programs that had never been rebroadcasted since and that were not available anywhere (internet, public domain). She found that 6 TV programs among 50 were particularly well identified by the majority of the participants (using a four title propositions forced-choice procedure, 4 AFC). Interestingly, two

participants spontaneously retrieved the title of one TV program before the forced-choice task. Finally, some participants were able to recall information related to the video clips (e.g. day of the week it was broadcasted, duration of the episode, etc.). Taken together, her findings suggest that **re-exposures to the stimuli are not necessary to maintain information in memory for a lifetime.**

In another study, she tested participants' ability to recognize pictures (drawings) they saw very few times (one to three times, 2s each) 12 years earlier (Larzabal, Tramoni, Muratot, Barbeau, & Thorpe, 2018). Memory was assessed using recognition memory tasks (a forced-choice task and a yes/no task). Surprisingly, on average participants successfully recognized the stimuli during testing. The participants who saw the drawings more times (three times) than the others (only once) were systematically among the best subjects. This study showed that **very few exposures (three) to a material seem to be sufficient to recall that information at least a decade later.**

2.2.2 *Long-term memory for noise*

Another of my colleagues from the M4 team, Jaya Viswanathan, investigated memory formation within the auditory domain by using meaningless auditory stimuli (Gaussian noise). Such stimuli are particularly interesting because they cannot be consciously rehearsed. A previous study showed that humans, who repeatedly listened to meaningless auditory patterns during a learning session, greatly remembered these patterns over several weeks without conscious rehearsal (Agus, Thorpe & Pressnitzer, 2010). It provided strong evidence that, even in the auditory domain, a **sensory stimulus can be memorized over an extended period without the possibility of rehearsal.**

During her thesis, Jaya replicated this experiment but used modified versions of the learned noises at test. Specifically, she exposed participants to looped and scrambled versions (at 10 or 20ms bin size) of learned cyclic noises 4 weeks after learning. She demonstrated that participants were significantly better at recognizing the Gaussian noises, whatever their versions (original, looped or scrambled), compared to novel cyclic noises (discrimination tasks) (Viswanathan, Rémy, Bacon-Macé, & Thorpe, 2016). She provided further evidence that **information can be stored implicitly** and suggested that in the auditory domain, **neurons might be encoding small bits of information** (10-20ms bits) separately.

2.2.3 *Memory for repeated images via RSVP*

Back to the visual domain, a post-doctoral student of our M4 team investigated the effect of repetition on the recognition of images that were briefly presented in very rapid succession (rapid

serial visual presentation, RSVP) (Thunell & Thorpe, 2019a, 2019b). Adult participants had to detect repetitions in streams of thousands of natural images presented on a tactile screen and then to identify the repeated image among a choice of two (2-AFC). During the encoding phase, the images reoccurred a certain number of times (the framerate varied between 2 to 120 images a second) interleaved with a varying number of distractors (from 1 to 5). Her results showed that the performance of adults in detecting the repeated images increased with the number of presentations of the repeated images up to ceiling level around seven (Thunell & Thorpe, 2019a). She also established that performance was well above chance level even with only two presentations. Interestingly, she demonstrated that the ability of participants to detect and memorize repeated images was preserved even with irregular sequences (i.e. varying number of distractors between two presentations of the target image) (Thunell & Thorpe, 2019b).

Taken together, the authors reported **impressive abilities of adult humans in detecting and remembering repeated images presented in RSVP streams over a time-course of few minutes** and showed that only **2 presentations of an image sufficed for participants to recognize that image**. Their findings also indicate that temporal regularity is not necessary to form a memory trace for repeating images in RSVP streams and that the target images can be spaced in time during the learning session (i.e. they do not need to be presented in succession).

To summarize

Altogether, these findings give weight to the claim that, in adult humans, the **strength of the memory does increase roughly linearly with the number of presentations**, as suggested in the M4 project (claim #3) (Thunell & Thorpe, 2019a, 2019b). These studies also provided clear evidence that a **memory can be formed despite a very low number of exposures to the sensory inputs** (claim #4) (Larzabal et al., 2018, Thunell & Thorpe, 2019a) and that a **memory can sustain very long delays without being reactivated in-between** (claim #1 and 2) (Agus et al., 2010; Larzabal et al., 2017; Viswanathan et al., 2016). The authors proposed that such impressive skills should depend on **highly selective neurons** that were generated during learning and that **remained “dormant” during the entire period of rest** (claim #6). This suggests that **a very small number of presentations of the sensory inputs would be enough for neurons to become selective and to maintain a trace of the memory for a long time period**.

3 Presentation of the thesis

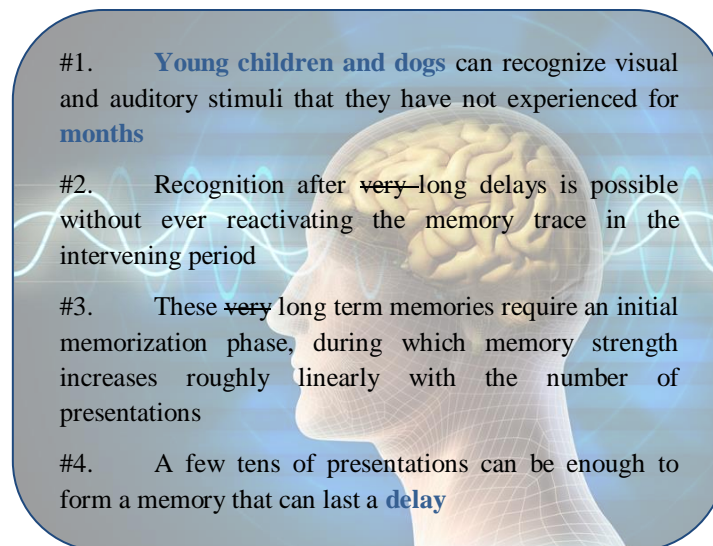
In this thesis, we were interested in **examining if some of these claims that are already partly elucidated in adult humans would also apply to humans early in life as well as to a non-primate animal model.**

Selective neurons have been identified in a wide range of animal species using electrophysiology methods (e.g. place cells, grid cells, etc.). Moreover, as reviewed above, empirical demonstrations of the Hebbian theory as well as the discovery of synaptic plasticity, LTP and LTP were allowed thanks to animal models. Therefore, there are great reasons to believe that these processes are the ones implemented in the formation of memories in animals (including humans). However, whether **selective neurons can sustain a memory trace for a long time period in animals as well as in young children is still unclear.** In other words, it is not established that a neuron of an early developing brain system and of a non-primate animal can remain “dormant” for months or even years. To the best of our knowledge, the concept of “grandmother” cells has never been investigated apart from adult humans. And this is not the purpose of this thesis either, since we were not recording the activity of unitary cells. However, even at a behavioral level, there is relatively few evidence that animals can form long lasting memories for arbitrary sensory information, as it will be reviewed later on in this manuscript (note that we are not talking about noxious stimuli (e.g. electric chocks) in the context of fear conditioning for example, but we are simply talking about ordinary sensory stimuli (e.g. images, sound patterns, etc.)). Similarly, whether long-term sensory memories exist early during childhood is still debated, especially regarding the different types of memories (a point that will be reviewed later on) and regarding the phenomena of infantile and juvenile amnesia.

3.1 Objectives of this thesis

Therefore, the first challenge of this thesis was to **explore whether long-term memories for arbitrary and complex but realistic stimuli exist in animals other than primates as well as early during the development of humans.** Along with findings from the literature, this would help understanding if the predictions of the models and theories fit with the observations at a behavioral level. The second challenge was to **examine the conditions required for the formation of these long-term memories.** This could have direct implications for educational purposes.

In order to investigate these questions, we specifically focused on four of the M4 claims that were slightly reformulated according to the populations we studied and to temporal constraints:



3.2 Models of research of this thesis

In this thesis, we concentrated our investigations on two models of research: young children (of 3 different age groups: 18-month-olds, 24-month-olds and 4-year-olds) and a non-primate model; the domestic dog (2 different age categories: puppies and adults).

As mentioned above, what is intriguing and fascinating with human memory is that adults typically remember nothing about personal experiences that occurred during their first years of life (a phenomenon called infantile amnesia). Paradoxically, babies, infants and children must be learning a lot. As it will be reviewed in this thesis, babies learn to recognize people, objects, odors, sounds, etc. They also acquire procedural memories (e.g. walk) and of course, language. Interestingly, these learnings are achieved before they are able to recall specific episodic memories and interestingly too, most of these learnings will last their lifetime which is not the case for episodic memories. One explanation is that the establishment of long lasting memories may require repetitions as suggested by one of the M4 claims, even early during childhood. As it will be discussed in this thesis, personal experiences occur only once and may require the child to rehearse the event internally to keep a trace of that event, an ability that is missing early in life. So, multiple repetitions of the information/events/stimuli may be all that is needed for young children to create long lasting memories. We also suggest that the strategies implemented during learning play a role in the storage of the information. To test these claims, we conducted a series of experiments involving cross-modal sensory inputs: object names. We tested the number of repetitions required to form a memory trace of the name of a novel object at different ages during development. We tested children after a 30min delay and were able to bring some participants back to the lab after 1 month or 6 months without rehearsal in between, to test if the memory trace lasted in time. We also tested the efficacy of two main learning strategies.

As a direct extension of the claims of the M4 project, we sought to examine the generality of the repetition-based learning process in a non-primate animal: the domestic dog. As it will be reviewed in this thesis, thanks to their impressive skills to communicate and cooperate with humans, dogs are now widely used in animal cognition. Moreover, recent studies revealed that individual dogs were capable of learning to fetch hundreds, even a thousand of objects by name. To do this, dogs must have formed memories of the associations. However, the question of how long these individual dogs were capable of retaining the information without rehearsal in between had not yet been addressed. Moreover, these studies have not investigated how many repetitions were required to form these memories. Finally, whether the learning strategies implemented to reach such impressive skills are similar to those used by humans is also still unclear. These were the main questions we focused on in this thesis. We recruited 40 dogs who were naïve about the word-referent concept. Thus, the first challenging step was to train these dogs on that task during a 6-month period. Then, only dogs who were reliably familiar with the fact that word forms can refer to objects could be used to investigate the issues addressed above. Unfortunately, well below our expectations, only one dog mastered the task after the training phase, and we therefore ran our series of experiment only with that dog subject. We were able to explicitly test the number of exposures to novel pairings that were necessary for him to remember these associations after a delay. We were able to examine if the memory trace decayed with time or remained relatively stable months after learning without rehearsal. Finally, we also attempted to identify how this dog learned novel word-object associations, i.e. what was the strategy that was the most efficient for him to acquire the name of a novel object. We assumed that having a better comprehension of the processes underlying learning in dogs can have implications for dog trainers.

In sum, this thesis should help understand if some of the claims of the M4 project also apply to other living organisms that have less complex or less mature brains.

3.3 Our experimental paradigm

The paradigm implemented in this thesis was the **word-referent paradigm**, that is, the fact that word forms can refer to specific items (or category of items). This paradigm involves two sensory modalities: the auditory and the visual domains. This paradigm is relevant for humans since humans are early in life accustomed to attribute a word form to each specific item of the environment. We will see in this thesis that word learning is also ecologically relevant for dogs who share our daily life and who also, more or less naturally, learn the meaning of specific words (e.g. words referring to specific actions: sit, turn left, stop here, bring back, etc.). However, this paradigm is a **highly demanding task** because it requires participants **(1) to encode the visual**

features of the objects (i.e. to create a representation of the objects), (2) to encode the auditory properties of the sound patterns (i.e. to create a representation of the word forms) and (3) to link the two representations together. At a neuronal level, regarding the theories and models mentioned above, this paradigm would imply the formation of **selective neurons within the visual brain areas** (i.e. selective neurons to each novel object), **selective neurons within the auditory pathways** (i.e. selective neurons to each novel word form), **and highly integrated selective neurons in multimodal association areas that are high up in the brain system** (i.e. “concept” neurons for each word-object pair).

In this thesis, we used objects and labels that were unknown to the participants to ensure that at the time of the experiment, they did not already have internal representations of these stimuli (note: we cannot exclude the possibility that the objects or word forms may evoked something to the participants, but with our method we were 100% sure that participants had never encountered the arbitrary associations as unitary “concepts” prior to the task (e.g. they could not have experienced the “rivou” as the label for a jumping green monster-like object before)). The second reason was to control that no re-exposure with these stimuli happened during the intervening periods, which would have reactivated the memory trace.

Finally, using such a difficult task involving complex bimodal stimuli enabled us to highlight differences across developmental, which would probably not have been the case with mere isolated images or isolated sound patterns or even with easier paradigms.

Definition clarification:

Most of our experiments about long-term memory in young children implemented a 30min delay between the learning and the testing phases. One could wonder whether it is really fair to talk about a ‘memory trace’ after a 30-minute delay? In the literature, “*long-term memory*” is not clearly defined. The traditional models about memory consider that the duration of a “*short-term memory*” ranges from few seconds to a minute. Short-term memory is sometimes distinguished from *working memory* that requires the individual to temporarily and actively hold and/or manipulate the information. As soon as the information persists after longer delays, and is accompanied by physiological changes and/or the synthesis of new proteins, the memory is considered as “long-term”, although there is no consensus about the exact timeframe (some researchers talk about minutes, other hours, others a night of sleep, etc.). In our protocols, the option of maintaining the information in working-memory was ruled out by the distractive activities performed by the children during the 30 minute delay. Therefore, it can be reasonably assumed that the participants had already started to consolidate the information into more permanent representations.

PART I



LEARNING STRATEGIES AND LONG-TERM MEMORY OF WORD-OBJECT PAIRS IN YOUNG CHILDREN



Abstract

In this part of the thesis, we will examine the question of long-term memory formation in young children in the context of a complex bimodal sensory concept: the word-object concept. This concept refers to the idea that word forms can designate specific items of the environment. Thus, this thesis will be tightly related to the development of the human being and will be at the interface between memory formation and word learning.

Therefore, it seemed essential to me to first provide the reader with substantial knowledge about the development of the human brain (Chapter 1), about memory formation during development from both a neural (Chapter 1) and a behavioral (Chapter 2) perspective and about the word-object concept (Chapter 3).

Precisely, we will start this “journey” with the earliest stages of brain development, from gestation to early childhood, but we will pay special attention to the development of the neural structures that encompass the mnemonic system (Chapter 1). In this first chapter, we will focus on the different types of memory, the different brain components that constitute each type of memory and their presumed roles, and finally their development during childhood. Different viewpoints, models and theories will come to support the state-of-the-art. In a second chapter, we will try to correlate the development of these neural structures with the onset and development of declarative memory, as assessed by behavioral studies. We will see that rudimentary forms of memory already emerge very early in life and then progressively improve with age. We will focus on some of these memory skills and will address the question of whether they truly tap into declarative memory or not. In a third chapter, our “journey” into the literature will take the path of the semantic memory branch of the declarative memory. We will specifically update our knowledge about word learning during childhood by reporting both ancient and recent studies, theories and computational models that should help the reader comprehend how children rapidly and fascinatingly acquire language. Next, we will be particularly interested in the word-referent concept that will be at the heart of the experimental work of this thesis. We will mainly focus on two word-learning strategies: ostensive labeling and inferential reasoning (i.e. “fast-mapping”). For both strategies, we will review how and when word learning translates into word retention. Finally, we will try to understand which brain structures may be involved in each of these learning strategies and how they may be responsible for the discrepancy of these two learning methods to induce long-term retention.

Next, we will get to the heart of the matter by presenting in detail the six experiments I carried out during my thesis. Throughout these experiments, we sought to shed light on some of the factors

and conditions necessary for the formation of long-term memories for word-object pairs in early childhood. This experimental part tested whether some of the provocative claims of the M4 project already apply to early developing brain systems. Specifically, we addressed the question of whether recognition after a delay is possible without reactivating the memory trace in the intervening period (recognition tests were performed after a 30-min distractive period and again after one month or 6 months for some participants) (#claim N°2). We also attempted to understand whether the memory strength increases roughly linearly with the number of presentations (#claim N°3). To answer that question, word-object pairs were presented a variable number of times to the participants and we explored whether there was a correlation between the number of presentations of the pairs during learning and the performance for each pair during testing. We also investigated the minimal number of presentations that were sufficient to form a memory trace that could survive a delay (#claim N°4).

In this experimental part, we also addressed questions that were outside the scope of the M4 project. Precisely, we wanted to better understand the effect of different learning conditions and learning strategies on the retention of object names. Specifically, we developed a tightly controlled procedure, which enabled us to compare the effect of ostensive naming and inferential reasoning on word retention at key stages of development. We also examined the effect of various variables on the formation of a memory for word-object pairs (e.g. attentiveness during learning, language level, exposure to another language, etc.). Finally, we tried to link our findings to the existing knowledge about the putative brain mechanisms provided by the literature as reviewed in the three first chapters, in order to merge our behavioral observations with neuroscientific processes.

CHAPTER I.

NEURAL SUBSTRATE OF THE HUMAN MNESIC SYSTEM DURING DEVELOPMENT

1 Brain development from gestation to early adulthood

The nervous system develops via the interaction of several synchronized processes, some of which are completed before birth, while others continue into adulthood. Data from animal and post-mortem studies tell us much about these processes. More recently, magnetic resonance imaging (MRI) has opened the field of investigations of brain changes during development in living people, which considerably improved our knowledge about the main developmental events a human undergoes before reaching adulthood.

1.1 Prenatal brain development

The sequence of events during the prenatal development of the human brain closely resembles that of many other vertebrates. The nervous system begins with a process of **neurulation**, which is a folding process of the ectodermal tissue onto itself to form a hollow cylinder called the **neural tube** (completion by 3-4 weeks of gestation). From 4 to 12 weeks, the neural tube differentiates along three dimensions: length, circumference and radius. The length dimension gives rise to the major subdivisions of the central nervous system: the **forebrain** and the **midbrain**, both arising at one end of the neural tube, and the **spinal cord** at the other end [Figure 5]. The front end of the neural tube progressively forms bulges and convolutions. The first bulge (from the frontal part) mutates into the cortex (telencephalon), the second starts to differentiate into the thalamus and hypothalamus (diencephalon), while the third turns into the midbrain (mesencephalon). The other bulges give rise to the cerebellum (metencephalon) and to the medulla (myelencephalon) (e.g. Dehaene-Lambertz & Spelke, 2015; Johnson & De Haan, 2015).

Meanwhile, sensory and motor systems originate from the circumferential dimension of the neural tube. On the top side, the sensory cortex can be distinguished while the bottom side roughly

corresponds to the motor cortex. In between, the various association cortices and the higher sensory and motor cortices are aligned.

Finally, the complex layering patterns and cell types are differentiated along the radial dimension of the neural tube. Moreover, within **proliferative zones** - close to the hollow portion of the neural tube (which will become the ventricles of the brain) - cells proliferate, migrate from their origins and journey to destinations within the cortex where they then differentiate into particular neuron types. These proliferative zones correspond to the ventricular zone (the first to be formed) and the subventricular zone (contributes to the development of the neocortex). Prior to birth, neurons are generated at a rate of more than 250,000 per minute (e.g. Dehaene-Labertz & Spelke, 2015; Johnson & De Haan, 2015; Lenroot & Giedd, 2006).

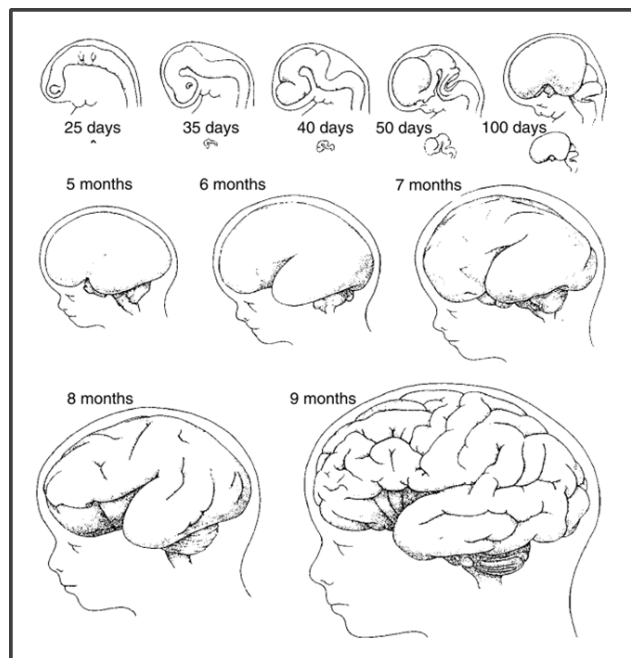


Figure 5 - Drawings of the human prenatal brain development (embryonic and fetal stages). The forebrain, midbrain and hindbrain originate as swelling at the head end of the neural tube. *From Johnson & De Haan, 2015.*

From this early stage, interactions between cells are critical, such as the transmission of electrical signals between neurons. It is assumed that the **waves of neural activity** intrinsic to the developing organism might play an important role in the emergence of brain structures before the individual is confronted to the sensory inputs from the external world (Shatz, 1996). Moreover, the assumption that the newborn brain may be structurally more developed than previously thought is currently a matter of debate. Initial studies of neonatal neural networks reported only a dense local connectivity within sparse segregated modules and few long-distance connections, suggesting that the neural connectivity essentially undergoes substantial remodeling after birth. For instance, a study examined the maturational changes of the cortical connectome in subjects

(premature neonates, term-born neonates, six-month-old infants and adult subjects), using diffusion tensor MRI (Tymofiyeva et al., 2013). They found basic modular network topology in newborn brains but observed increasing brain network integration and decreasing segregation with age in term-born subjects. In contrast, other studies revealed large-scale dynamic functional networks analogous to those seen in adults that would already be effective at birth. Precisely, a “*rich club*” of interconnected cortical hubs (in regions including the dorsal, medial frontal and parietal cortex, precuneus, hippocampus and insula), previously reported in adults, would already be present by 30 weeks of gestation (Ball et al., 2014). According to those authors, the number of connections between “rich-club” regions and the rest of the cortex increases significantly during 30 and 40 weeks of gestation, allowing information transfer across the cerebral network. They assumed that “rich-club” organization precedes the emergence of complex neurological functions, and that environmental stress alters the network architecture of premature extra-uterine life (Ball et al., 2014).

1.2 Postnatal brain development

One of the major postnatal developmental characteristics is the imposing increase in the total volume of the human brain from birth to early adulthood. This is unlikely to be due to additional neurons since the genesis of a vast part of them (except in the dentate gyrus) and their migration to their corresponding brain areas takes place predominantly during the prenatal developmental period (around the seventh month of gestation) (Rakic, 1995). Moreover, after migrating, a period a rapid cell death is observed, reducing the neural number by half from 24 weeks of gestation to 4 weeks after birth (Lenroot & Giedd, 2006). Only a comparatively reduced number of neurons will originate postnatally, primarily within the dentate gyrus of the hippocampus.

1.2.1 *Synaptogenesis*

Beginning around the 20th week of gestation, the increase in volume is already mainly attributable to a substantial growth of synapses, dendrites and fiber bundles [Figure 5]. Neurons’ dendritic trees increase in size and complexity which gives the neurons their specificity and specialization. Additionally, a significant proliferation of the density of synaptic contacts between cells contributes to refine the organization of the neuronal network. Using a phosphotungstic acid method which stains synaptic profiles selectively (fixed in tissues up to 36h postmortem), Huttenlocher and co. reported for the first time a considerable growth in the postnatal density of synapses regionally within the human cerebral cortex (Huttenlocher, 1979). This synaptogenesis boost begins around birth in all cortical areas, but the most rapid bursts of increase and the peak of density occur at different ages in different brain regions. For example, while the visual cortex’s

maximum density is reached between 4-12 months of age, the synaptic density of the prefrontal cortex increases very slowly and only reaches its peak after four years of age. Overall, by 2 years of age the level of synapses is on average approximately 50% greater than that typically measured in adults (Huttenlocher, 1979). This latter postmortem finding was confirmed by recent MRI studies (Lenroot & Giedd, 2006).

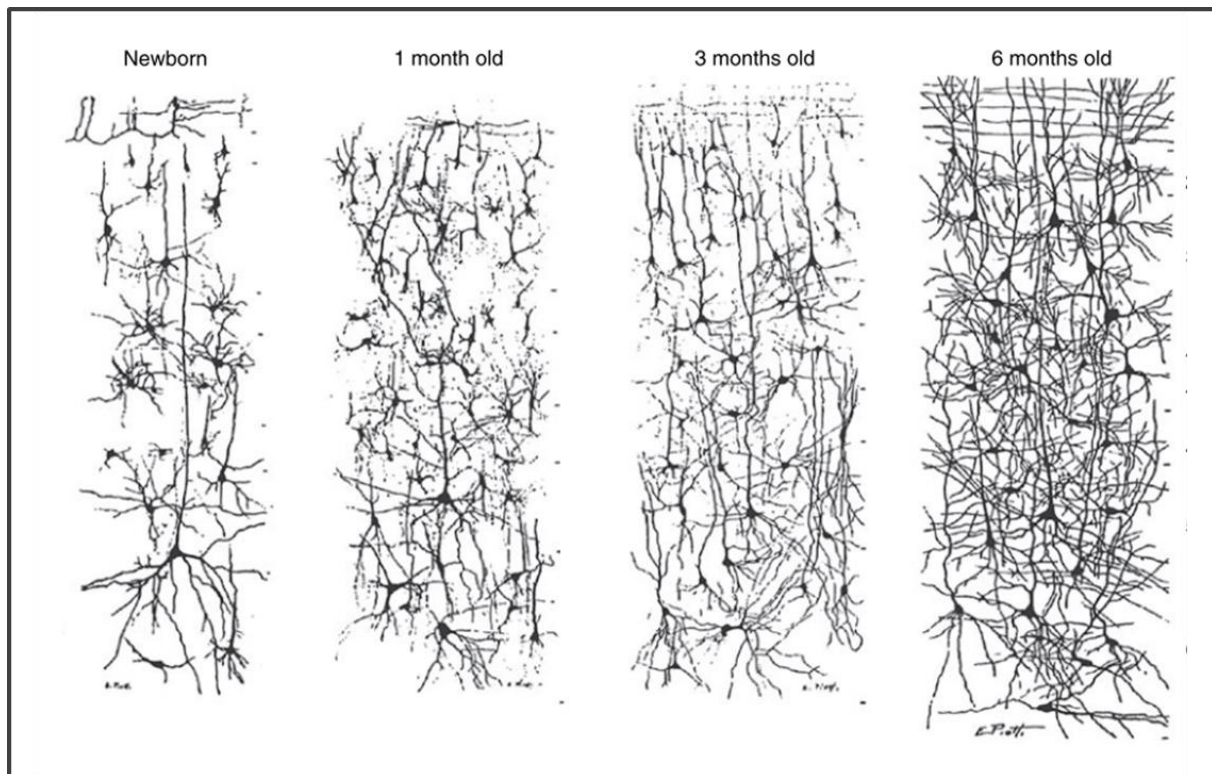


Figure 6 - Drawings of the cellular structure of the human visual cortex on Golgi stain preparations from Conel (1939-1967). From Johnson & De Haan, 2015.

1.2.2 *Myelination*

Another additive process that contributes to the brain's bulk is **myelination** [Figure 7]. Myelination is an increase in the fatty sheath that surrounds the neuronal pathways – and therefore the nerve fibers - and constitutes the **white matter**. Myelin is known to efficiently increase the speed of transmission of the electric signal (by as much as 100 times) between neurons. Sensory areas tend to myelinate earlier than motor areas. Cortical association regions myelinate last and the myelination process continues into the second decade of life. Even if most major tracts are myelinated by early childhood, axons within the cortex (and in some regions such as the arcuate fasciculus located near the temporal lobe) continue to myelinate into the second and third decades of life (Yakovlev & Lecours, 1967). Although the rate of white matter increase varies with age, there is no apparent reduction of it in any brain region until the fourth decade (e.g. Giedd &

Blumenthal, 1999; Johnson & De Haan, 2015). Nonetheless, under-myelinated connections in the young human brain are still capable of transmitting signals.

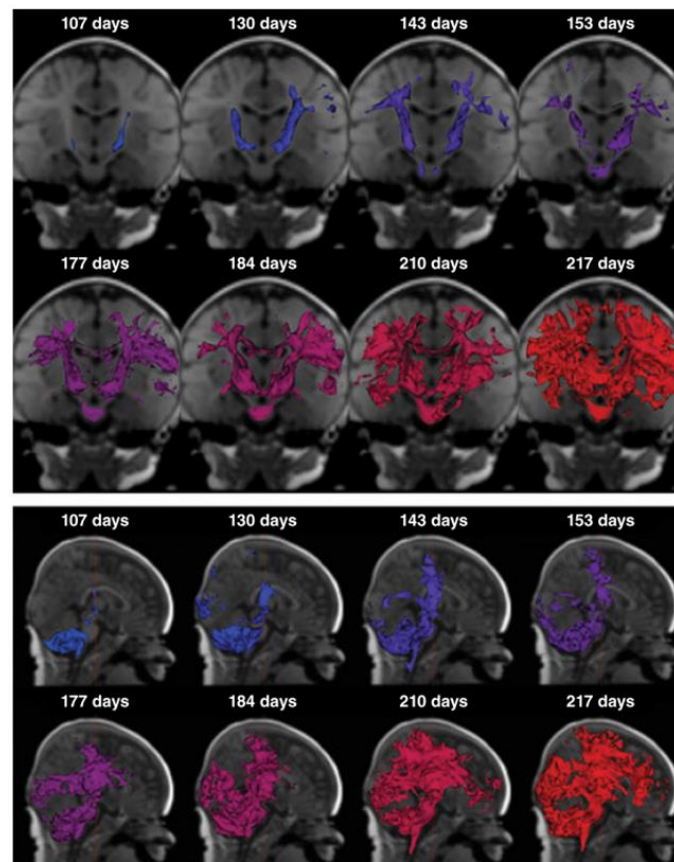


Figure 7 - The expansion of myelinated fibers over early postnatal development as revealed by a new structural MRI technique. *From Johnson & De Haan, 2015.*

1.2.3 *Brain metabolism: glucose uptake*

Finally, studies using positron emission tomography (PET) techniques reported a sharp rise in overall resting brain metabolism (uptake of glucose) after the first year of life (e.g. Chugani, Phelps, & Mazziotta, 1987). This increasing rate of glucose metabolism around 4-5 years of age for some cortical areas even exceeds adult levels (approximately 150% above adult levels).

1.2.4 *Regressive events*

Human postnatal brain development is also characterized by dramatic **regressive events**. Among them, the glucose metabolism described above reduces to adult levels after about 9 years of age for most cortical regions. For some brain areas, this **decrease of glucose uptake** may be caused by the decline of the synaptic contacts. Indeed, the rise in synaptic density is also subsequently followed by a period of **synaptic loss** (i.e. **synaptic pruning**), for which the timing varies

between cortical regions [Figure 8]. For example, synaptic density in the visual cortex returns to adult levels between 2 and 4 years, whereas regions of the prefrontal cortex do not reach adult levels before 10-20 years of age (Huttenlocher, 1979). It is suggested that the initial overproduction of synapses plays a critical role in the apparent plasticity of the young brain. This rise and fall pattern of synaptic density is likely unique to primates (M. H. Johnson & De Haan, 2015). The synaptic pruning observed between 2-20 years old is accompanied by a slight **decrease in neuronal density** (Huttenlocher, 1979). In contrast to the synaptic density, this event is largely distributed among animals. In rodents and other vertebrates neuronal loss is shown to be even more significant than in humans.

Finally, numerous **neurotransmitters** also show this additive and subtractive developmental pattern, specifically the excitatory intrinsic glutamate transmitter, the inhibitory intrinsic transmitter (GABA) and the extrinsic serotonin transmitter (M. H. Johnson & De Haan, 2015).

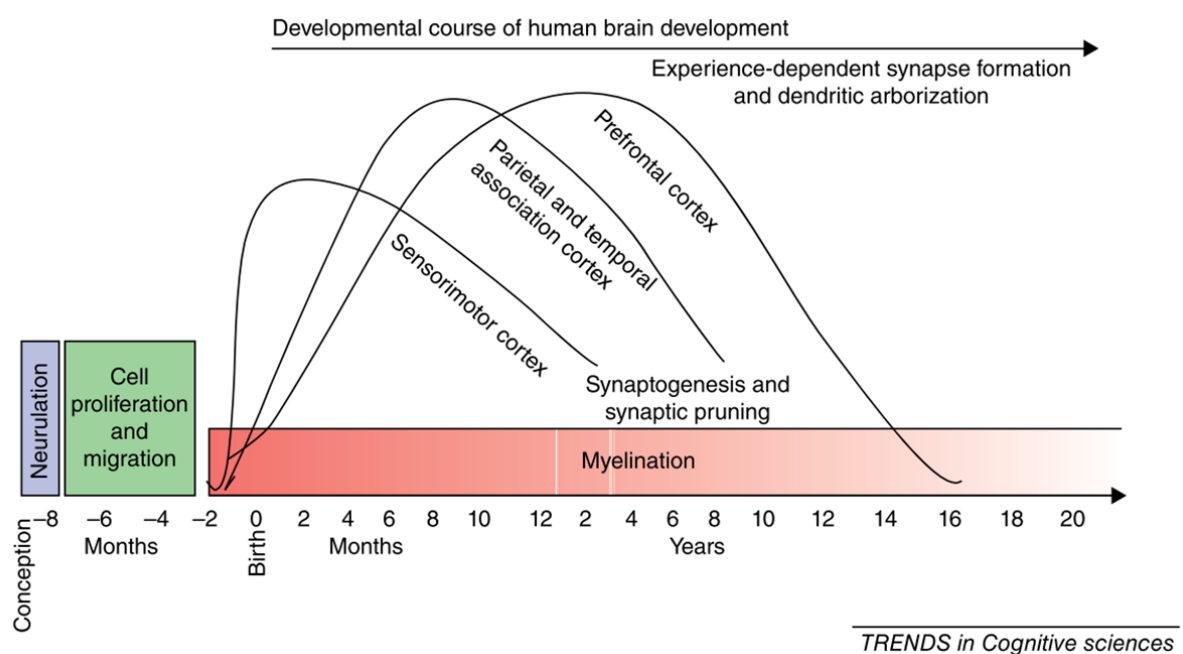


Figure 8 – Illustration of the approximate timeline for some of the most important events in human brain development, including the characteristic rise and fall of synaptic density. *From Johnson & De Haan, 2015*

To conclude, it is important to stress that these **rise and fall developmental patterns** first do not encompass all events (e.g. myelination is not affected by a subsequent fall), and second that they are **dynamic processes**; in other words that these processes are not distinct and separate progressive and regressive phases. Alternatively, the dynamic interplay between progressive and regressive events largely account for the relatively rapid brain growth in the first 2 years of life. At this age, an infant's brain has achieved 80% of its adult weight. Around 5 years of age, brain size is approximately 90% of adult size and at age six years, 95% (Lenroot & Giedd, 2006).

Finally, one should keep in mind that these overall developmental events are highly variable across individuals. Indeed, current sophisticated brain imaging techniques (e.g. MRI) evidenced that there is considerable variation in structure and function in normal adult subjects (e.g. Tramo et al., 1995). For example, healthy children of the same age may have 50% differences in brain volume, which supports the need to be cautious when linking functional implications with absolute brain size (Johnson & De Haan, 2015; Lenroot & Giedd, 2006). The lateral ventricle volumes are reported as showing the highest variability of brain morphometric measures. Sexual dimorphism is also markedly observed during brain development, for example there is a higher increase of white matter with age in males than in females (Giedd & Blumenthal, 1999). Amygdala volume increases significantly with age only in males whereas hippocampal volume increases significantly with age only in females (Lenroot & Giedd, 2006).

The following section will focus specifically on the development of the brain systems involved in long-term memory formation. Clearly, elucidating the postnatal structural development of the brain components assigned to memory formation would allow a greater understanding of the emergence of the different forms of memory across the life span.

2 The mnesic system and its development

The neural basis underlying memory began to be elucidated in the 1950s with clinical stories of single individuals, such as the well-known case of HM (a pseudonym for Henry Gustav Molaison). HM suffered from severe memory loss subsequently after treatment for medication-resistant epilepsy in which he had a surgical resection of a large part of both temporal lobes (Scoville & Milner, 1957). Following the surgery, HM was unable to form new memories for events and had no recollection of his everyday life. He no longer remembered people he just met, couldn't recognize the route that led to his new house, and read books or magazines over and over again without any sense of familiarity. Surprisingly, however he remained fully capable of recalling remote events from well before his surgery (e.g. incidents from his early school years) but for the two years preceding the operation, the memories he had were only vague (Milner, 1968). Also unexpectedly, he was able to learn and form new memories for new perceptual-motor skills (e.g. acquire a mirror-reading skill with a tracing rate comparable to that of control subjects). Finally, his comprehension of language was undisturbed; he could even get the punchline of jokes that played with semantic ambiguity (Milner, 1968). Among others, this unfortunate clinical case had a historical repercussion on the understanding of memory. This specific pattern of amnesia associated with temporal lobes removal illustrated that memory is perhaps not a unitary function as previously thought and that different types of memory might be

subserved by different brain systems (Tulving, 1972). Specifically, noted clinical cases highlighted the distinction between **declarative memories** (also often called **explicit** memories, or relational memories) and **procedural memories** (also often called **implicit** memories, **non-declarative** memories or unconscious memories) (Cohen & Squire, 1980).

On one hand, declarative memory refers to memories that can be brought to mind and that can be explicitly and consciously recollected (“*knowing what*”). It primarily relies on the medial temporal lobe (MTL).

On the other hand, procedural memories are typically expressed as changes in perceptual and motor performances, for which it is difficult to put words on (“*knowing how*”). They are mediated by different and sometimes overlapping brain circuits, such as the striatum, cerebellum, basal ganglia and brain stem.

“Explicit memory is roughly equivalent to ‘memory with consciousness’ or ‘memory with awareness’. Implicit memory, on the other hand, refers to situations in which previous experiences facilitate performance on tests that do not require intentional or deliberate remembering” (Schacter, 1989)

Other researches refined the knowledge about long-term memory by implementing additional divisions within the declarative and procedural forms of memory (Squire & Zola-Morgan, 1991). On one hand, *declarative memory* encompasses **semantic memories** for decontextualized facts and general knowledge which could be defined as an individual’s mental thesaurus that comprised organized knowledge, meaning and referents for words, symbols or dates, rules, formulas, concepts, etc. (e.g. the name of a monument, the date of birth of a celebrity, etc.); and **episodic memories** for faithful personal events that can be dated temporally and that spatially and temporally rely on the context during which those memories were formed [Figure 9] (Tulving, 1972). Thus for this latter type of memory, an individual should be able to recall details about “*what*”, “*where*” and “*when*” the event took place as well as being capable of giving precise information about the specific “*source*” of the event. Episodic memory also enables an individual to be consciously aware of the event, as well as being gifted to mentally travel through the episode.

On the other hand, *procedural memory* involves a range of different skills including **motor learning** (e.g. learning to play a musical instrument), **conditioning** (e.g. associational learning), **visual discrimination learning** (e.g. puzzle-solving skills) and **perceptual priming** (i.e. priming cues) [Figure 9] (Squire, 1992; Squire & Zola-morgan, 1991).

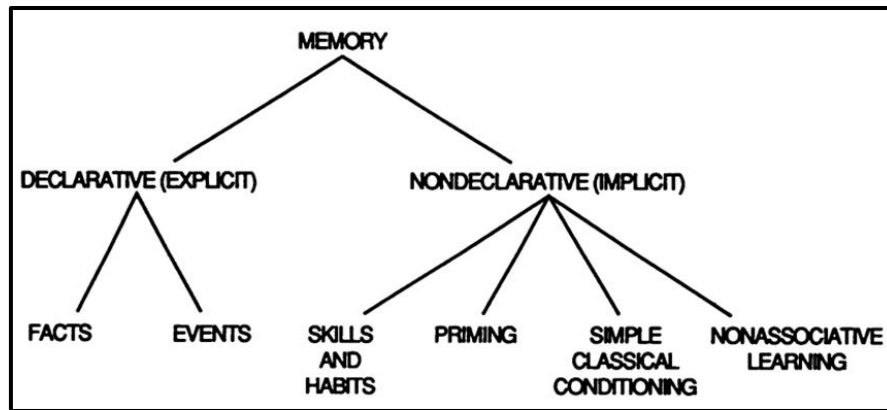


Figure 9 – Classification of long-term memory developed by Squire and Zola-Morgan. Declarative (explicit) memory refers to conscious recollection of facts and events and depends on the integrity of the medial temporal lobe. Non-declarative (implicit) memory refers to a collection of abilities and is independent of the medial temporal lobe. Non-associative learning includes habituation and sensitization. *From Squire & Zola-morgan, 1991*

Nevertheless, this classification is not universally admitted and there are currently still some debates regarding different memory models, different definitions about the various memory types and even different terms to adopt (e.g. for some authors autobiographic memory refers to episodic memory while for others it is a distinct type of memory).

For instance, another influential theory (**the complementary learning system, “CLS”**) postulates that declarative memory is primarily made of two differentially specialized memory systems that do not differ in the kind of information processed but rather in the time of integration: a system specialized in rapid acquisition of specific events (~episodic memory) and a system that slowly extracts and integrates statistical regularities across those events (~semantic memory) (e.g. McClelland, McNaughton, & Reilly, 1995). This theory holds that the slow integration process of arbitrary new information is necessary to prevent “catastrophic interferences” with existing knowledge networks (represented by synaptic connections among neocortical neurons). Therefore, this theory postulates that interwoven experiences are specifically required for the formation of “semantic memories” (note that general knowledge that makes up semantic memory can be based on multiple learning events). However, it admits that new information which is consistent with prior knowledge can still be learned rapidly without interfering with the existing structures (McClelland, 2013; Sharon, Moscovitch, & Gilboa, 2011).

“If empirical dissociations were the criterion for differentiating memory systems, our field of memory might soon become a taxonomic science resembling botany” (R.G. Crowder, *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving*)

Finally, it is a matter of debate whether applying the same frameworks to early-developing human beings is really appropriate. For developmental scientists, the distinction between declarative and

non-declarative memory and their traditional definitions are untenable from developmental and comparative perspectives (Rovee-collier, 1997). A major reason arises from the fact that it is not clear how to assess conscious recollection of information in an organism that is not yet able to talk. Thus, researchers must indirectly infer memory from changes in infants' behavior as a function of their experience. Many researchers tried to determine the earliest age at which infants could perform tasks qualified as measuring declarative memory. Consequently, this paradigm-driven approach essentially focused on determining the start point of performance on a given task but put to rest the neuroscientific processes underlying developmental changes.

“One of the unmistakable characteristics of an immature science is the looseness of definition and use of its major concepts. In experimental psychology, [...] we can measure our progress by the number and generality of empirical facts and the power and scope of our theories, and we can assess the lack of progress by the degree of ambiguity of our most popular terms.” (Tulving, 1972)

2.1 Procedural (implicit) memory

2.1.1 *The neural substrate of procedural memory*

Procedural memories rely on various brain areas. For example, motor skills and conditioning are supported by the basal ganglia, the cerebellum and other motor structures. Instead, perceptual priming (i.e. the increased ease with which a perceptual stimulus is processed) has been linked to the sensory cortex.

2.1.2 *Development of the neural bases of procedural memory*

Most researchers agree that procedural memory is in place within the first months of postnatal life, with only little development in performance beyond 3 years of age.

For instance, 3-year-olds demonstrated evidence of implicit priming by identifying blurred images of animals they have seen on a children's book 3 months earlier quicker (perceptual facilitation) than control children who had not previously read the book (Bullock Drummey & Newcombe, 1995). In contrast, no evidence of explicit recognition of the pictures of animals could be established. In a second experiment, the authors showed that explicit memory significantly improved between 3 and 5 years, whereas priming did not. Investigators concluded that procedural memory develops earlier than declarative memory in humans.

Neuroimaging studies revealed that areas of the striatum, cerebellum, basal ganglia and brain stem are activated during non-declarative memory tasks and are functionally mature early in life (Thomas et al., 2004). However, children (7- to 11-year-olds) showed greater subcortical

activation while adults mostly activated cortical regions during implicit memory tasks, supporting evidence for developmental differences in the recruitment of brain structures (Thomas et al., 2004).

2.2 Declarative (or explicit) memory

2.2.1 *The neural substrate of declarative memory*

In adult humans, declarative memory depends on a multi-component network recruiting temporal structures, namely the **medial temporal lobe** (MTL). The MTL operates with cortical systems to maintain long-term memories and to retrieve those memories. This **temporal-cortical network** consists of anatomically related structures [Figure 10]: the **hippocampal system** (i.e. CA1, CA2, CA3 and CA4 subfields, dentate gyrus and subicular complex, [Figure 12]), the adjacent **perirhinal**, **entorhinal** and **parahippocampal cortices**, along with the cortical **diencephalon** (i.e. thalamus and hypothalamus) and regions of the **prefrontal cortex** (e.g. Cohen & Squire, 1980; Eichenbaum & Cohen, 2001; Squire & Zola-morgan, 1991).

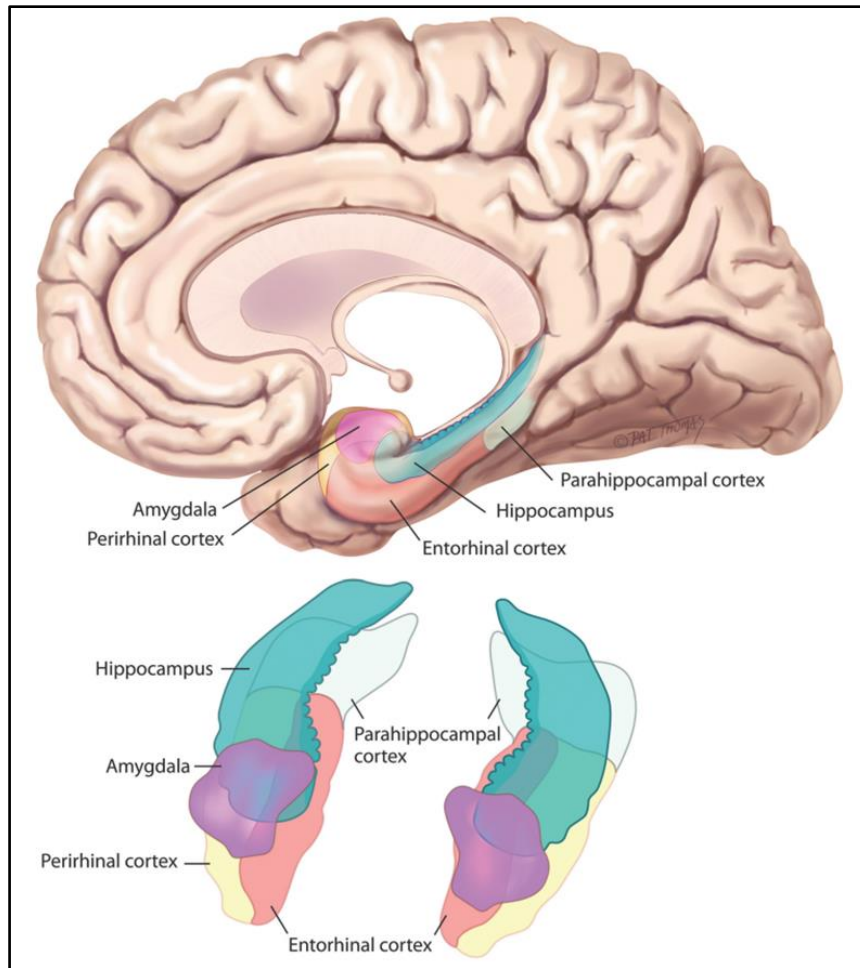


Figure 10 – Illustration of the medial temporal lobe. It consists of the hippocampus (blue-green: CA fields, dentate gyrus and subicular complex) superiorly and the parahippocampal gyrus inferiorly. The entorhinal (brown) and perirhinal (yellow) cortices form the medial and lateral components, respectively of the anterior portion of the parahippocampal gyrus, while the parahippocampal cortex (off-white) forms the posterior portion. *From Raslau et al., 2015 (adapted from Purves D. et al. *Principles of Cognitive Neuroscience*, 2008).*

The prefrontal cortex and the diencephalon are responsible for the **retrieval process of memory formation**. Diencephalic structures, such as the medial thalamus and mammillary bodies, play a critical role in the process of recognition, for example (e.g. Aggleton, Dumont, & Warburton, 2011 for a review).

For its part, the MTL appears to be particularly involved in the **consolidation** process of information in higher-order associational cortices, through feedback projections. More precisely, the hippocampus lies at the end of a cortical processing hierarchy and is the ultimate recipient of convergent inputs from all sensory modalities [Figure 11] (Lavenex & Amaral, 2000, see also box 4 in the general introduction of this thesis). Indeed, it receives only highly polysensory integrated information originating from associational connections within the perirhinal, parahippocampal and entorhinal cortices. These areas themselves have already significantly integrated information from unimodal and polymodal cortical inputs (e.g. temporal, frontal, parietal lobes). The information processed in the hippocampus is then projected back to its

adjacent cortices (entorhinal cortex first and perirhinal and parahippocampal cortices thereafter), that in turn spread out to the neocortical association areas that gave rise to their inputs (Lavenex & Amaral, 2000). Thus, these **cortical regions** would serve as the **final repositories of highly selective neurons that are the neuronal support of long-term memories**, as mentioned in the introduction of this thesis (Squire & Alvarez, 1995).

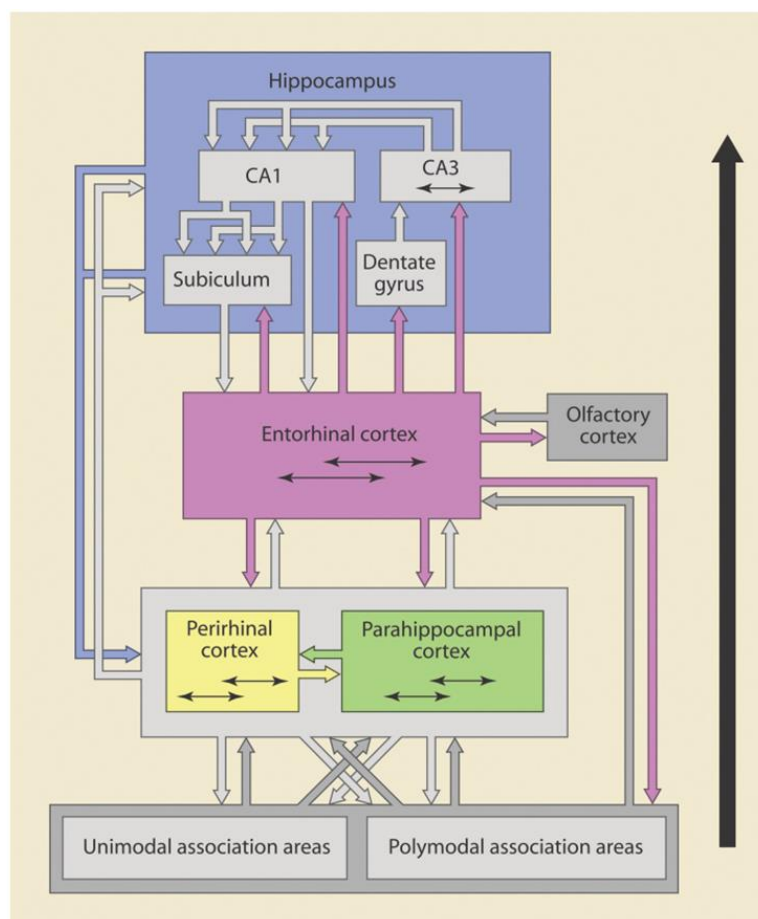


Figure 11 – The net flow of information arises from the perirhinal and parahippocampal cortices to the entorhinal cortex and then to the hippocampal formation (CA fields, dentate gyrus and subiculum). Considerable information processing occurs within and among the subregions of the parahippocampal gyrus before hippocampal formation involvement. *From Raslau et al., 2015 (adapted from Lavenex & Amaral, 2000).*

The MTL circuitry would only play a temporary role in the establishment of long-term memories within the neocortex (Squire, 1992). According to some authors, this would concern both semantic and episodic memories as recapitulated in the so-called **Standard Model** (Squire & Alvarez, 1995). Note that there is an alternative theory, postulating that only semantic information no longer depend on the hippocampus at long-term (Nadel & Moscovitch, 1997). This theory is defined as the **Multiple Trace Theory**. The debate is still not closed today.

Moreover, the **complementary learning system** (CLS) framework introduced earlier, similarly postulates that novel information would always depend initially on the hippocampus, and then

gradually be integrated into networks within the neocortex (McClelland et al., 1995). As previously stated, this mechanism is supposed to be slow and gradual to avoid incorporation of novel information in the neocortex interfering with existing knowledge already hosted by the neocortex (“*catastrophic interference*”). Consequently, new memories (i) would first be stored via synaptic changes in the hippocampal system (i.e. synaptic consolidation); (ii) these changes would then support reinstatement of recent memories in the neocortex; (iii) neocortical synapses would change slightly on each reinstatement and; (iv) remote memory would finally be based on accumulated neocortical changes (McClelland et al., 1995). Accordingly, memories would become hippocampus-independent over periods of time that range from weeks to years. Recently, the CLS model has been refined and claims that neocortical circuits can also rapidly acquire new information provided that they are consistent with prior knowledge (McClelland, 2013). In this case, the information would be rapidly and directly stored in the neocortex independently of the hippocampus.

What are the roles of the different protagonists of the consolidation process?

As explained above, the consolidation process of declarative memory is predominately a slow process that relies on a complex temporal-cortical network. But what do we know about the role of the different components of this circuitry?

Although the cortices adjacent to the hippocampus are certainly more than interfaces funneling information within the neocortical-hippocampal loop, the exact functions of each component of the MTL system in the memorization process are still not totally well established and sometimes give rise to conflicting and contradictory results, as it will be reviewed below.

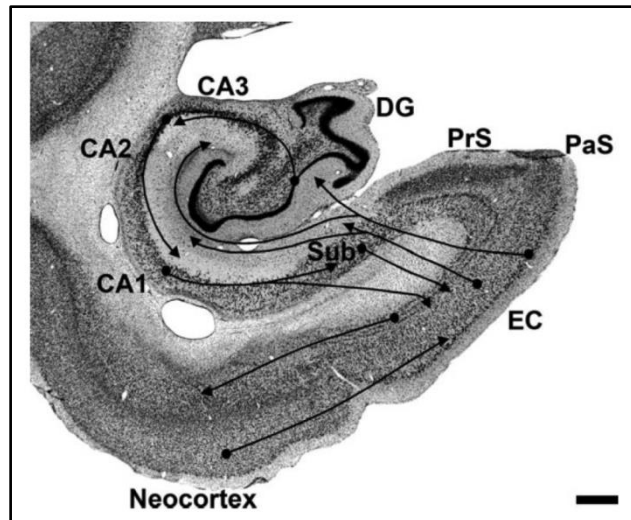


Figure 12 – Schematic representation of the main serial and parallel pathways within the hippocampal formation and surrounding the hippocampal system. EC: entorhinal cortex; DG: dentate gyrus; CA3, CA2, CA1: fields of the hippocampus; Sub: subiculum; PrS: presubiculum; PaS: parasubiculum. Scale bar=1mm.

But in the first place, it is largely admitted that the **hippocampal system** plays a crucial role in **relating or combining together information** from multiple sources (e.g. Eichenbaum, Schoenbaum, Young, & Bunsey, 1996; Squire & Zola-Morgan, 1991). For example, associating a stimulus within a spatial and temporal context - thus representing a new event (i.e. **episodic memory**) - or relating different elements to be remembered as a pair (e.g. a name and a face) are strongly supposed to be hippocampus-dependent (Squire, Stark, & Clark, 2004). To do so, each field of the hippocampus has attributed a specific function attributed to it [Figure 12].

Since O’Keefe’s pioneering discovery that CA1 *place-cells* fired according to a rat’s location in the environment (O’Keefe & Dostrovsky, 1971), it is now largely admitted that the **CA1 field** of the hippocampus is implicated in **spatial memory**. Moreover, CA1 field would play a critical role in integrating several inputs into a **unitary crystallized representation** (Ribordy, Jabès, Banta Lavenex, & Lavenex, 2013; Rolls, 2010). For its part, the CA3 field and its associated network is assumed to store the different parts of a representation separately, so that activation of only a small part of the network would facilitate the retrieval of the whole representation (a phenomenon known as *pattern completion*⁵) (Gold & Kesner, 2005). The different parts of a representation are associated into a network but remain distinct from each other, allowing a large number of memories to be stored but avoiding interference between memories (Rolls, 2010). This mechanism would implement temporal order in memories and prevent an individual from memory loss (Morris, 2007). Finally, the **dentate gyrus** and its projections to CA3 have been shown to be responsible for the *pattern separation*⁵ phenomenon (Morris, 2007).

⁵**Pattern completion:** ability to retrieve a complete memory when only a part of the information is available. Lesion or inactivation of NMDA receptors in the CA3 region of rodents disrupts their ability to retrieve information necessary to navigate in the environment when only a subset of environmental cues present during training is available during testing (Gold & Kesner, 2005).

Pattern separation: encoding of information as distinct from similar or closely related information. Damage limited to the dentate gyrus disrupts the ability of rats to distinguish or separate sensory inputs that are very similar, such as the spatial information about two food dishes located very close to each other (Gilbert, 2001). There is recent evidence of the implication of the dentate gyrus and its projections to CA3 in the process of pattern separation in humans (Bakker, 2008).

Whilst essentially all theories agree to say that the **hippocampal system** is strictly mandatory for the formation of **episodic memories**, there are conflicting viewpoints about its implication into the formation of semantic memories. This conflict first arose from patients with damages or even a resection of this area (e.g. H.M.) who exhibited only mild impairment on semantic knowledge tests (e.g. Milner, 1968; Vargha-Khadem, 1997). Instead, it appeared that the severity of their impaired semantic knowledge was related to the extent of anterolateral temporal damages (Schmolck, Kensinger, Corkin, & Squire, 2002). In parallel, some researchers showed that knowledge about general facts or judgments of **familiarity** about recently presented items, was rather supported by the cortices adjacent to the hippocampus, particularly the **perirhinal** and **entorhinal** cortices (Brown & Aggleton, 2001). Specifically, the anterolateral entorhinal cortex and the perirhinal cortex would function as a hub allowing integrated conjunction of features. The perirhinal cortex would be involved in the recognition process of an object as a known identity. Finally, others proposed that the **anterior temporal lobe** serves as a representational hub for linking associative semantic knowledge together and may be critical for supporting unique semantic associations (Patterson, Nestor, & Rogers, 2007; Sharon et al., 2011).

According to these viewpoints, the episodic and semantic components of long-term declarative memory are partly dissociable, with only the context-rich episodic component being fully dependent on the hippocampus and the context-free semantic component being definitely stored independently of episodic memories (Squire et al., 2004 for a review).

Unlike this assumption, few other studies revealed that patients with damage limited to the hippocampus had marked deficits in knowledge about general facts that occurred after the onset of the lesions (Squire et al., 2004 for a review). Therefore, in an opposed viewpoint, **semantic and**

episodic memory seem both dependent on the **hippocampal system**. An explanation to the impressive semantic knowledge of patients with developmental amnesia, who sustained limited hippocampal damage early in life and suffered from a severe loss of episodic memory, could arise from functional reorganization or compensation through learned strategies (Vargha-Khadem, 1997).

To summarize, nowadays most of the components of the medial temporal lobe are assigned a function in the memorization process, although there are still some debates. Recent neuroimaging studies along with stereo-electroencephalography (single cells recordings from pharmacology-resistant epileptic patients) opened a fruitful area of research that will undoubtedly refine our knowledge about the exact role that should be attributed to the numerous actors of memory formation. However, while those studies are difficult to be carried out in human adults or even in nonhuman primates, exploring the ontogeny of the human memory through the functional development of the underlying brain components is an even harder challenge.

2.2.2 *Development of the neural bases of declarative memory*

The ontogeny of the different components of the declarative memory of humans is not well elucidated. The lack of neuroimaging studies on babies and young infants are largely responsible for this caveat. Indeed, it remains very difficult to image infants as they engage in cognitive tasks. These difficulties stem in part from the low compliance of young subjects, the impracticality of studies involving extensive training of infants and the impossibility of giving verbal instructions (Dehaene-Lambertz & Spelke, 2015). Further difficulties arise from the small size and immaturity of the infant brain, which affects the characteristics of MRI images and of electrical activity. Actual knowledge about the emergence and development of the structures described above are partly based on studies conducted on non-human primates as well as on behavioral studies. The theory is simple: the onset of a cognitive trait should coincide with the emergence in functionality of the brain structure responsible for this trait.

Originally, Schacter and Moscovitch speculated that the MTL is not functional for the first or first and second years of life (Schacter & Moscovitch, 1984). They hypothesized that infants initially rely on procedural memories, with precocious declarative memory emerging only when the neural system matures (around 8-10 months of age). This proposal was further supported by studies conducted on infant monkeys, showing that they failed to learn a cognitive task (namely the delayed non-match to sample task, i.e. DNMS) until they were 4 months old and did not reach adult levels of accuracy by the end of the first year. In contrast, infant monkeys of 3 or 4 months of age were able to learn “visual habits” as easily as adults (Bachevalier & Mishkin, 1984). A

similar experiment was conducted on human infants (12-32 months old) and adults, and demonstrated that the dissociation in ability of infants on the DNMS task versus discrimination tasks closely resembled the dissociation previously reported with infant monkeys (Overman et al., 1992). Twelve to fifteen month-old infants required extensive training to perform the DNMS task (no infants did so until 15 months) and participants even had troubles reaching adult levels of performance at 6 years of age. These findings fortified Schacter and Moscovitch's assumption that the performance of human infants on "*early-system tasks*" (i.e. habituation/novelty preference and conditioning) precedes their ability to perform "*late-system tasks*" (e.g. DNMS, object search, cross-modal recognition).

In a similar theoretical vein, Nelson (1995) suggested that the striatum, cerebellum and brain stem - which are already functional at birth as reviewed above - allow very young infants to demonstrate procedural learning such as visual expectation, operant and classical conditioning tasks (Nelson, 1995). He also proposed that an immature form of explicit memory mainly reliant on the hippocampus - what he called a "*pre-explicit*" memory system - would already develop in the first few months and would confer some adult-like memory abilities, such as novelty preferences in the visual-paired comparison (VPC) procedure. He argued that maturation of the medial temporal lobe system, the surrounding cortical areas and the projections between these areas, and their targets in the prefrontal cortex are required for more sophisticated forms of explicit memory, such as solving DNMS tasks, deferred imitation and cross-modal recognition memory. He suggested that around 8-10 postnatal months, the performance on these tasks corresponds with a transition from reliance on a pre-explicit immature system to an adult-like declarative memory system, probably involving the dentate gyrus of the hippocampus.

In contrast, other studies contradicted these protracted MTL maturational theories. Researchers proposed that many of the medial temporal lobe components already develop early. According to these authors, the cells that make up most of the hippocampus (except the dentate gyrus) are formed in the first half of gestation and presumably journeyed to their adult locations by the end of the prenatal period (Seress, 2001). Moreover, at birth, the cytoarchitecture of the hippocampal formation already resembles that of an adult since neurons began to connect by the end of the second trimester.

Other opponents to the stepwise maturational theory argued that there is no evident qualitative change in the nature of infants' memory late in the first year of life. For instance, Rovee-Collier and coworkers claimed that only little evidence of a sudden improvement in performance after the proposed critical age period (around 8-10 months) is generally observed (Hartshorn et al., 1998). They showed that long-term memory improvement during early childhood increases

monotonically and suggested that it is rather a continuous and unitary process which does not consist of multiple systems that develop at different rates (Rovee-collier & Cuevas, 2009). Their proposal is based on their studies using the mobile or train conjugate reinforcement task [Figure 14 & Figure 15]. This procedure is however refuted by the proponents of the multisystem view, who are reluctant to accept it as reflecting declarative memory. They argued that the mobile conjugate reinforcement task is a motor skill and thus a procedural memory task reliant merely on the cerebellum and subcortical structures (see Johnson & De Haan, 2015 for a review).

Despite this disagreement, there is a larger consensus about a **protracted maturation** of one component of the MTL, namely the **dentate gyrus** of the hippocampus (e.g. Jabès & Nelson, 2015 for a review). Indeed, neuroanatomical studies conducted on monkeys first showed patterns of dissociated hippocampal development, with some areas (specifically the CA2, subiculum and their associated subcortical connections) developing first, whilst the CA1 field of the hippocampus followed by the CA3 field and finally the dentate gyrus along with their associated cortical projections, mature progressively and only later on (Bachevalier, 2015; Jabès & Nelson, 2015; Lavenex, Banta Lavenex, & Amaral, 2006). Although the granule cell layer (where excitatory neurons are generated) of the dentate gyrus is present by the 12th week of gestation, **neurogenesis** and **cell migration** continue until after birth [Figure 13] (Eckenhoff & Rakic, 1988). Indeed, at birth, the dentate gyrus includes only about 70% of the adult number of cells, thus roughly 30% of the neurons are generated postnatally (Lavenex et al., 2006). The **volumetric expansion** due to a considerable increase of **synaptic density** of the dentate gyrus is also dramatically delayed in human infants, beginning at 8-12 months after birth and reaching a peak at 16-20 postnatal months (e.g. Seress, 2001 for a review). Successively to a relatively short period of stability, excess synapses are pruned and adult levels are only reached at about 4-5 years of age (see Bauer, 2004 for a review). Moreover, the direct or indirect projections from the dentate gyrus to the hippocampal regions exhibit equally protracted developmental profiles. Consequently, the whole circuitry is presumably only functional when the connections are efficient (around 4-5 years in humans), which could explain the late emergence of certain types of hippocampus-dependent memory functions (e.g. **spatial** and **non-spatial relation memory**, **episodic memory**, etc.) (Jabès & Nelson, 2015).

Furthermore, **inhibitory interneurons** from the dentate gyrus resemble those of adults very late in development (not before 2- to sometimes 8-years of age) (Lavenex et al., 2006; Richmond & Nelson, 2007; Seress & Abraham, 2008 for reviews). However, it is assumed that these GABAergic interneurons play a critical role in memory processes, attention and broadly cognitive functions since these cells are implicated in **gamma oscillations** (e.g. Richmond & Nelson, 2007).

Indeed, the high-frequency gamma-band oscillatory activity (30-80 Hz) has been shown to increase during completion of cognitive tasks, such as DMS (delayed-match-to-sample) or working-memory tasks (Kahana, 2006). It has also been reported that increased gamma synchrony between the rhinal cortices and hippocampus during the encoding phase could predict subsequent recall (Kahana, 2006). Consequently, it can be easily claimed that the prolonged development of the dentate gyrus and of inhibitory interneurons precludes, to some extent, adult-like cognitive functions, such as memory or attention, during infancy and early childhood.

In addition to the protracted development of the dentate gyrus, the **prefrontal cortex**, largely responsible in the **retrieval process of memory formation**, also develops slowly and later on. For instance, the density of synapses in the prefrontal cortex increases roughly at about 8 months after birth and reaches a peak around 15-24 months of age. Synapses do not adopt adult morphologies until 24 months. Synaptic pruning to adult levels is extended until puberty (Bauer, 2004 for a review; Huttenlocher, 1979). Additionally, myelination in the frontal cortex continues into adolescence and adult levels of some neurotransmitters are not measured until the second or third decades of life (Bauer, 2004 for a review).

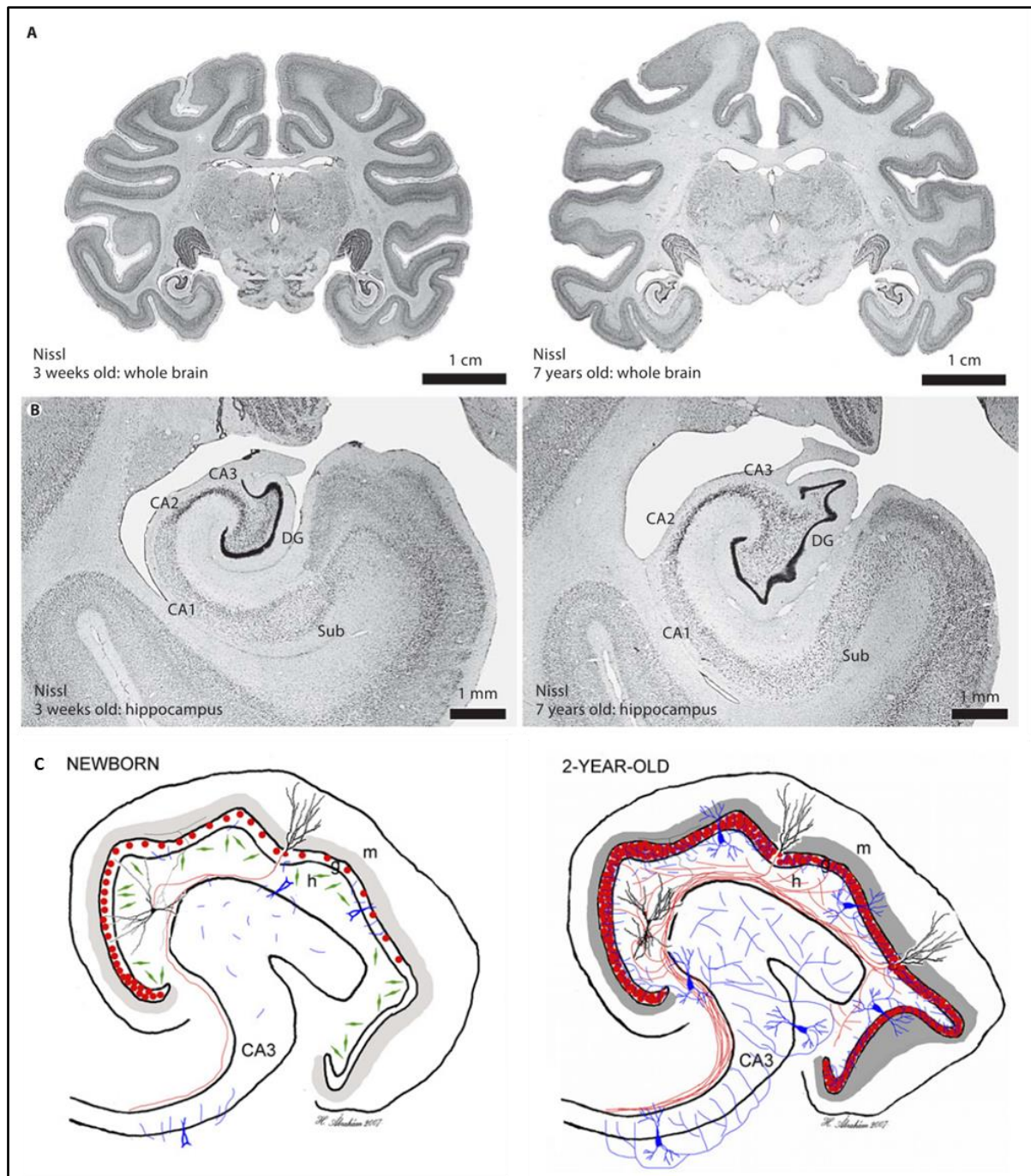


Figure 13- **A.** Coronal Nissl-stained sections through the brain of a 3-week-old (left) and a 7-year-old (right) monkey at comparable mid-rostro-caudal levels of the hippocampus. **B.** Higher-magnification photomicrographs at the same level illustrating four major subdivisions of the monkey hippocampal formation (DG= Dentate gyrus; CA3, CA2, CA1 = fields of the hippocampus; Sub=subiculum). Note in particular the large increase in size and complexity of the dentate gyrus between 3 weeks and 7 years of age. **C.** Schematic drawing of the developmental changes in the hippocampal dentate gyrus between birth and 2-years of age. **Newborns:** granule cells (red dots) express the calbindin binding protein (CaBP-D28k) within the granule cell layer. Hilus (h): migration of the granule cells toward their final position through the hilus. Mossy cells (excitatory cells, black) have few spines and no thorny excrescences, indicating sparse connection with granule cells. The mossy fiber bundle is also sparse (red line). The axon arborization of the inhibitory basket cells (blue) is sparse. **2-year-olds:** the axons of the granule cells (red dots) form a rich collateral network with the hilus (h). The mossy fiber bundles (red lines) innervate both inhibitory (blue) and excitatory cells (black). The morphology of the mossy cells (black) displays complex spines with thorny excrescences (which is a characteristic of mossy cells in adults). The axons of the basket cells (blue lines) form a rich network among excitatory cells. Note that the hippocampal dentate gyrus is much larger in the 2-year-old than in the newborn. Source: Figure courtesy of Laszlo Seress & Hajnalka Abraham. Adapted from Richmond & Nelson, 2007.

In conclusion, since the network that supports long-term declarative memory in humans involves medial temporal and cortical structures, the current predominant view emphasizes that its functionality depends on the maturational processes of its underlying components along with the functional maturity of the synaptic connections. It has been proposed that the hippocampal-subcortical circuits involving specifically the CA2 and subiculum mature the earliest, at birth or shortly after, and would subserve the emergence of the earliest “hippocampus-dependent” memory functions, what Nelson originally called “*pre-explicit* memory” (Jabès & Nelson, 2015). However, regarding the relatively late functional maturity of the prefrontal components and of the dentate gyrus of the hippocampus, as well as the prolonged maturation of the connections between those brain areas, declarative memory in young infants is supposed to be mostly rudimentary, with significant development occurring only from second year onwards (the most important morphological changes occur around the 4th year of age in humans).

Consequently, it can be argued that the increased improvements of long-term retention skills during childhood coincide with the gradual and prolonged maturation of the underlying components and networks. Nevertheless, it is important to remain cautious since those theories tend to transfer the organization of the memory systems of adults and non-human primates onto infants and young children, which may be a limitation. Neuroimaging studies should allow a better characterization of the brain memory systems that support cognitive functions during development. Indeed, memory tasks likely engage multiple memory systems. Thus, a lack of maturity of one distinct pathway may be masked in some tasks by a compensatory activity involving other pathways, especially in young children.

In the next section, we will try to link the development of the brain components that mediate memory formation with the development of specific memory functions during the first years of life.

CHAPTER II:

RELATION BETWEEN BRAIN AND LONG-TERM MEMORY DEVELOPMENT: A FOCUS ON DECLARATIVE MEMORY

The brain is the physical substrate for cognition and behavior. Nevertheless, correlating particular brain areas with specific cognitive functions is rarely straightforward, as evidenced by the complexity of interactions between brain regions, neurons, neurotransmitters systems, dynamic synapses, etc. This is even more challenging with a developing brain system. This section will try to enlighten the effects of maturation of the neural structures during development on the emergence of long-term **declarative memory**. Studying *when* and *how* declarative memory emerges during development is of central interest since it is suggested that the late development of declarative memory provides an explanation for the phenomenon of *infantile amnesia*⁶.

***Infantile amnesia:** children under 2-years of age are unable to form or store episodic memories for recall later in life. Consequently, adults do not have access to personal memories dating from their first two years of life.*

***Childhood amnesia:** during the next 3 to 5 years, only few episodic memories will survive until adulthood (e.g. Bachevalier 1992, Bauer 2007; Nadel & zola Morgan 1984, Newcombe, 2007).*

However, as it will be discussed below, some primitive forms of declarative memory seem to be already established early in life. A major challenge in the study of declarative memory is to determine whether infant memory tasks truly tap declarative memory or instead procedural memory.

“Tasks require recall (i.e. declarative memory) when the events/information tested are entirely novel at the time of the learning session, the subjects are not allowed to practice before they are tested and the actions/information are not modeled/repeated again after the initial learning session” (McDonough, Mandler, et al., 1995).

A little bit of history...

The study of memory development was initiated with **Piaget** in the 1950s. Piagetian theories made a strong prediction about a lack of mental representations of objects and events by 18-24 month of age. Instead, they proposed that infants were living in a “here and now” world that included physically present entities that had neither futures nor pasts. Indirectly, they implied that infants **lacked declarative memory**. Even for older children, who should have the capacity to construct mental representations, Piaget assumed that they required the cognitive structures to make those representations memorable. To illustrate his claim, he argued that seven year olds made errors in temporal sequencing, which suggested that they could not organize information temporally, and thus must be incapable at that age of telling a story from beginning to end appropriately.

In the late 1970s and early 1980s, researchers studied cognition in more **ecologically valid conditions**, and found that children as young as 3 years of age provided brief but well-organized reports of their own recent activities (K. Nelson & Gruendel, 1981, 1986). These findings opened the door for research on memory in children, even younger than age three years. It was claimed that if three-year olds already have well organized representations of past events, the capacity to construct those internal representation must have developed even earlier.

In this section, we will focus on the main cognitive paradigms and techniques used to assess the emergence and development of long-term memory (especially declarative memory) in human infants and young children. Some of the main paradigms will be reported below although many more should be added to the following overview. The hypothesized neural correlates underlying each cognitive function will also be reported.

1 Mobile conjugate reinforcement task: an operant conditioning task?

The mobile conjugate reinforcement task is at the center of a debate attempting to determine whether it taps into declarative or procedural memory.

This paradigm consists of two phases. During a learning phase, a ribbon connects the infant’s ankle to a mobile, so that the mobile moves when the infant kicks [Figure 14]. After a delay, changes in the kicking response are used as a way of assessing what the infant remembered, that is, more kicking than in baseline indicates the infant recognized the mobile (during baseline and memory tests, the ribbon is disconnected).



Figure 14 – Experimental design used with 2- to 6-month-olds in the mobile conjugate task, shown here with a 3-month-old. From left to right: (a) Baseline – the ankle ribbon and mobile are connected to different hooks, and kicks do not move the mobile; (b) Acquisition – kicks conjugately move the mobile via an ankle ribbon connected to the mobile hook; (c) Immediate retention/long-term retention test – the ankle ribbon and mobile are again connected to different hooks. During the test, infants who recognize the mobile, kick to move it even though they cannot. *From Rovee-Collier & Cuevas, 2009.*

A modified version of the task for older infants was developed, precisely a lever that moved a toy train for 2s each time the child pressed it [Figure 15]. Overall, Rovee-Collier and co. found a linear increase in the length of retention over the first 18 months of postnatal life (Hartshorn et al., 1998).



Figure 15 - Experimental arrangement used with 6- to 24-month-olds in the operant train task, shown here with a 6-month-old. Each lever press moves the toy train for 2 sec (1s for older infants) during acquisition; during baseline and all retention tests, the lever is deactivated, and presses do not move the train. *From Rovee-collier & Cuevas, 2009.*

This mobile conjugate reinforcement task has been largely labeled as a procedural memory task (e.g. Bauer, 1996; Mcdonough, Mandler, Mckeet, & Squire, 1995) and was attributed to the early

developing cerebellum and certain deep nuclei of the brainstem (e.g. Nelson, 1995). Rovee-Collier counter-argued that procedural learning is characterized by a gradual and incremental learning resulting from a reinforced and/or prolonged practice (Rovee-collier & Cuevas, 2009). She claimed that in the mobile task, learning is instead rapid rather than gradual or incremental. For instance, both 3- and 6-month-olds typically doubled or tripled their baseline response rate within a few minutes (Hartshorn et al., 1998) and exhibited significant retention for 3 days at 3 months of age (after 9 min of training) and for 5 days at 6 months of age (after 6 min of training).

Since the neurological correlates underlying this task have not been directly studied, the debate about whether structures from the MTL are recruited to solve the task or whether it is simply mediated by motor skill regions remains unresolved.

“Support for a developmental hierarchy has only been inferred from the memory performance of adults with amnesia on priming and recognition/recall tests in response to manipulations of different independent variables.[...] Implicit and explicit memory follow the same developmental timetable and challenge the utility of conscious recollection as the defining characteristic of explicit memory” (Rovee-Collier, 1997).

2 Recognition memory: Visual Paired Comparison (VPC) tasks

One of the earliest-emerging memory functions is visual recognition memory, precisely, the process by which an item is recognized as having previously been seen.

This procedure was originally developed by Fantz as a method of examining early perceptual functions (Fantz, 1956). In 1956, Fantz discovered that infants had a proclivity to respond to novel stimuli; that is, when given a choice between a familiar stimulus and a novel one, infants older than 2-3 months would look at the unfamiliar stimulus for longer, which could reflect their “preference” for this one over the familiar one. This discovery quickly blossomed into a major tool for studying early memory (Fagan, 1973) and has been a mainstay ever since.

In the Visual Paired Paradigm (VPC)⁷, the participant is **first familiarized to a stimulus** for a set period of time, and their memory is tested after a delay by presenting the **familiar stimulus alongside a novel one** (the participant remains entirely uninstructed about the subsequent “test”). During this test trial, **preferential looking** (measured as the total looking time toward one stimulus compared to the second) is used to assess **recognition memory** [Figure 16].

⁷This paradigm resembles the DNMS task, although great performance on the VPC task has been shown to emerge earlier in life than on the DNMS tasks, and monkeys with neonatal hippocampal lesions performed accurately on the DNMS task but not on the VPC task (Pascalis & Bachevalier, 1999). Similar results were provided by human patients, such as patient YR, who suffered from discrete hippocampal damage and who demonstrated impaired novelty preference after delays of 5 and 10s but who successfully recognized the same types of stimuli in DMS tasks following delays up to 10s (Pascalis, Hunkin, Holdstock, Isaac, & Mayes, 2004). Thus, it has been speculated that only the VPC task is fully reliant on the hippocampus and that DNMS tasks are perhaps mediated by alternate strategies that are independent of the hippocampus (Pascalis & Bachevalier, 1999).

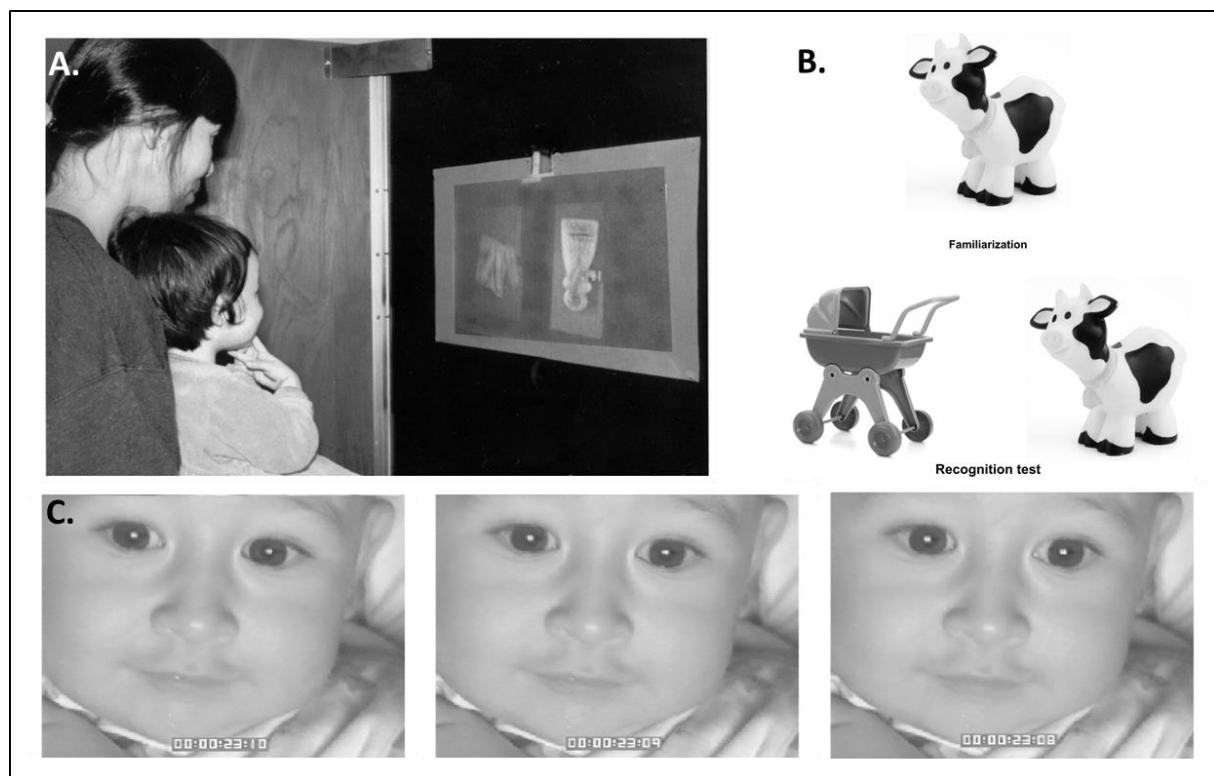


Figure 16 – (A) Example of the experimental setting. The infant is watching pictures; a camera located above the screen record the eye movements. (B) Example of stimuli used in the VPC task. The infant is presented with the sample for a familiarization period. Thereafter, the participant is confronted with the familiar stimulus and a new stimulus. The time spent fixating on each stimulus is recorded. Position of the novel stimuli in the recognition test is counterbalanced between left and right. (C) Videoframes (3/100 sec) of an infant's eye movements during the retention tests of the VPC task. Corneal reflections were used to determine positioning of gaze: the child looked at the stimulus on their right (left panel), in the center (middle panel) and at the stimulus on the left (right panel). Adapted from Pascalis & De Haan, 2003.

It has been found that 3-to-4-day-old infants looked preferentially at a novel face than at a familiar one after a two minute delay (Pascalis & De Schonen, 1994), and 3-month-olds (males but not females who were delayed in their face-processing ability) exhibited the same pattern over a 24-hour delay (Pascalis et al., 1998), indicating in both cases recognition of the familiar face. Similar results were established in 15-to-30-day-old rhesus monkeys (Bachevalier, Brickson, & Hagger, 1993) but not in monkeys with neonatal lesions of the hippocampal formation and parahippocampal cortex (after delays of 30s to 24h) (Pascalis & Bachevalier, 1999). This suggests that medial temporal structures significantly contribute to visual recognition memory at this very early age in monkeys. Likewise, adult amnesic patients did show visual recognition abilities after a 2-min delay but not after a 1 hour interval, when tested with similar preferential visual looking paradigms (McKee & Squire, 1993; see also Pascalis et al., 2004). In addition, human patients with developmental amnesia caused by bilateral hippocampal damage sustained early in life showed reduced looking at novel stimuli following delays of 30 sec and 120 sec, which indicates a deficient novelty preference especially when a delay is introduced between familiarization and test (Munoz, Chadwick, Perez-Hernandez, Vargha-Khadem, & Mishkin, 2011). Toddlers who experienced prolonged febrile seizures (which severely injured their hippocampus) demonstrated a hippocampal integrity-related degree of novelty preference after a 5-min delay (i.e. participants with the smallest mean hippocampal volumes revealed the most dramatic drop in novelty preference) (Martinos et al., 2012). The authors controlled and demonstrated that the impairments were not due to a transient effect of the seizure. Together, these results suggest the existence of a **primitive visual recognition memory** shortly after birth while a more adult-like recognition memory system would only emerge after a few months of life. These findings are in line with Nelson's assumption of a "pre-explicit" memory, which he believed to be different from procedural memory and already present after three postnatal months at least in human infants (Nelson, 1995). It also suggests that the hippocampus and surrounding tissues likely contribute to recognition memory and consequently that the MTL memory circuitry is functioning to some extent during early infancy. Researchers correlated the early emergence of incidental recognition memory with the **early maturation of the subiculum** – a hippocampal field - (Jabès & Nelson, 2015), whereas others suggested that novelty preference is, rather, mediated by the **perirhinal and parahippocampal cortex** – cortices surrounding the hippocampus - (Bachevalier, 2015), which are thought to be critical for familiarity judgments in adults (e.g. Brown & Aggleton, 2001). Somehow consistent with this latter idea, some authors advised to remain cautious about a conclusive role of the hippocampus in recognition memory since some variables in VPC task, such as the extended period of stimulus encoding during familiarization, could have overridden its incidental nature (Munoz et al., 2011). Consequently, an extensive familiarization could have

promoted strength-based recognition or even perhaps recollection, which logically rendered the task hippocampal-dependent. Another explanation could be that VPC, rather than providing a measure of incidental recognition, assesses an implicit, information-gathering process modulated by habituation, for which the hippocampus is also partly responsible, independent of its role in recognition (Munoz et al., 2011). Despite the need for additional data to determine the precise neural correlates of recognition memory across development, the hypothesis that at least part of the hippocampal circuitry is relatively mature soon after birth seems less controversial (Jabès & Nelson, 2015).

3 Basic relational memory functions

3.1 Spatial relational memory

Relational memory refers to the representation of arbitrary or accidental relations among the constituent elements of an event (e.g. the relation between different objects on a picture or the temporal relations between actions (i.e. deferred imitation)). Relational memory seems to emerge later than recognition memory, depending on the type of relational memory.

For example, **spatial relational memory** (or allocentric spatial memory), the ability to remember locations in relation to distant environmental cues (thus in the absence of local cues marking the location), is already present in 9-month-old macaque monkeys (Lavenex & Lavenex, 2006) but only emerges during the second or third year of life in human children. For instance, the ability of children to locate a hidden toy in a sandbox when only distant environmental cues are available has been reported to emerge at around 21 months of age (N. Newcombe, Huttenlocher, Drummey, & Wiley, 1998). Moreover, children have been shown able to find one reward location among four possible locations in an open-field arena in the absence of local cues by 25-39 months of age (Ribordy et al., 2013). In contrast, 18-23-month-olds did not perform well on that task and 25-42-months-old were not capable of discriminating three reward locations among 18 possibilities in absence of local cues. Together, this finding show that **allocentric spatial memory** is only **rudimentary** in children **under 3.5 years** of age and **absent before 2 years of age**. Given that spatial relational memory is a fundamental component of **episodic memory** (i.e. the “where” component of episodic memory), the authors suggested that the delayed ability to remember closely related spatial information of the environment coincides temporally with the phenomena of *infantile and childhood amnesia* (Ribordy et al., 2013). Nonetheless, the fragile improvement observed between 2 and 3.5 years of age could be linked to the maturation of the CA1 field of the hippocampus. Indeed, it has been suggested that the gradual maturation of CA1 might subserve

the incremental emergence of basic relational memory functions (such as spatial and nonspatial relational memory) observed between the second and third year of age in human children (Jabès & Nelson, 2015).

3.2 Deferred imitation

Deferred imitation is another type of basic relational memory. It corresponds to the ability to **reproduce actions previously performed by an experimenter**, that is, the ability to process relational information to create a **representation of novel unrelated actions**.

Precisely, in this paradigm, participants are first given a set of objects to play with for a baseline assessment of spontaneous actions. Then, an adult uses the objects to model a specific sequence of target actions to produce an interesting result (e.g. construction of a rattle) [Figure 17]. After a delay, the objects are re-presented to the participant and the number of target actions spontaneously produced and in the correct order is reported. This number is compared to the baseline to evidence memory.

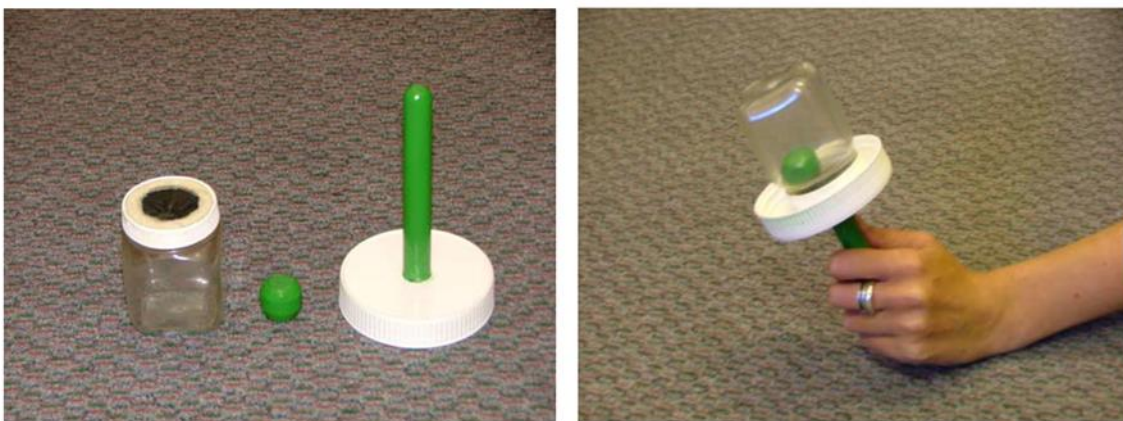


Figure 17 – Example of a deferred imitation task used with 18- to 30-month-olds: infants are shown how to make a rattle. There are three target actions: put the ball into the container, put the handle on the top of the container, and shake the handle. Infants don't have the opportunity to copy the actions immediately after they saw them being demonstrated. *From Herbert & Hayne, 2000.*

It has been demonstrated that deferred imitation already **emerges between 6 and 12 months of age** and gradually improves during the second year of life. Indeed, 6-month-old infants were able to reproduce after a 24-h delay, a sequence of three actions that they have previously seen a substantial number of times (Barr, Dowden, & Hayne, 1996). Conversely, when the experimenter only briefly demonstrated the actions with the puppet, **6-month-olds** exhibited only **immediate imitation** (i.e. they did not demonstrate retention after a 24h delay) whereas **12-month-olds** reproduced the actions significantly above chance after a **24h delay** (Barr et al., 1996). A striking

improvement in deferred imitation memory skills are demonstrated around **8-10 months**. For instance, Carver and Bauer showed that 9-month-olds as a group were capable of remembering target actions after a delay of one month (infants had a total of 3 pre-exposure sessions spaced in time (43h in mean between sessions) with the experimenter modelling the sequence of events twice in succession during each session and 1 re-exposure session one week later) (Carver & Bauer, 1999). However, only 45% of the infants recalled the temporal order of the events after this one month delay (Carver & Bauer, 1999) and no retention was observed over longer delays (Carver & Bauer, 2001). In contrast, 10-month olds seemed able to encode and retrieve some event representations over delays of up to 3 and even 6 months (thus retention was tested when they were 13 or 16 months old) (Carver & Bauer, 2001). This sharp improvement supports the idea that the system underlying long-term ordered recall emerges near the end of the first year of life.

But what type of memory is involved in deferred imitation tasks?

Overall it is proposed that deferred imitation tasks rely on declarative memory. Alternatively, one could have speculated that subjects simply learned a sensorimotor association between an object and an action by observation alone, so that presentation of the object might prime the production of the target actions (see Mcdonough, Mandler, Mckeet, & Squire, 1995 for an interesting discussion). In this case, deferred imitation might be similar to skill learning or conditioning which would not depend on the MTL circuitry. To determine whether deferred imitation is mediated by priming effects, by a stimulus-response association between an object and an action, or truly by declarative memory processes, amnesic patients with a damaged hippocampal system underwent deferred imitation tasks (that consisted of three actions as it is commonly used with infants) (McDonough et al., 1995). Following a delay of 24h, the impaired performance of those adult-onset cases of amnesia along with relatively similar deficiencies observed in patients with developmental amnesia (although some residual memory remained) (Adlam, Vargha-khadem, Mishkin, & De Haan, 2005) reinforced the assumption that deferred imitation is well and truly a hippocampal-driven paradigm. Hence, these findings suggest that the hippocampus is already functioning, in part, at the ages when children succeed on deferred imitation tasks.

4 Beyond long-term memory development: do infants and young children generalize?

As illustrated above, some forms of rudimentary declarative memory seem to be already implemented in early childhood. It has also been shown that across development, memory skills increase - sometimes linearly, sometimes brutally - as reflected by the delay between last learning (or practice) and testing, the number of exposures (or looking time) necessary to encode the information and obviously the level of accuracy reached during testing. We discussed above the maturational processes that might be responsible for these improvements and exposed how recent research came to the assessment that the human hippocampus may be more mature earlier in life than previously thought. Here, we would like to go further in our understanding of long-term memory development during childhood by questioning the **flexibility of learning**, that is, the ability to **generalize** learning to novel situations. Flexibility is fundamental since it allows an individual to extend their learning to novel instances of the environment or to novel situations of their everyday life. Moreover, being able to generalize suggests that learning is not limited to the specific context encountered during learning. Since the flexible use of memory is an ability attributed to the **hippocampal formation**, the onset of flexibility should strengthen our understanding of its maturational stages during development.

In the case of the mobile conjugate task, a study showed that 3-month-olds exhibited retention after a delay of 24h provided that the mobile was the same as that used during original training (Greco, Rovee-collier, Hayne, Griesler, & Earley, 1986). However, no retention was demonstrated if more than one object was substituted into the mobile during the delayed recognition test (Hayne, Greco, Earley, Griesler, & Rovee-collier, 1986). Similarly, if either the color or the form of the alphanumeric characters (e.g. “2” category vs “A” category, [Figure 18](#)) that composed the mobile were changed between learning and testing, no evidence of retention was apparent in 3-month-olds (but if learning encompassed those dissimilar stimuli, performances were preserved) (Hayne, Rovee-collier, & Perris, 1987). Likewise, 6- to 7-month-olds exhibited no forgetting after a delay of 14 days when the training and test mobiles were the same (Hill, Borovsky, & Rovee-collier, 1988). However, memory retrieval was precluded by changes in the mobile after this delay (Hill et al., 1988). Together, these studies show that the retrieval of memories in infants below 1 year of age is dependent on the context in which the memory was formed and that flexibility in learning is sparse at that age.



Figure 18 – Left panel: Reinforcement phase during initial training with an exemplar from the « 2 » category. Right panel: exemplar from the “A” category used as the memory prime. *From Hayne et al., 1987.*

Relatively similar patterns of impairment were reported in regard to deferred imitation. It has been shown that although 12-month-olds were able to accurately reproduce the three actions after a 24-h delay (Barr et al., 1996), their performance was disrupted if the puppet used to learn the sequence of actions was different from the one used during the testing phase (e.g. a pastel pink rabbit puppet vs a pale gray mouse puppet) (Hayne, MacDonald, & Barr, 1997). Similar findings were reported in 18-month-olds; however 21-month-olds performed equivalently whatever the degree of dissimilarity (Hayne et al., 1997; see also Hayne et al., 2000 for similar results). On the other hand, changing the test stimulus had no effect on performance at 18 months of age. Together, these findings suggest that over the course of their second year of life, infants become increasingly able to use novel cues to retrieve their memory of a prior event.

Finally, similar observations have been made with respect to recognition memory tasks. For example, a change in context (e.g. a different colored background, **Figure 19**) during testing disrupted recognition of the images (i.e. no novelty preference was reported) at 6 and 12 months of age, but not at 18 and 24 months of age (Robinson & Pascalis, 2004). However, if infants were familiarized with a picture on multiple backgrounds, infants as young as 6 months of age demonstrated recognition despite a change in background at test (Jones, Pascalis, Eacott, & Herbert, 2011). This supports the idea that dissociating an item from its context (or encountering an item in many different contexts) during encoding may be critical for visual recognition to be flexible during infancy.

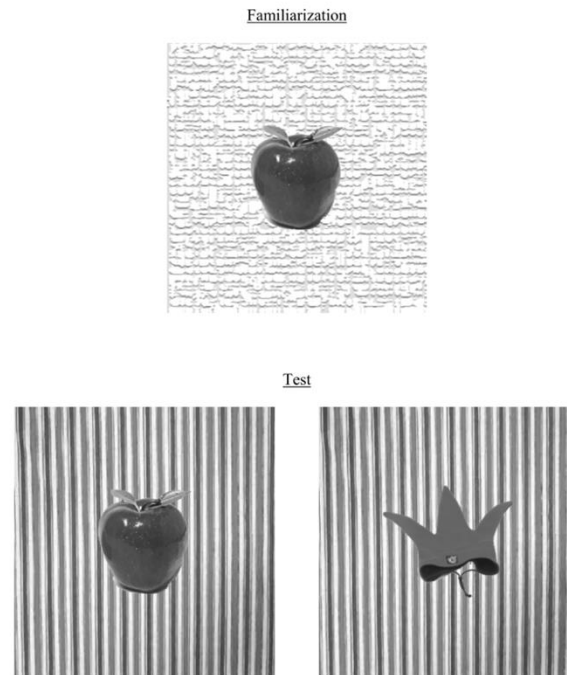


Figure 19 – Example of the stimuli used in the VPC paradigm to examine the effect of a background change on visual recognition memory. If the color of the background is changed from familiarization to test, 6- and 12-month-old infants showed no sign of remembering. *From Robinson & Pascalis, 2004.*

Additionally, when the familiarization and testing phases occurred in two different rooms, 6- to 9-month-olds exhibited impaired recognition of the images, but not 12- to 18-month-olds (Jones et al., 2011). Thus, representational flexibility in visual recognition memory, as measured with the VPC paradigm, is not present until 18 months of age regarding changes in color of the background but seems to emerge earlier – around 12 months of age - with respect to a change of experimental room.

Overall, these findings support the idea that before 1 year of age, memory is extremely **specific to the context** in which learning occurred and then **gradually becomes flexible**, allowing generalization to novel instances, conditions and situations (Jabès & Nelson, 2015). This **flexible use of memory** is thought to depend on the **hippocampal formation** as evidenced by clinic cases of adult patients with hippocampal lesions and monkeys with neonatal hippocampal lesions (including the dentate gyrus and a portion of the parahippocampal region) (Pascalis, Hunkin, Bachevalier, & Mayes, 2009). As already described above, the maturational stages of the different components within the hippocampal formation are time lagged. It might be that the progressive improvement in the ability to generalize is closely correlated to the development of the different components of the hippocampal formation, especially the CA3 field of the hippocampus and the dentate gyrus, respectively responsible for the pattern completion and pattern separation phenomena.

To summarize

Data from behavioral and neurobiological studies suggest that the differential maturation of specific brain circuits subserve the gradual emergence of declarative memory functions across development (Jabès & Nelson, 2015 for a review). Some hippocampal circuits are relatively mature at birth (e.g. the subiculum) and could subserve a “*pre-explicit*” form of declarative memory (e.g. recognition memory) until the end of the first year of life (~10 months). The protracted maturation of the CA1 field of the hippocampus and its related circuits might subserve the emergence of basic, context-specific relational memories (e.g. non-spatial and spatial relational memory) during the second or third year of life (Ribordy et al., 2013). This rudimentary form of declarative memory will gradually become flexible, complex and context-independent (e.g. representational flexibility) which may be a consequence of the later maturation of the CA3 field and the dentate gyrus as well as their associated circuits (Jabès & Nelson, 2015). Ultimately, when the declarative memory circuitry as a whole has reached a certain level of maturation, building and maintaining long term representations of everyday experiences should be possible.

In fact, the formation of a memory after a single event critically depends on the integrity and functionality of the MTL structures, in particular the hippocampus. This is coherent with the fact that the onset of long-lasting episodic memories and likewise the offset of childhood amnesia is thought to occur during the 3rd year of life (Hayne et al., 2000).

Nonetheless, some authors advised to remain cautious about the assumption that changes in cognitive skills may reflect a change in the maturity of the hippocampus or surrounding cortices. They argued that the changes in performance could also be due to the maturation of the pathways leading to these structures or simply to a change in the manner infants process information (Robinson & Pascalis, 2004). Behavioral data in conjunction with neuroimaging measures applied to very young infants (e.g. EEG (electroencephalography), fNIRS (functional near-infrared spectroscopy), MEG (magnetoencephalography)) will certainly further our comprehension of memory development in the close future.

CHAPTER III:

RELATION BETWEEN BRAIN AND SEMANTIC MEMORY DEVELOPMENT: A FOCUS ON WORD LEARNING

Introduction

In the previous chapter, we exposed in detail how brain maturation, especially maturation of the medial temporal lobe and its projections, may subserve the gradual development of long-term declarative memories. However, across the various protocols they utilized, authors rarely addressed the question of what kind of declarative memory was postulated (see Newcombe, 2015), possibly because the definitions do not properly apply to non-verbal populations as already mentioned. As introduced earlier, declarative memory consists of two separable subtypes: *episodic memory* (or personal context-related memories) and *semantic memory* (or decontextualized information about general facts and knowledge). We introduced the phenomena of infantile and childhood amnesia, namely the fact that memories for personal events (episodic memory) formed during the first years of life do not (or poorly) persist until adulthood. At first glance, those phenomena appear to be contradictory to the findings that are in favor of an early establishment of declarative memory (e.g. Jabès & Nelson, 2015). Therefore, it has been argued that declarative memory is primarily semantic whereas episodic memory would develop later, perhaps only after the second year of life (Newcombe, 2015). This claim, at least, reconciles early declarative memory establishment with infantile amnesia since semantic memories do not seem to be (or at least seem to be less) concerned by the infantile and childhood amnesia phenomena. But what accounts for the memory loss (or memory inaccessibility) of personal events?

Many explanations can be given, such as the developmental regressive events described in the first chapter, the immaturity of some brain components that mediate episodic memory, the inability for preverbal infants to verbalize the episodes (Simcock & Hayne, 2002), and obviously the uniqueness of the event, that is, the fact that it occurred only once in the infant's life. As reviewed previously, memory requires repetition to be maintained at long (or even very long) term. This repetition-driven mechanism strengthens the neuronal connections between neurons

involved in the integration of the information, allowing a better access to that information but also preventing those synapses from being pruned. In the case of episodic memory, adults and older children often narrate their personal experiences to other people, or sometimes actively retrieve their memory; for example, when looking at a photo of the event, or mentally replaying the episode, etc. Ultimately, even if the event occurred only once, those repetitive reactivations contribute, in a similar way to direct re-exposures to the sensory inputs, to the (re-)consolidation processes that are critical in the update and maintenance at long-term of the episodic memory. Pre-verbal children are basically not, or less, subjected to such “indirect reactivations”, which may also explain the discrepancy between pre-verbal children and adults to recall personal events at long-term. In contrast to episodic memory, repetition (in the sense here of direct re-exposure to the sensory information) is usually abundant in semantic learnings. Semantic learning consists of general knowledge, facts, dates, definitions, vocabularies, etc. Typically, across repetitive exposures to the information, an individual will isolate the specific information from the details that make up the context. The information will gradually become context-independent and the individual unable to recall *when* and *where* he/she acquired that knowledge as well as incapable of recollecting details about the context. As a consequence, semantic information that is repeated during childhood should survive the pruning and regressive events and should be durably maintained in the brain. As evidenced, one major constituent of semantic memory, **word learning**, emerges very early in life and survives until adulthood. Word learning represents the arbitrary mapping of phonological and semantic representations, that is, the knowledge about the meaning of words. Word learning is a fundamental building block in the acquisition of language.

1 Generalities about language acquisition

Language enables people to communicate between each other using a code composed of sounds, symbols and words to express meanings, ideas or thoughts. (Kuhl, 2004). The world’s languages consist of around 600 consonants and 200 vowels. Each language uses a unique set of about 40 distinct elements, called phonemes, which change the meaning of a word (e.g. “bat” vs “pat”) (Ladefoged, 2004). During the first 18 years of their life, a human being will acquire about 60,000 words, but the major steps of language acquisition mainly occur early in life. Language includes primary functions, such as perception and processing of incoming speech (i.e. discrimination, analysis and comprehension of the meaning of an incoming sound pattern) and production of meaningful speech output. Language also encompasses secondary functions such as reading and writing (M. H. Johnson & De Haan, 2015). So, the primary roles of language are to receive, comprehend (i.e. word reception) and produce audible sound patterns (i.e. word production).

1.1 Word reception

Word comprehension involves the brain area **Wernicke** – named after the neurologist who discovered it in the years 1870s -, located in the posterior portion of the left temporal lobe [Figure 20]. People with damage in this area can usually speak but their speech is often incoherent and deprived of sense (Johnson & De Haan, 2015 for a review). Wernicke's area is responsible for the processing of speech-relevant sounds (i.e. words) that one hears. It is highly connected to Broca's brain area, the structure specifically involved in word production, by the **arcuate fasciculus** (a large bundle of nerve fibers).

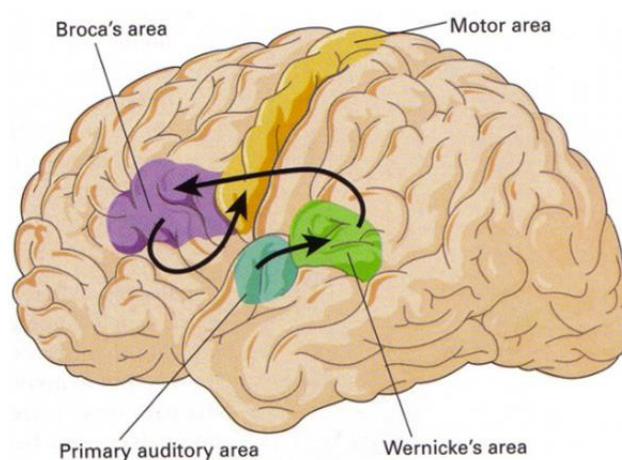


Figure 20 – Illustration displaying the Broca's and Wernicke's brain areas, as well as the primary auditory areas and the motor areas. They form a loop that is fundamental in language (reception and production). Not to scale. *From Courtesy of Neuroscience & Cloud*

The emergence of this primary language function occurs very early in life as attested by parental reports, observational studies and experimental studies on word comprehension. Generally, infants manifest the **first signs of word comprehension** at around the **middle of their first year of life** (Benedict, 1979; Bergelson & Swingley, 2012; Fenson et al., 1994; Friedrich & Friederici, 2011; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Tincoff & Jusczyk, 1999). At this stage, they start figuring out that **words can specifically refer to people or objects** from their surrounding environment (Bergelson & Swingley, 2012; Kuhl, 2004; Tincoff & Jusczyk, 1999; Oviatt, 1980). Typically, **word comprehension increases gradually and linearly** during the **second year of life** [Figure 21]. Children comprehend about 50 words at 11 months of age, and by 16 months, their receptive vocabulary repertoire has increased to 170 words (Fenson et al., 1994). From that age and until their sixth year of age, children learn an average of 5 new words a day; and thereafter more than 10 new words a day (P. Bloom, 2000). Importantly, studies reported large variabilities in the onset and development of word comprehension. Reports from parental

vocabularies inventories (MCDI, namely *MacArthur Communicative Development Inventories*) revealed that this variability increases dramatically with age.

1.2 Word production

As recently introduced, language production is mediated by a specific brain area, **Broca's area** [Figure 20]. This brain region was discovered in 1861 by Paul Broca, a French neurosurgeon, and it was the first area of the brain to be associated with a specific function. This discovery originated from patients who were able to understand spoken language and who did not suffer from any motor impairments of the tongue or mouth but were affected in their ability to speak. They could not produce complete sentences nor express their thoughts in writing. Autopsies of their brains revealed sizable lesions in the left inferior frontal cortex, a distinct region that Paul Broca identified as the “language center”. Broca and Wernicke regions form a highly connected loop, located in the majority of the cases in the left hemisphere.

Contrary to vocabulary reception, spontaneous word production is minimal before 12 months of age, even if babbling generally arises around 6 months of age [Figure 22]. The infant's first words are seen as “**proto-words**”⁸ because they are deprived of a referential link between a phonologically specified word form and the representation of a semantic meaning (Nazzi & Bertoncini, 2003). The first “**genuine words**”⁸ are usually produced around the child's first birthday [Figure 22] and encompass names for people and objects. Then, children exhibit a slow increase in word production, from fewer than 10 words at 12 months, to an average of 40 words at 16 months [Figure 21] (Fenson et al., 1994). Typically, children produce only one or two new words a week. Word production is then generally characterized by a **sharp acceleration around 18 months of age** (a phenomenon called “*vocabulary spurt*”, “vocabulary burst”, “*naming explosion*” or “*lexical boom*”), that typically occurs when the cumulative vocabulary production exceeds 50 words (Benedict, 1979; Goldfield & Reznick, 1990). Other studies showed that the vocabulary burst is not a universal phenomenon. Some “late talkers” display a slow monotonic rise in vocabulary that is best fit by a linear function, while other children are reported to show a series of small bursts, resulting in a step function (Goldfield & Reznick, 1990). Authors proposed that children who learn words at a more gradual pace may be using a strategy that attempts to encode broad ranges of experiences (i.e. lexicons are distributed to nouns, verbs, modifiers, pronouns rather than restricted to nouns as is the case for children with a typical vocabulary spurt profile).

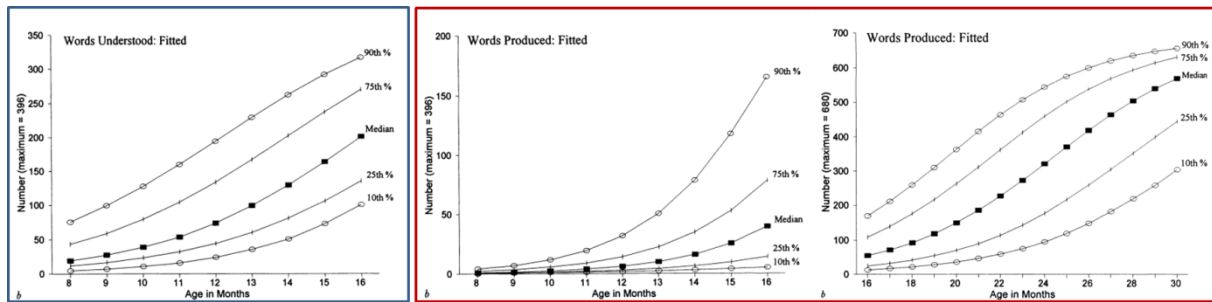


Figure 21 – Left (blue frame): fitted values of the number of words of the Infant form (MDCI) reported to be comprehended by children at each month; middle and right (red frame): fitted values of the number of words of the Infant form (MDCI) reported to be produced by children at each month. On each graph, median values are spread of score distributions. Adapted from Fenson et al., 1994.

Despite this, it has been suggested that **strong maturational constraints may mediate the onset of speech production** (e.g. Fenson et al., 1994). Indeed, the **acceleration in vocabulary production** may be related to the **increase in synaptic connectivity** and to **changes of the cortical activity pattern** that is known to occur in the same period of time (~17-20 months after birth) (Bates et al., 1992).

⁸**“Proto-words”**: pairing phonetically underspecified sound patterns to specific objects following repeated simultaneous presentations. Proto-words are characterized by pure associative links (Nazzi & Bertoncini, 2003).

“Genuine words”: pairing phonetically specified sound patterns to distinct object categories (at the child-basic level at first). Genuine words are characterized by a referential understanding and the representation of semantic categories (Nazzi & Bertoncini, 2003).

After the vocabulary spurt, word production exhibits a **regular pattern of growth across the 16-30 months of age** [Figure 21]. Between 18-20 months, children generally produce their first word combinations. At this time, these combinations are restricted to “telegraphic speech”, namely, few content words devoid of inflectional marking (e.g. “mommy sock”, meaning “mommy’s sock”) or grammatical function words (e.g. “give me cookie”, meaning “give me a cookie”) (Fenson et al., 1994). Around 30 months of age, a tenfold increase in productive vocabulary over the last 15-month period (i.e. 16-30 months) is observed, with 30-month-olds producing nearly 400 words. By **3 years of age**, most normal children have mastered the **basic morphological and syntactic structures** of their native language (Fenson et al., 1994). When reaching adulthood, humans’ lexical repertoire encompasses about 60,000 words. The extraordinary speed of language acquisition by young children was puzzling for traditional theories of learning.

1.3 Theoretical and computational models to explain early word learning

For many decades, researchers have tackled the question of how infants effortlessly solve the complex task of word learning, a task that even recent artificial intelligence approaches struggle to encrypt. Impressively, infants develop strategies that were unpredictable by the main historical theorists. It is now established that infants use **computational strategies to detect the statistical regularities and prosodic patterns of language inputs** (i.e. invariant sequences within domains, e.g. phonemes) (Kuhl, 2004). Precisely, infants first approach language with a set of initial **perceptual abilities**. Then, they are “primed” to learn the **regularities of linguistic inputs** directly from social exchanges, by combining **pattern detection** and **computational abilities** (also called *statistical learning*⁹) [Figure 22]. Nevertheless, infants’ perceptual and learning abilities are also constrained, since they cannot perceive all physical differences in speech sounds and are not capable of learning all possible stochastic patterns in language input (Kuhl, 2004). Despite all of this, they develop a remarkable sensitivity to acoustic patterns that are important for language.

⁹*Statistical learning: Acquisition of knowledge through the computation of information about the distributional frequency with which certain items occur in relation to others, or probabilistic information in sequences of stimuli, such as the odds (transitional probabilities) that one unit will follow another in a given language (Kuhl, 2004).*

The neural networks of early word learners become rapidly sensitive and selective to patterns that reflect native language (a phenomenon called **neural commitment**¹⁰) (Kuhl, 2004). In this way, exposure to a specific language sharpens infants’ perception of auditory stimuli near phonetic boundaries in that language. It proposes that the initial coding of native-language patterns may interfere with the learning of new patterns (e.g. a foreign language) because they do not conform to the neural pathways already established [Figure 22] (Bates et al., 1992; Kuhl, 2004).

¹⁰*Neural commitment: Learning results in a commitment of the brain’s neural networks to the patterns of variation that describe a particular language. This learning promotes further learning of patterns that conform to those initially learned, while interfering with the learning of patterns that do not conform to those initially learned (Kuhl, 2004).*

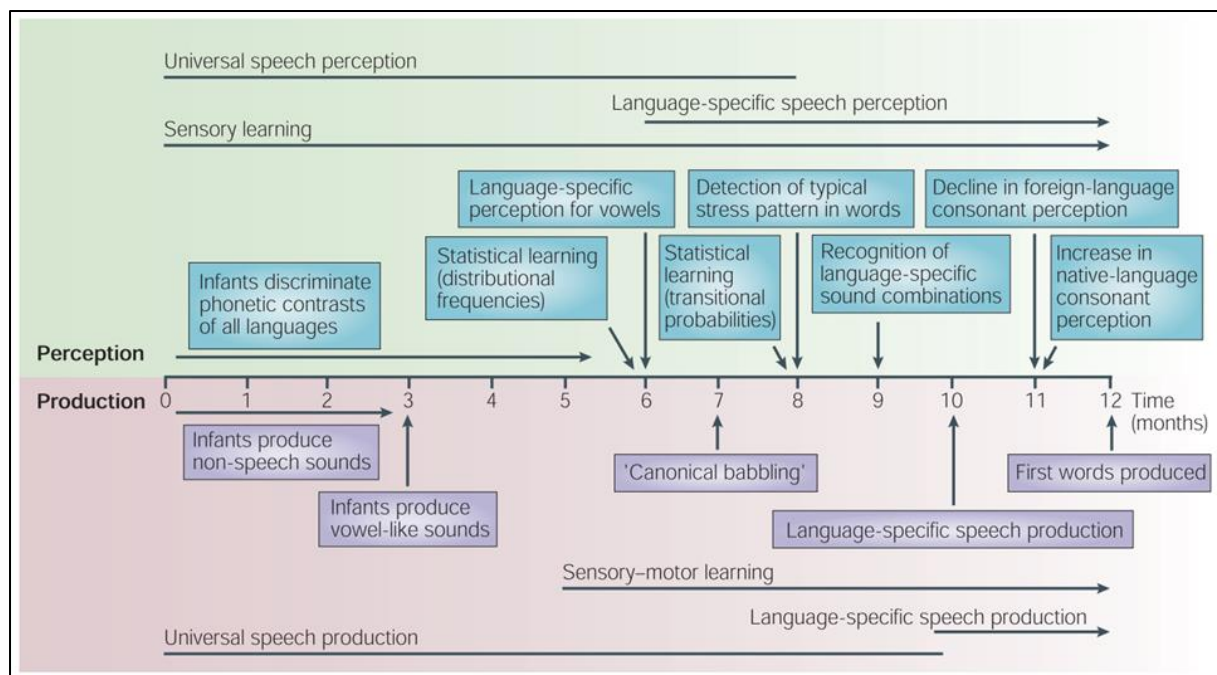


Figure 22 – The universal language timeline of speech-perception and speech-production development. This figure shows the changes that occur in speech perception and production in typically developing human infants during their first year of life. *From Kuhl, 2004.*

Specifically, by about 10 months, infants' ability to discriminate nonnative contrasts diminishes, as revealed by electrophysiological studies (Event Related Potentials (ERP) studies). It has been shown that American infants showed discrimination of both English (native) and Spanish (non-native) consonant contrast at 7 months in the **N250-550 response**¹¹, whereas 11-month-olds did so only for native contrasts (Rivera-Gaxiola, Silva-Pereyra, & Kuhl, 2005). Earlier exposure to speech does not preclude nor delay *neural commitment* (Pena, Werker, & Dehaene-lambertz, 2012). Indeed, it has been demonstrated that pre-terms do not appear to benefit from their earlier exposure to speech. Thus, the change in response to native and non-native speech depends rather on maturational age than on duration of experience (Pena et al., 2012). This finding suggests that the impact of speech exposure depends primarily on the state of the brain at the time of the experience.

¹¹*Specific ERPs markers of particular aspects of word learning (Friedrich & Friederici, 2011, 2015):*

P200-400: *positive component within 200 to 400ms evoked by visual stimuli, over posterior brain regions. In infants, this neural signature would reflect either experience-related ease to process visual stimuli or visual expectancy.*

N200-500: ***perceptual priming effect.** Lateral-frontal increased negativity (or reduced positivity) indicating perceptual familiarity to a spoken word (i.e. mental representation of a word form). In the context of cross-modal word-picture priming: both pictures and words can facilitate the perceptual processing of contextually expected words, which implies the presence of a neural connection between the object representation and its labeling word. Note that this effect is delayed in younger children whose neuronal network is less myelinated (~N300-800)*

N400: ***semantic priming effect (i.e. word comprehension).** Centro-parietal distributed negativity. Incongruous priming enhances negativity. In the context of cross-modal word-picture priming: the meaning of the word is pre-activated by the picture content, which implies the presence of referential connections between words and their meanings, and thus, the existence of at least simple genuine words. Note that this effect is delayed in younger children whose neuronal network is less myelinated (~N600-1200).*

In brief, early learning promotes future learning that builds on the patterns already encoded, but limits future learning of patterns that do not conform to those already encoded. A computational approach modeled early word-learning from a neuronal perspective.

The model uses the simplest form of associative learning, i.e. Hebbian learning. According to this **connectionist model**, sound patterns (word forms) and visual displays (objects) are initially randomly and weakly connected (Mcmurray, Horst, & Samuelson, 2012) [Figure 24, top]. In other words, connection weights start from small random values (e.g. when a child hears a novel word and is confronted to a multitude of unknown objects). Over time (i.e. across multiple learning experiences) the spurious connections will be pruned, while the genuine ones will be strengthened [Figure 24, bottom]. Precisely, inputs (words and objects) will be associated with an internal lexical unit if both are active in a temporal contiguity; otherwise the connection will decay. For example, repeated simultaneous exposition to a word (e.g. “dog”) and an object (a dog) will strengthen the connection between these two inputs while at the same time, the connection between the word “dog” and the object *tree* (which are not, or less, present simultaneously) will be weakened. Pruning prevents potential connections between unrelated

word forms and referents to be maintained at long term and simultaneously refines the connectivity between related words and referents. Over time, such changes build a system of links that encompass many words and objects. If a linkage becomes sufficiently reinforced, the pathway will be activated each time the word form (e.g. “dog”) is heard again, allowing the word to activate the appropriate concept without external support. However, even within this simple approach, the *connectionist model* assumes that there are layers of complexity. According to this computational model, auditory and visual inputs are indirectly connected between each other via lexical concepts (McMurray et al., 2012) [Figure 24, bottom].. These lexical concepts work like lemmas (i.e. abstract representations that connect other representations). Their presence means that learning requires at least two connections (word → lexicon; lexicon → object).

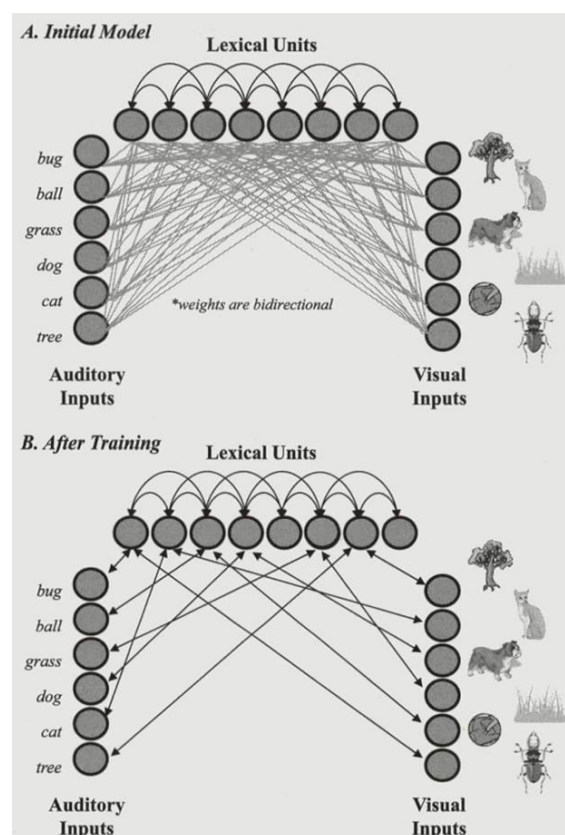


Figure 23 – Architecture of the connectionist model, before (A) and after (B) learning. *From McMurray et al., 2012*

In sum, the extraordinary and mainly effortless aptitude young humans have to acquire language is a fascinating topic. This becomes even more interesting when one considers the fact that their brain systems are engaged in dramatic changes and maturational processes. Researchers showed that language development is shaped by a sequence of age-based trends that are presumably constrained by maturational factors and refinement of the neural pathways. Substantial progress has been made in understanding the initial phases of language acquisition, at a behavioral, neural and computational level.

2 The word-referent concept

As indicated earlier, one of the first stages of language acquisition is the comprehension that phonetically specified sound patterns (words) can refer to item-unique things or to categories of things of the external environment (Quine, 1960). Such acquisition of genuine words constitutes a part of the semantic knowledge of infants. In this section, we will review studies related to the word-referent concept and we will specifically focus on the effect of two word learning strategies to promote long-term retention of novel word-object pairs. But first of all, we wanted to make precisions about the onset of the word-referent concept during childhood since recent studies have brought substantial insight into this field.

It has been shown that 10-14-month old infants attend more to objects when language (e.g. labeling) accompanies their inspection of those things (Baldwin & Markman, 1989). This tendency for **language to sustain infants' attention and examination of objects** helps to explain how infants take the **first step in establishing word-object relations**, i.e. the step of linking the language sounds they hear to the objects they see. Pioneer studies demonstrated that infants begin to treat words and objects as related by 11-12 months of age (Oviatt, 1980).

More recently, using a “*looking-while-listening*”¹² method (see Fernald, Zangl, Portillo, & Marchman, 2008 for the method), researchers established that around the middle of their first year, infants already comprehend the meaning of several words, specifically a range of food and body-part terms (Bergelson & Swingley, 2012). As evidenced, when hearing their parents labeling an object of a picture book, children tended to direct their gaze to the correct picture of the set of pictures. Does all of this mean that a 6-month-old child can already understand the genuine sense of a word?

¹²**Looking-while-listening (LWL):** methodology that uses real-time measures of the time course of infants' or young children's gaze patterns in response to speech for assessing word comprehension. Typically, infants are presented with visual displays, usually two discrete images, one of which is labeled in a spoken sentence such as “look at the [label]”. Eye movements are analyzed off-line, frame by frame, from the videotapes by trained coders who are naïve about the stimuli. Similarly to “preferential looking”, this method does not require automated eye-tracking technology. However, on the contrary to “preferential looking”, LWL yields high-resolution measures of speech processing from moment to moment, rather than relying on summary measures of looking preference. Specifically, children's gaze patterns are time-locked to the speech signal (Fernald et al., 2008).

The answer to that question arises from recent ERP studies. Indeed, the specific **N400 semantic priming effect**¹¹ was truthfully observed in infants as young as 6-month-olds who were exposed to pictures of novel objects repeatedly paired to pseudo-words during a learning session (Friedrich & Friederici, 2011). However, when testing this lexical knowledge the next day, this neural signature disappeared, indicating that the memory trace didn't survive 24h (it is only from 14months of age that this neural signature survives a delay of 24hours). Interestingly, the N400 was not reported in infants of age 3 months who underwent the same task (Friedrich & Friederici, 2015). Instead, the **N200-500**¹¹ brain correlate indicating familiarity to perceptual word forms was already found in these 3-month-old participants (Friedrich & Friederici, 2015). This finding supports evidence for a precocious establishment of mental representations of word sounds, of visual displays and of neural-based connections between them. However, it also indicates that at 3 months of age, a semantic processing stage is not yet reached. The authors are in favor of an associationist view, assuming that the ability of 3-month-olds to link words and visual displays together is caused by a primary learning mechanism. This mechanism would enable the creation of associative connections between the perceptual representations of objects and words but not of referential connections with a semantic component. This would constitute the neural base of the “proto-words” originally described by Nazzi and Bertoncini (Nazzi & Bertoncini, 2003). On the other hand, the base of “genuine” word comprehension arises around 6 months of age and would be characterized by tangible referential connections.

Once children expect words to be related to things in the world - in other words, once they acquired the word-referent concept - their receptive vocabulary increases gradually and linearly as described previously. However, we will see below that in this first stage of word acquisition, learning the name of a new object is a slow and **time-consuming process**. In the following section, we will focus on the different word learning strategies children rely on to build their receptive vocabulary repertoire and we will expose how these strategies may evolve across time. In this thesis, since we were particularly interested in two learning strategies (ostensive labeling and “fast-mapping”) the following review will mainly focus on these procedures. For each of these strategies, we will review how word comprehension translates into word retention and we will examine the putative underlying (brain) mechanisms.

2.1 Word learning strategies

When a young word learner is confronted with a scene that comprises multiple items, they may map a novel word they hear to any element of the scene. Thus, they face a large number of possible referents. How will they determine the referent (or more so the meaning) of the new

word? To illustrate this point, Quine imagined a stranger who hears a native say “gavagai” while pointing to a scene. To what does “gavagai” refer: the rabbit, the grass, a tree, the rabbit’s ears, or perhaps the beauty of the whole (Quine, 1960) ?

There are at least three main ways to solve this referential indeterminacy problem, which correspond to the main word learning strategies known to induce receptive lexical knowledge in young children.

More precisely, if the child encounters multiple situations where “gavagai” is heard and one element of the scene is the only constant element among the alternatives, then they may be able to detect this regularity (i.e. **co-occurrence frequency**) across these multiple individually ambiguous scenes. Ultimately, via **cross-situational statistics** they may be capable of identifying the right meaning for the word “gavagai” (Smith & Yu, 2008). If the word learner is now provided with a larger knowledge about several elements of the scene, they might then use an **elimination process** to solve the problem (also called *disjunctive syllogism* (Halberda, 2006)). Specifically, they should logically exclude the elements of the scene for which they already know the names and infer that novel word relates to the sole unnamed item. This strategy is well-known under the term of “**referent selection**”, which consists of the first step of the “**fast-mapping**” paradigm initiated by Carey and Bartlett (Carey & Bartlett, 1978). Lastly, the adult may offer overt social cues indicating **unambiguously** to which element of the scene the “gavagai” refers. For example, they may **ostensively** point at the corresponding item of the scene while saying “gavagai”. The onset and efficiency of these three strategies will be detailed below.

2.1.1 *Cross-statistical word learning (or cross-situational word learning)*

In their very first stages of receptive vocabulary learning, infants are similar to strangers who do not know the native language. In everyday contexts, children encounter many words, many potential referents but often limited cues indicating which word goes with which referent (Smith & Yu, 2008). In such **highly ambiguous learning contexts**, some young learners may just ignore the information and wait for contexts containing more explicit information about the referential bindings. In contrast, other young learners may take advantage of the information available to them, to use in the future when faced with subsequent ambiguous contexts (Smith & Yu, 2008).

Smith and Yu reported for the first time that infants are capable of learning word-object pairs simply by **tracking statistical regularities across multiple and individually ambiguous word-scene pairings**. They showed that the indeterminacy problem is not solved in a single experience but across multiple experiences, not for a single word-object pair but for a set of many pairings. In their experiment, 12- and 14-month-old infants were taught 6 word-referent pairs via a series of

30 individually ambiguous trials. On each trial, two word forms and two potential referents were presented with no information about which word went with which referent, as illustrated by the following example [Figure 24]. Note that in this paradigm, word-object pairs were consistent across trials, meaning that the referent assigned to a given word was constant. In total, each correct pairing occurred 10 times. Immediately after learning, infants underwent 12 test trials (2 test trials per word). Testing word learning typically consists of providing the child with 2 (or more) objects (or images of the objects) and measuring the percentage of time the child spends looking at each object after label onset (i.e. *preferential looking* after label onset). In older children, it is rather a measure of *preferential reaching* that is employed (e.g. “show me (or give me) the [label]”). In this study, the percentage of looking time toward each object was measured at each trial and revealed a significantly greater looking time at targets than at distractors.

The following example illustrates the cross-situational learning paradigm used by Smith & Yu [Figure 24]: the child hears the unknown words “bat” and “ball” while looking at a scene that involves a BAT and a BALL. In the absence of any other information, the young word learner cannot know to which of the two items the “ball” refers. However, if they subsequently view another scene containing a BALL and a DOG while hearing the words “ball” and “dog”, and if they registered the information from the previous trial, they should be able to combine the co-occurrence frequencies from the two streams of data, to correctly map “ball” to BALL.

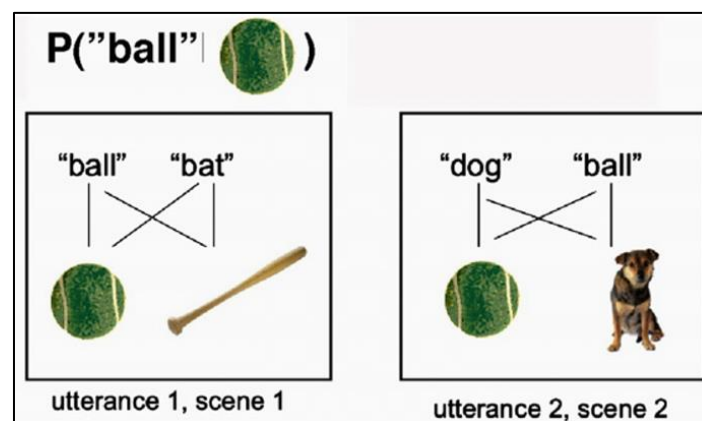


Figure 24 – Associations among words and referents across two individually ambiguous scenes. If a young learner calculates co-occurrence frequencies across these two trials, they can identify the proper mapping of “ball” to BALL. From Smith & Yu, 2008.

This word learning strategy is totally implicit and passive because it does not require the child to determine the referent; the child implicitly accumulates co-occurrence statistics to learn the mappings (McMurray et al., 2012).

According to the authors, this **cross-situational learning mechanism** is **rapid** and **efficient**, and may play a substantial role in early lexical learning (Smith & Yu, 2008). However, a recent work

showed that if a young word learner makes an **initial false hypothesis** about the possible referent of a given word, they will not perform better than chance when selecting among potential referents on a subsequent trial (Aravind et al., 2018). The authors argued that learners manifest a strong tendency to retain information about their initial hypothesis. So, if their initial hypothesis is erroneous, they will **struggle to recognize the missed alternative on a following trial**. In other words, they will struggle to recover. According to these authors, cross-situational learning is not a rapid word learning process and children require multiple instances of co-occurring frequencies to verify their initial associative hypothesis (Aravind et al., 2018).

2.1.2 *Ostensive labeling (“ostensive naming”, or “unambiguous naming”)*

In everyday life, children not only encounter ambiguous word learning situations, they also often face situations where the name of an object is explicitly given to them; a word learning process called “ostensive labeling”. Parents frequently provide ostensive definitions for their infants when they are about 9 months of age, because from this age, infants begin to look in the direction indicated by their parents (e.g. a parent pointing at the moon while saying “look this is the moon”) (Murphy & Messer, 1977). Ostensive labeling typically includes at least two components: **non-verbal external cues** (e.g. eye-gaze toward the referent, pointing or holding-up the object, etc.), and **labeling** (Baldwin & Markman, 1989). Sometimes, the object is presented in isolation from any potential distractor. In this case, labeling alone is sufficient to unambiguously induce the referential linkage. Hence, it goes without saying that ostensive labeling corresponds to an **unambiguous** word learning strategy. This procedure enables a great ease in figuring out the specific referent of a given label, which is presumed to considerably help increase children’s lexical repertoires (Baldwin & Markman, 1989).

Nevertheless, even if the word-object association is unambiguously given to the child, this latter also requires a lot of exposures to both the novel word and to the novel object to exhibit comprehension on test immediate trials (Oviatt, 1980).

For instance, 13-month-olds have been shown to be able to successfully associate the name of a novel object to the pairing after nine exposures but only if favorable circumstances were provided during learning (i.e. absence of preferential trials that tested for a possible innate preference for the target, presentations made in a bloc rather than in interspaced trials, etc.) (Woodward et al., 1994; see also Hollich et al., 2000 for similar results). Moreover, since children in these experiments had only one novel object name to learn, it is not clear whether the accuracy measured during testing truly demonstrates word learning or merely children’s ability to recall which object was given a special treatment (see Axelsson & Horst, 2013). When controlling for

this potential confounding factor by teaching children two novel word-object pairs, it has been demonstrated that 13- to 17-month-olds but not 8- to 12-month-olds, recognized the pairings, provided that they were taught at least 10 to 12 times each (Gurteen, Horne, & Erjavec, 2011; Werker et al., 1998). The same results were found when stimuli consisted of abstract images displayed on a screen (rather than 3D objects) (Schafer & Plunkett, 1998). Overall, these studies indicate that **before 18-months of age**, learning the name of a new object via ostensive labeling is a slow and **repetition-based process**. Indeed, children seem to require at least ten or more unambiguous repetitions of the linkages in order to manifest word comprehension on immediate test trials.

Intriguingly, this repetition-driven pattern seems to be no longer vital anymore after 18 months of life. Indeed, 18-month-olds showed comprehension on immediate looking-while-listening test trials after being presented with only three* (Houston-Price, Plunkett, & Harris, 2005) or four (Bion, Borovsky, & Fernald, 2013) ostensive teaching trials that encompassed two word-object pairs. **(in this study, only the labels were novel; objects were already familiar to the participants).*

This finding is particularly interesting given that around 18 months of age children undergo the developmental phenomenon of **vocabulary spurt**, that is, the onset of the extraordinary productive vocabulary burst described previously. One may speculate that both patterns are correlated: the change in lexical production (vocabulary burst) would accompany changes in lexical comprehension. Moreover, vocabulary spurt and rapid word comprehension also coincide in time with **changes at the neural level** and with a **reorganization of the neural substrate** of language processing brain regions. Indeed, this time period is characterized by a steep increase in synaptic density in many relevant cortical regions (Bates et al., 1992). According to the authors, this would enable a larger capacity for information processing and storing. Consequently, the quantitative shifts in word learning observable at a behavioral level (vocabulary spurt and improvements in vocabulary reception) may be caused by a qualitative shift of the mechanisms responsible for word learning at the neural level.

Furthermore, it has been suggested that the neural system mediating early language comprehension (and perhaps production) may be distinct from later emerging lateralized systems whose engagement may only be necessary for more advanced linguistic skills (Dapretto & Bjork, 2000; Mervis & Bertrand, 1994). In other words, the hemispheric lateralization of the neural substrates responsible for language functions coincides with the striking improvements in word comprehension and production. Precisely, a study using ERP investigated the neural correlates of processing known and unknown words in children before and after the vocabulary spurt

phenomenon (Mills, Coffey-Corina, & Neville, 1993). Results showed that the amplitude of ERPs from 200 to 400ms of **pre-vocabulary spurt children** were significantly larger for known than unknown words. These differences in amplitude were **broadly distributed over anterior and posterior regions of both left and right hemispheres**. In contrast, in children who **underwent vocabulary spurt** (i.e. for 20-month-olds who had vocabularies of over 150 words), ERP differences from 200 to 400ms were more **focally distributed over temporal and parietal regions of the left hemisphere**. A subsequent analysis comparing children with different vocabulary sizes at an equal age demonstrated that this shift in brain response was **related to vocabulary, not to age** (Mills et al., 1993).

To summarize, despite a large diversity of the methodologies used (e.g. looking-while-listening, preferential reaching, forced-choice tasks, etc.) and of the parameters manipulated across studies (e.g. age, number of word-object pairs to learn, number of exposures to each pair, etc.), the literature clearly suggests that there is a noteworthy developmental step in the **child's ability to master word learning processes that encompass ostensive labeling during the second year of life** (Gurteen et al., 2011; Hollich et al., 2000; McMurray, 2007; Reznick, 1990; Werker, Cohen, Lloyd, Casasola, & Stager, 1998; Woodward et al., 1994). This improvement in vocabulary comprehension coincides with the phenomenon of vocabulary spurt, and both may depend on maturational advances such as reorganization of the cortical structures and lateralization of the functions devoted to language. Researchers suggested that the regions responsible for first-language learning in young children are not necessarily the ones responsible for language use and maintenance in adulthood (see also Johnson & De Haan, 2015). The interactive specialization view proposes that the brain activity underlying language functions becomes more focal with experience, as language skills become more efficient and automated. Alternatively, it can be suggested that children may need to have undergone advanced maturational processes to rapidly and more efficiently incorporate new words in their lexicon.

Also intriguingly, around this same age period, another language-related skill emerges: the ability to infer a novel word to its referent by logically excluding familiar items, a phenomenon known as “**referent selection**” or “**disambiguation ability**” (Halberda, 2003).

2.1.3 *Referent selection (or “disambiguation ability”)*

In 1978, Carey and Bartlett's pioneering study revealed that 3-4 year old preschoolers could infer the referent (an unfamiliar color) from an unknown word (a new color name) when hearing their teacher ask them: “go and get the *chromium* tray, not the blue one, the *chromium* one” (Carey & Bartlett, 1978). More surprising, subsequently to this one-trial lexical experience, children

remembered this semantic information (the chromium color) more than 1 week later, when they had to recall which one of a set of six unfamiliar color chips was the chromium color (Carey & Bartlett, 1978). The authors named this remarkable skill “fast-mapping”, given the extreme rapidity with which children linked the novel word to its referent and retained this mapping over time. This discovery inspired a proliferation of research that has tried, for many decades, to discover the underlying mechanisms of this learning procedure, in children but also in adults and more recently in unrelated animals (e.g. the domestic dog).

In Carey and Bartlett’s pioneer study, children were provided with an **explicit lexical contrast** (“the chromium tray, *not the blue one*”). In contrast, the current majority of studies investigating “fast-mapping” use only a **familiarity contrast**. Specifically, children are presented with a novel object and one (or more) familiar object(s) when they hear the novel word [Figure 25]. They tend to logically exclude the already named object(s), and by deduction select the unfamiliar item as the referent. This procedure is called “referent selection” (or “disambiguation ability”) and constitutes the first step of the whole “fast-mapping” procedure that requires children to retain the information at long term (i.e. “referent retention”, a point that will be detailed later). Thus, disambiguating a situation is only the first step in apprehending the meaning of a new word.

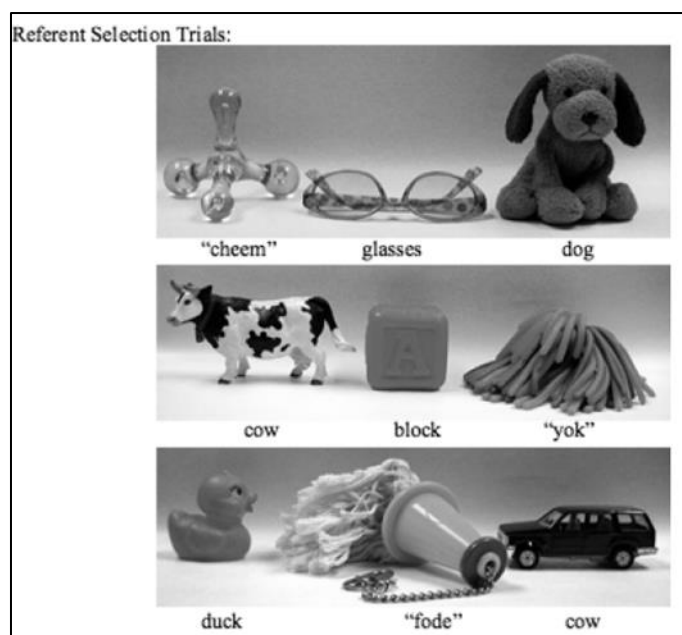


Figure 25 – Example of stimuli used by Horst & Samuelson on the “referent selection” trials. *From Horst and Samuelson, 2008.*

One overwhelming question is when exactly does this capability emerge during development? Researchers all agree that disambiguation biases are not present at the very onset of lexical development, precisely **not before the vocabulary spurt phenomenon**, as initially demonstrated by Mervis and Bertrand (Mervis & Bertrand, 1994). However, there are conflicting findings in the literature about the exact onset of the disambiguation ability. Halberda was the first to use online

looking-time measures to study referent selection by 14- to 18-months old and concluded that children were successful in disambiguation by the age of 17 months (Halberda, 2003), a result supported by other studies using comparable measures (e.g. Houston-Price, Caloghris, & Raviglione, 2010; White & Morgan, 2008). On the contrary, other studies failed to demonstrate an increase in looking at the novel object after hearing a novel label in children participants of 16- to 23-months of age (Bion et al., 2013; Mather & Plunkett, 2009, 2011). In spite of this, in Mather & Plunkett's research, 16-month-olds could map a novel label onto a novel object, provided that the novel label had no familiar phonological neighbors, suggesting that linguistic principles may govern "referent selection" abilities (Mather & Plunkett, 2011). Indeed, it is argued that the ability of mapping novel labels onto novel objects is constrained by (perhaps innate) linguistic heuristics. The main principles governing this **constraint approach** described in the literature are the following: **Mutual Exclusivity**¹³ (Markman & Wachtel, 1988), **Contrast**¹⁴ (Clark, 1988), **Whole Object Assumption**¹⁵ (Markman & Wachtel, 1988), **Pragmatic Account**¹⁶ (Diesendruck & Markson, 2001) and the **Novel-Name Nameless-Category Principle (N3C)**¹⁷ (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Mervis & Bertrand, 1994). Some of these principles are not available in the first stages of language acquisition, suggesting that a minimum experience with language is required to solve the "referent selection" by exclusion trials. This may help to explain why the onset of this ability is shown to be closely correlated to vocabulary spurt, thus also to a substantial size of the receptive lexical repertoire. As already described above, the increased lexical repertoire and the vocabulary spurt phenomenon are themselves supposed to be mediated by changes and reorganization within the underlying brain structures. Note that around two years of age, children who are mostly post-vocabulary spurt word-learners all generally exhibit excellent disambiguation abilities (e.g. Bion et al., 2013; Horst & Samuelson, 2008).

Note also that the **constraint approach** has limits and is still largely debated. Despite this, at least it raises fundamental questions that go beyond the frames of this thesis. For example, once children master basic-level terms, how is the taxonomic constraint relaxed to learn superordinates? When and how do children violate the whole-object assumption to learn object-related properties? How do children ignore mutual exclusivity to learn synonyms or superordinates? (see McMurray et al., 2012 for a review).

¹³**Mutual exclusivity (ME):** assumption that every object has just one name. In other words, words do not have overlapping references (Markman & Wachtel, 1988). Upon hearing a novel label, ME motivates a word-learner to reject objects that already have a known label.

¹⁴**Contrast:** principle that all lexical entries contrast in meaning (Clark, 1988). A word-learner using Contrast would avoid taking the novel word to be synonymous with the lexical entry for the known object.

¹⁵**Whole Object Assumption:** assumption that the novel word refers to the whole entity rather than to a part or attribute of it (Markman & Wachtel, 1988).

¹⁶**Pragmatic Account:** hypothesis that the speaker should use familiar terms when available. When presented with a familiar (e.g. a brush) and a novel object and asked for the novel item (e.g. “show me the dax”), a word-learner utilizing Pragmatics would reason as follow: “if the experimenter had wanted me to pick up the [brush], they would have asked me to show them the [brush]. Given that they asked me for a dax, they must have wanted me to give them the other object” (Diesendruck & Markson, 2001).

¹⁷**Novel-Name Nameless-Category Principle:** the N3C is the principle that word-learners are positively motivated to map novel labels to objects that do not yet have a name (Golinkoff et al., 1992). It postulates the strategy of “Map-Novelty-to-Novelty” (Mervis & Bertrand, 1994). This principle predicts that children tend to select the unnamed object “for the positive reason that children seek names for objects that are previously unnamed” (Golinkoff et al., 1992).

To summarize, in this section, we stressed the **gradual nature of word learning** and provided evidence from the overall literature that the **strategies that shape word learning have different developmental onsets**. The comprehension that everything has a name, the impressive gains in understanding spoken words as well as the rise in disambiguation abilities may coincide with vocabulary growth and maturational processes during the second year of life. In spite of this, for word learning to be truly efficient, children must be capable of retaining and recalling the words acquired through these different learning strategies after substantial delays.

2.2 Long-term storage of word-referent pairs in young children

For word learning to be successful, children not only need to understand the meaning of a new word or to disambiguate a language-related situation, they also need to **retain this knowledge for later use** (Wojcik, 2013 for a review). In this thesis, again, since we were mostly interested in *ostensive naming* and *referent selection*, the following section will only focus on the effect of these two learning strategies in retaining the names of novel objects (i.e. we will not discuss the *cross-situational* learning strategy).

Unfortunately, only a handful of studies have investigated how word learning translates into more permanent memory traces. In the literature, testing mostly occurs immediately after learning. Therefore it cannot be considered as measuring word retention, but rather *word comprehension*. **Word retention** was demonstrated in **3 and 4 year-old children** for up to **a week** (Carey & Bartlett, 1978; Markson & Bloom, 1997; Vlach & Sandhofer, 2012; Waxman & Booth, 2000) or **a month** (Carey & Bartlett, 1978; Markson & Bloom, 1997; Vlach & Sandhofer, 2012). However, in these studies, children were only taught one novel word-object association. Previous studies have demonstrated that labeling an object during training increases its salience and thus enhances its chances of being chosen by the child in subsequent tests (Baldwin & Markman, 1989). As already argued above, the accuracy reported during these long-term memory tests may simply reflect recalling which object was given a special treatment during the encoding phase rather than demonstrating actual word retention (Axelsson & Horst, 2013).

2.2.1 *Ostensive labeling: long-term retention of ostensively labeled objects*

In the context of ostensive labeling, there is no clear evidence from the overall literature that children who were taught at least two novel associations simultaneously demonstrated retention after a delay (Wojcik, 2013 for a review). To the best of our knowledge, only one recent study has addressed this question. In this study, the author demonstrated that 30- to 34-month-olds had significant retention and generalization abilities when tested **1 min** plus **1 week** after learning (Wojcik, 2017). Precisely, participants were trained on four novel objects, each of which was displayed in isolation four times on a screen (it consisted of 4 blocks of 4 exposures) and word labels were spoken in carrier phrases (one labeling per visual exposure). About one minute after training, children underwent a first *encoding test*. It consisted of 4 blocks of 4 test trials (i.e. each pair was tested 4 times) and looking behaviors were measured to assess comprehension. Participants then performed *retention and generalization tests* one minute after the encoding test (for a subgroup of subjects) or one week after the encoding test (for a second subgroup of subjects). Performance was significantly above chance after both delays and generalization

accuracy was overall significantly lower than retention accuracy. Here, the author showed for the first time that few ostensive labeling trials were sufficient to promote a long-term storage of the information in 30-month-olds (Wojcik, 2017). Nonetheless, it might be that the encoding test performed immediately after learning induced a rehearsal of the information which consequently reinforced the mappings, independently of the labeling trials. It is in fact well established that retrieval greatly strengthens memory representations (Roediger & Butler, 2011 for a review).

Overall, there is a serious lack of research into the question of whether and how ostensive labeling promotes the long-term retention of new word-object associations in childhood. The studies performed in this thesis aim to fill this gap.

2.2.2 “Fast-mapping”: *referent retention*

In the context of “fast-mapping”, were the objects correctly mapped by exclusion by the subjects also later correctly recognized (i.e. “*referent retention*”? From the overall literature, 24-month-olds systematically failed to demonstrate retention of the fast-mapped words after a five-minute delay (e.g. Horst & Samuelson, 2008) or even immediately after learning (e.g. Bion, Borovsky, & Fernald, 2013). 24-month-olds could retain the mappings for 5 minutes, provided that (i) the objects were inferred by logical exclusion and deliberately labeled several times (i.e. ostensively named) by an experimenter holding them up (Experiment 2, Horst & Samuelson, 2008); or that (ii) ostensive naming was provided in addition to the logical selection and the children’s attention drawn toward the referent by illuminating the target and/or covering the familiar competitors [Figure 26] (Axelsson et al., 2012); or finally that (iii) the objects were already familiar to the participants prior to the learning phase (Kucker & Samuelson, 2012, see also Houston-Price, Plunkett & Harris, 2005 for a discussion).

In one study, 24- and 30-month-olds demonstrated successful retention after a 24h delay, but the pairings were reviewed with ostensive labeling after the “referent selection” trials. So, it is unclear whether the inferential process itself sufficed to promote retention or whether the success was mainly attributable to the revised trials (Goodman, McDonough, & Brown, 1998). Overall, there is a consensus that **around two years of age, a single disambiguation situation is not enough to induce retention** of the name of an object, even after a minimal delay. There is no clear evidence about the number of times 2-year-olds should disambiguate a situation encompassing the same novel object to retain the name of that object.

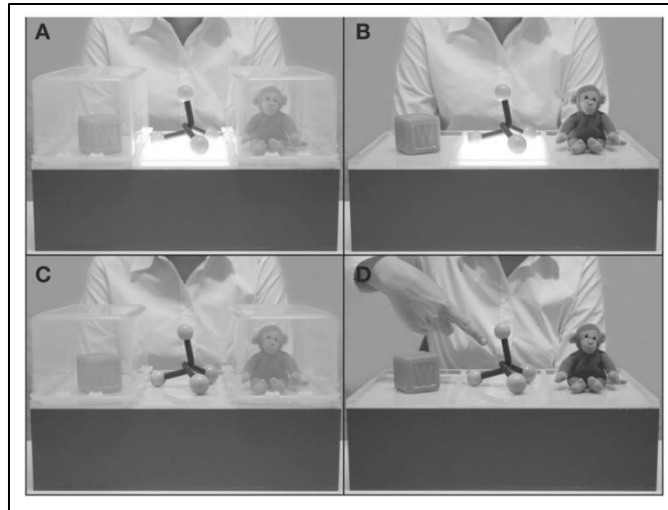


Figure 26 –Examples of feedback that followed the referent selections: (A) highlight-target/dampen-competitors, (B) highlight-target, (C) dampen-competitors, (D) control (pointing). From Axelsson et al., 2012

Why are fast-mapped objects not retained by 2-year-olds?

One explanation is that during the “referent selection” trials, the attention of child participants has to be shared between the referent and the familiar objects (i.e. the competitors), which may not allow a proper encoding of the physical properties of the referent itself (Horst & Samuelson, 2008). Worse, children may determine the referent by paying exclusively attention at the familiar objects, though this hypothesis has not yet been investigated. However, what has been shown is that the number of competitors didn’t affect the ability of 30 months old children to form the initial mappings; however only those who encountered fewer competitors during “referent selection” (a maximum of 2 competitors) exhibited retention after a 5-min delay (Horst, Scott, & Pollard, 2010). Moreover, it has been demonstrated that the saliency of the target needs to be increased to attract participants’ attention toward it during learning. Holding up, playing with, or pointing to the target as well as labeling it, illuminating it, covering the competitors or reducing the number of competitors, etc., likely help sustain children’s attention on the target object. In turn, it may facilitate the encoding process of the visual features of the target, of the auditory properties of the label and of the relational link between the two, so that the information could be recalled after a few-minutes delay.

An alternative explanation is that referent selection is not isomorphic to learning. It may only constitute an online process that is independent of long-term word learning. Perhaps all that counts for children is to arrive rapidly at the right inference for their immediate communicative requirement without the need of learning and remembering the linkages (McMurray et al., 2012). According to their *dynamic associative word learning model*, the authors compared the referential ambiguity to a situation-time problem that must be solved within the context of a single inferential

event, but building long lasting linkages would only be solved over developmental time as children encounter multiple presentations of the elements to be mapped (McMurray et al 2012).

Overall, it appears that the term “*fast-mapping*” initially adopted by Carey and Bartlett may truly help young children understanding a live conversation by disambiguating word-related situations but is definitely not synonymous with “*fast-learning*”. At least, this definition does not apply for children around 2 years of age. Improvement seems to occur at the end of the second year of life. Indeed, Bion et al. found that 30-month-olds, but not 24-month-olds, showed a fragile evidence of retention during “referent retention” tests carried out just after the disambiguation trials (Bion et al., 2013). The authors employed the *looking-while-listening paradigm* to assess retention and found that 30-month-olds looked nearly above chance to the targets on those test trials. However, in this study, children underwent 4 “referent selection” trials for each of the two pairs, which may explain the discrepancy with previous findings (where children underwent a unique “referent selection” trial per pair). Moreover, participants had only two pairs to retain compared to four in Horst and Samuelson’s experiments. In another study, Zosh et al. demonstrated that older children (36-42-month-olds) successfully recognized the names of the objects learned via inferential reasoning (Zosh, Brinster, & Halberda, 2013). However, since participants were tested immediately after learning, this finding is not conclusive about an actual storage of the information at long-term. It only provides evidence about the emergence of retention abilities for fast-mapped objects. Interestingly, in this study, the authors also showed that the retention rate was higher for the objects learned via inferential reasoning than for objects learned via instruction (i.e. ostensive labeling) (Zosh et al., 2013). This finding suggests the possibility of a developmental shift in the strategies that support word learning (Zosh et al., 2013).

2.3 Which neural bases support the formation of a memory for ostensively-named and fast-mapped words?

Earlier, we reviewed contradictory points of view about the involvement of the hippocampus in the formation of semantic memories. On one hand, some researchers and clinicians claim that semantic memory is stored by the same brain systems involved in episodic memory (i.e. the hippocampal system and the MTL)(see Squire, 2004 for a review). However, the dominant viewpoint is that the consolidation process of semantic knowledge is only supported by the brain structures surrounding the hippocampal system (the entorhinal and perirhinal cortices) (Brown & Aggleton, 2001; Schmolck et al., 2002). On the other hand, other researchers and theories assume that rapid acquisition of new words may not even necessitate the involvement of the MTL

structures at all (e.g. McClelland, 2013; Sharon et al., 2011). According to this alternative viewpoint, new knowledge would be directly stored into the neocortex.

To further emphasize the debate, one may wonder if ostensibly named objects and fast-mapped objects depend on the same brain structures? To date, there is very little evidence of this and the rare studies that addressed that question are mainly based on works involving amnesic patients.

In one study, four adult amnesic patients with **hippocampal system damage** showed severe impairments on “*explicit associative tasks*” (~ostensive labeling) but demonstrated restored performance to the level of age-matched healthy controls when a “fast-mapping” (~referent selection) procedure was used (Sharon et al., 2011). When tested after a delay of one week, the amnesic patients recognized the associations acquired via “referential selections” but were around chance level on the standard “explicit task”. Conversely, control subjects performed better on the “explicit associative memory task” (Sharon et al., 2011). However, note that the learning procedures employed by the authors were slightly different from those typically used with children. Indeed, “explicit associative tasks” consisted of displaying and labeling objects in isolation but participants (patients and healthy controls) were deliberately instructed to retain the associations [Figure 27]. For their part, “referent selection” trials required participants to answer questions that allowed them to infer which object corresponded to a particular name (e.g. adults: “is the [label]’s tail pointed up?”; children: “where is the [label]?”) [Figure 27]. This is why authors argue that their word learning procedures encompass explicit learning in the first case and implicit learning in the second.

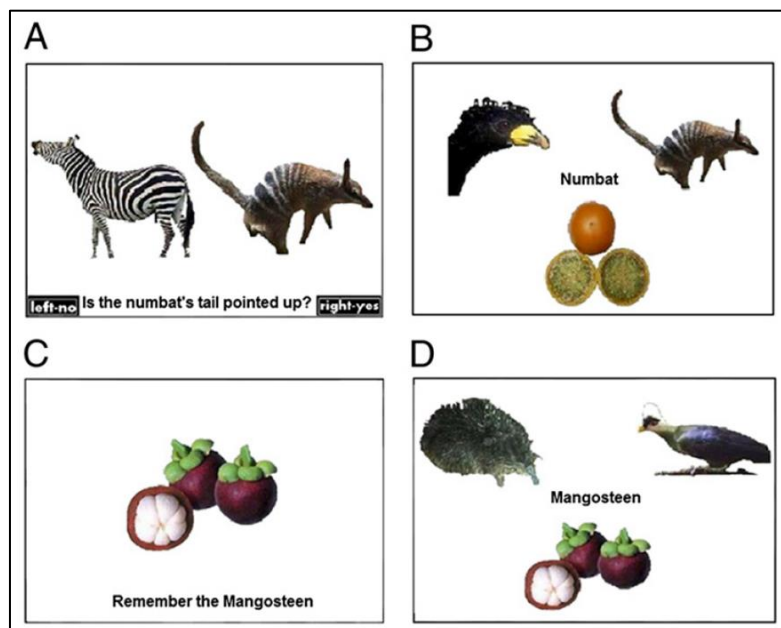


Figure 27 – Examples of stimuli used in the fast-mapping (FM) and explicit experiments. (A): previously unknown target study phase trial in the FM experiment. (B): Recognition test trial in the FM experiments. (C): Previously unknown target study phase trial in the explicit experiment. (D): Recognition test trial in the explicit experiment. *From Sharon et al, 2011.*

The authors then included two additional patients with unilateral damage to the anterior temporal neocortex. These patients showed impairments on “referent retention” test trials, suggesting that associative learning through inferential reasoning would rather depend on extrahippocampal neocortical regions, such as the lateral and **anterior temporal lobes**.

In a following fMRI study using **multivariate pattern analysis (MVPA)**¹⁸ and involving healthy adults, the same first author found that “fast-mapping” induced memory performance through greater activity within the **anterior temporal lobe**, the **lateral occipito-temporal**, the **parieto-temporal neocortex**, and **ventro-lateral prefrontal cortex** (Atir-sharon, Gilboa, Hazan, Koilis, & Manevitz, 2015) [Figure 28]. By contrast, memory performance following explicit encoding elicited activity in **medial and dorsolateral prefrontal** and **parahippocampal cortices** [Figure 28]. The authors claimed that the performance obtained in their experiments could not be due to non-declarative mechanisms, such as priming, because testing typically consisted of an explicit recognition test (3-alternatives forced-choice trials, i.e. 3AFC), whereas priming is rather revealed by indirect implicit measures of improved performance across time.

¹⁸**MVPA=Multivariate pattern analysis:** method to analyze neural responses as patterns of activity reflecting the varying brain states. In contrast to simpler univariate measures, the MVPA method increases the amount of information that can be decoded from brain activity. Specifically, it encompasses a pattern recognition algorithm (a machine-learning classifier) that “learns” a functional relationship between brain response patterns and specific sensory inputs (i.e. stimuli) (values can be either discrete (classification) or continuous (regression)). This learned functional relationship is then used to predict unseen stimuli from a new dataset (“brain reading”).

To summarize, these authors showed that four hippocampal-damaged patients were no longer “amnesic” when learning consisted of a “fast-mapping” procedure, suggesting that “fast-mapping” might bypass the hippocampal system, and in contrast to explicit learning, allows a rapid incorporation of new knowledge into existing cortical memory networks (see also Coutanche & Thompson-Schill, 2014). In other words, this claim of a non-hippocampal learning supports the idea that rapidly acquired information can be supported by structures outside the MTL, especially in the anterior temporal lobe, when an incidental learning procedure is used (Sharon et al., 2011).

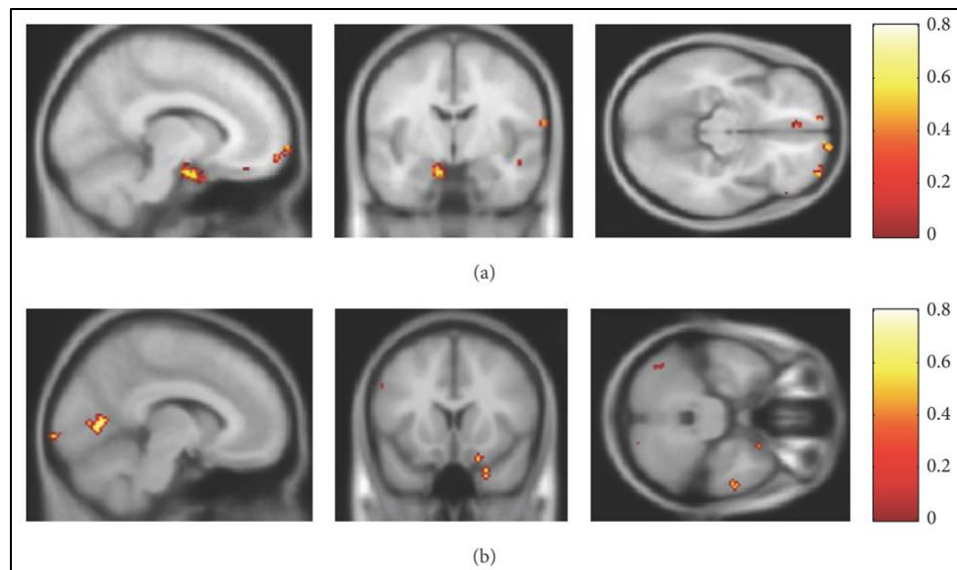


Figure 28 – Searchlight results (two-way classification between successful vs failed recognition conditions) for the explicit encoding condition (a) and “fast-mapping” (FM) condition (b) across participants. Explicit encoding (a) is associated with regions in the MTL including the hippocampus; bilateral ventral medial prefrontal cortices (VMPFC); right lateral prefrontal cortex; anterior cingulate and right posterior lateral temporal neocortex. FM (b) is associated with bilateral anterior temporal lobe; posterior lateral and inferior temporal neocortical regions; posterior inferior occipital cortices and frontal lobe involvement (orbitofrontal, dorsolateral, ventrolateral PFC but no VMPFC). *From Atir-sharon, Gilboa, Hazan, Koilis, & Manevitz, 2015.*

On the other hand, other studies severely contradicted this evidence. First, another study involving seven memory-impaired patients with hippocampal (N=6) or larger MTL (N=1) lesions failed to replicate the results obtained by Sharon et al., (they used the exact same paradigm and stimuli) (Smith, Urgolites, Hopkins, & Squire, 2014). In this study, patients were markedly impaired relative to age-matched healthy controls in both “fast-mapping” and explicit learning conditions (i.e. they did not benefit from a “fast-mapping” learning procedure). In another study, during which the “referent selection” trials mimicked protocols used with children (e.g. “click on the numbat”), hippocampal-damaged patients performed as well as healthy controls in inferring the their targets from the unknown words, but in contrast to the control subjects, these patients no longer retained the associations after a delay (Warren & Duff, 2014). The authors claimed that the **hippocampus is not essential for on-line “fast-mapping”** of novel words (i.e. referent selections), but is **necessary for the maintenance of arbitrary relational information**.

Second, it has been shown that **aged participants** with **reduced hippocampal grey-matter volume** (revealed by fMRI scans) manifested poorer retention for associations acquired through “fast-mapping” compared to associations acquired through explicit instructions (paradigm and stimuli identical to the one used by Sharon et al.) (Greve, Cooper, & Henson, 2014). The authors found that hippocampal volumes predicted memory performance in both the “fast-mapping” and explicit conditions, suggesting that both **learning procedures were supported by the same MTL structures** assumed to enable rapid associative learning. On the contrary to Sharon’s

evidence that the anterior temporal lobe would support fast-mapping, here the authors found no evidence that the volumes of this brain structure predicted memory performance in either “fast-mapping” or explicit learning (Greve et al., 2014).

Finally, patients with Down syndrome characterized by **hippocampal dysfunctions** and memory and learning impairments did not benefit from a “fast-mapping” learning method compared to healthy controls nor compared to an explicit encoding method (Sakhon, Edwards, Luongo, Murphy, & Edgin, 2018).

The discrepancy in the findings might arise from the fact that the populations studied are simply “functionally” not equivalent. For the same reason, it would be too prejudicial to infer early developing brain systems from the properties observed in either amnesic patients, patients with Down syndrome or old people. Moreover, since the “fast-mapping” and explicit encoding paradigms developed for adults differ from those employed with children, direct comparisons are unlikely.

Conclusion

In the first stages of language acquisition, word learning may be mediated by procedural memory or basic associational mechanisms and is deprived of any semantic content. From 6 months of age, learning a new word is associated with a referential (i.e. a meaning) component but the semantic priming effect disappears after a 24h delay, indicating that at that age, although word learning becomes semantic the knowledge no longer survives. Around 14-months of age, the referential connections between arbitrary sound patterns and visual displays seem to be maintained after a certain delay. Nonetheless, at a behavioral level, 14-month-old word learners require a bunch of presentations of the pairings to recall the information after a minimum delay.

A striking developmental switch is observed between 18 and 20 months of age, which is reflected by (i) the vocabulary spurt phenomenon, (ii) a rapid incorporation of novel words into the receptive lexicon and (iii) the emergence of inferential reasoning abilities in the context of word learning. Maturational processes accompany those cognitive changes (e.g. increased connectivity within the hippocampus, reorganization of the cortical structures, onset of the lateralization process, etc.) but there is no consensus about which the cause is and which the consequence effect is.

Moreover, word retention is still fragile at that age and has essentially been investigated immediately after learning. Thus, while around their second birthday children demonstrate the

ability to rapidly comprehend the meaning of a new word and to easily map labels onto referents, the memory processes underlying word learning and their development are still poorly studied and poorly understood. As a matter of fact, there are still conflicting discussions in the literature about the brain areas that underlie word learning and word retention. On one hand, theories and researches conducted on amnesic patients claim that the hippocampus only supports the memorization process of information acquired explicitly while words learned incidentally would be directly stored within neocortical structures. If so, why would 2-year-olds not be capable of recognizing fast-mapped words after a minimal delay, and conversely seemingly capable of retaining ostensibly named words that would depend on less mature brain systems? On the other hand, it has been proposed that the protracted maturation of the hippocampus is mainly responsible for the delayed ability of children to maintain semantic (as well as episodic) knowledge at long-term, whatever the learning method. For semantic information to survive longer delays, children would rely on multiple repetitions which would strengthen the neural pathways. If so, what would be the minimum number of exposures to a novel word-object pair and the learning conditions required for the association to survive after a delay?

The findings of previous research also indicate how crucial adequate learning conditions are for memories to be formed. Indeed, they suggest the aptness of ostensibly naming the objects as well as using external cues such as pointing or holding up the objects, to foster pairing recall in 24-month-olds after a minimum delay (see also Booth, McGregor, & Rohlfing, 2008). Nonetheless, how this fertile ground evolves during the third and fourth year of life is still another open question. Related to this issue, little is known about a potential gradual shift in efficiency of the word learning strategies presented above throughout development. Is “fast-mapping” still a “slow-learning” process later in development? The high variability in the methods used to measure word retention makes it difficult to compare the findings obtained by the different researches and thus, to draw the developmental history of long-term memory for word-object pairs following various learning strategies. This constitutes one of the challenges of this thesis.

By using a uniform and highly controlled methodology applied to child populations ranging from 18 months to 4 years of age, we will be allowed to directly compare the efficiency of two learning strategies (ostensive labeling and inferential reasoning (i.e. “fast-mapping”)) on memory formation at different stages of the development. This will also enable us to investigate the minimum number of exposures required at a given age to form a memory trace, which should indirectly, reveal much about the putative underlying brain mechanisms.

CHAPTER IV.

EXPERIMENTAL PROJECTS

1 Presentation of the project

In this thesis, we aimed to better understand the learning conditions and principles that govern the formation of a memory trace for cross-modal stimuli, such as word-object pairs, during early childhood. We also aimed to comprehend whether and how memory formation for semantic knowledge evolves during development. Finally, we liked to know how long young children could maintain “dormant” in memory, semantic information to which they were exposed again.

To address these questions, three children populations reflecting strategic developmental stages underwent different learning procedures and their knowledge for the newly acquired word-object associations was tested after a distractive period of 30-min, and re-tested after delays of either 1 month or 6 months (only for subgroups of subjects). Precisely, the populations consisted of **18-month-olds** (i.e. mainly pre-vocabulary spurt children with immature hippocampal systems), **24-month-olds** (i.e. mainly post-vocabulary spurt children with immature hippocampal systems) and **4-year-olds** (i.e. children who typically speak well, have a more mature hippocampal system and reach the offset of the infantile amnesia phenomenon). The accuracy of each age group in recognizing the pairings was compared to the other age groups, to a group of adult participants who underwent the same teaching methods under the same conditions, and to levels expected by chance.

The effect of different word learning procedures on retention was tested in this thesis. We especially focused on ***ostensive labeling*** and ***inferential reasoning*** (i.e. “fast-mapping”). For these two word learning methods, we attempted to investigate the **minimal number of exposures** that was required to trigger successful retention and intended to know whether accuracy was positively correlated to the number of presentations during learning.

In order to better understand how memories for sensory information are formed during childhood (and adulthood), for each type of learning, we also investigated the **effect of various variables** on the retention rates (e.g. vocabulary production scores, attentiveness during learning, age and socio professional status (for adults), gender, etc.).

Finally, for each experiment, we also tried to link our findings to the existing knowledge about the putative brain mechanisms provided by the literature as reviewed in the last three chapters, in order to merge behavioral observations with neurosciences. Unfortunately, studies on behavioral research during development rarely address the question of the underlying mechanism, for the simple reason that neuroimaging investigations on young children are sparse, as explained earlier. Developmental researchers can primarily speculate on the basis of discoveries made on adult populations, like amnesic patients or healthy elderly persons. Here, we similarly attempted to find rational explanations for the results we obtained on the basis of current knowledge predominantly gained from non-children populations.

2 Overall materials and methods

This part of the thesis consists of 6 experiments; three are achieved, and three are ongoing projects. Since the same methodology and apparatus are implemented throughout the different experiments, an overall methodological section is developed below. The specificity of the methods for each individual experiment will be recapitulated along with each corresponding experiment.

2.1 Collaborations

This thesis was carried out in collaboration with **Olivier Pascalis** and **Hélène Loevenbruck**, two researchers working at the LPNC (Laboratoire de Psychologie et NeuroCognition, CNRS, UMR5105) of Grenoble. Olivier and Hélène provided their useful expertise and advice which helped in designing the experiments. They also gave precious feedback about the results and proposed perspectives to the work. Each time I visited them, they put their Babylab at my disposal to facilitate the execution of the experiments. About half of the recruited participants underwent our experiments in the Babylab of the LPNC and the second half in the Babylab of CerCo as soon as the ethical committee for the protection of human (Comité de Protection des Personnes, CPP) gave its consent for starting the experiments at CerCo (12/2017). To avoid biases due to the experimenter, the same investigator (myself) ran all experiments. The apparatus and procedure were identical in the two localities and the experimental room and conditions were highly similar (small and cozy rooms adapted for children, toys available in a box for playing during the distractive phase, etc.). Each study involved participants from both localities. Since no differences in performance were found between the two places, we grouped the subjects into a same sample and did not further distinguish this variable in the analyses.

2.2 Participants

Overall, **278 children** and **89 adults** were recruited and underwent one of the six following experiments. Child participants consisted of **29 eighteen-month-olds**, **145 twenty-four-month-olds** and **104 four-year-olds**. In Toulouse, children were recruited in child-care centers and preschools. In Grenoble, children were recruited from the LPNC's internal Babylab dataset of parents who have consented for their children to participate in research studies. For each experiment, a certain number of children could not be included in the analyses for diverse reasons (that will be reported separately for each experiment). Eighteen-month-olds' and twenty-four-month-olds' receptive and productive vocabularies were measured using a French-word checklist (IFDC, a French adaptation of the short MCDI MacArthur-Bates devised by Kern, Langue, Zesiger, & Bovet, 2010). Adults constituted a control group, and were naïve to the purpose of the study to enable comparison with the groups of children. To that aim, adults were initially recruited as survey participants to evaluate the use of touch-screens in preschools. At the end of the experiment, adult subjects were asked whether they had anticipated the final retention test. In total, twelve out of the 89 subjects claimed they did. Their data were not included in the final analyses and additional participants were recruited. The children and adult participants were all native French speakers. Some participants were exposed to a second language in a regular or irregular frequency (that will also be detailed separately for each experiment).

2.3 Ethics

The project as a whole was approved by the French ethical committee for the protection of human subjects (CPP, IdRCB n°2017-A03515-48), and by two local Ethics Committees for Non-Interventional Research (Grenoble: CERNI, IRB00010290-2018-02-06-39; Toulouse: CERNI, 2017-059). All participants and their legal care-takers respectively gave oral and informed written consent before experimentation.

2.4 Stimuli

Prior to the task, color photographs of five familiar and eight novel stimuli were shown to the participant's parents to ensure that the child was familiar with the known objects and completely unfamiliar and unable to label the novel ones. Objects that did not respect these criteria were removed from the set. Familiar objects consisted of a cup, a spoon, a book, a hammer and a car. Novel stimuli consisted of manually modified toys, which do not have a proper label in French. The investigator randomly chose three familiar and three novel objects from the remaining set.

Auditory stimuli consisted of pre-recorded labeling phrases ending with a bi-syllabic pseudo-word repeated three times in a row. A list of twelve child-adequate pseudo-words was generated based on the work by Dohen et al. (2016). These pseudo-words were built so as to include sequences of syllables (each pseudo-word was composed of two syllables; each of them with a different first and second syllable) with high phonotactical probability in French children directed speech. Phonotactical probability was computed based on the methods of French child-directed speech corpora described in Monnin (2010). Pseudo-words were always preceded by an indefinite article (“*un*” or “*une*”, meaning “a”). This list of pseudo-words was also given to the parents prior to the task for novelty approval. Auditory stimuli were recorded by a female native French speaker and normalized for intensity using Audacity®, a free, Cross-Platform Sound Editor. The durations of all pseudo-words were highly similar (M : 490.17 ms, SD : 62.17 ms). The investigator randomly assigned a pseudo-word to each novel object to learn. Pairings of pseudo-words and objects were counterbalanced across participants.

2.4.1 *Learning stimuli*

During the learning session, stimuli were presented in short video sequences, each lasting precisely 15s, during which the object was briefly manipulated by an experimenter’s hand and labeled three times in a row before disappearing [Figure 29]. The first utterance was a simple carrier phrase designed to focus the attention of the child to the object. The following two utterances simply consisted of the pseudo-word preceded by its randomly chosen indefinite article (e.g. “Regarde! Ça c’est un *rivou*, un *rivou*, un *rivou*”; meaning “Look! This is a *rivou*, a *rivou*, a *rivou*”). Since many 18- and 24-month-olds were easily distracted during the learning phase, pronouncing the label 3 times in a row insured that young participants heard the name of each object properly. Moreover, when a child was not focused on the video-clip, the investigator draws his/her attention back to the screen by saying “hey look at here”. Only the hand of the experimenter was seen in the video, manipulating and leaving the object static on the table. The female voice was played when the object was static on the table.

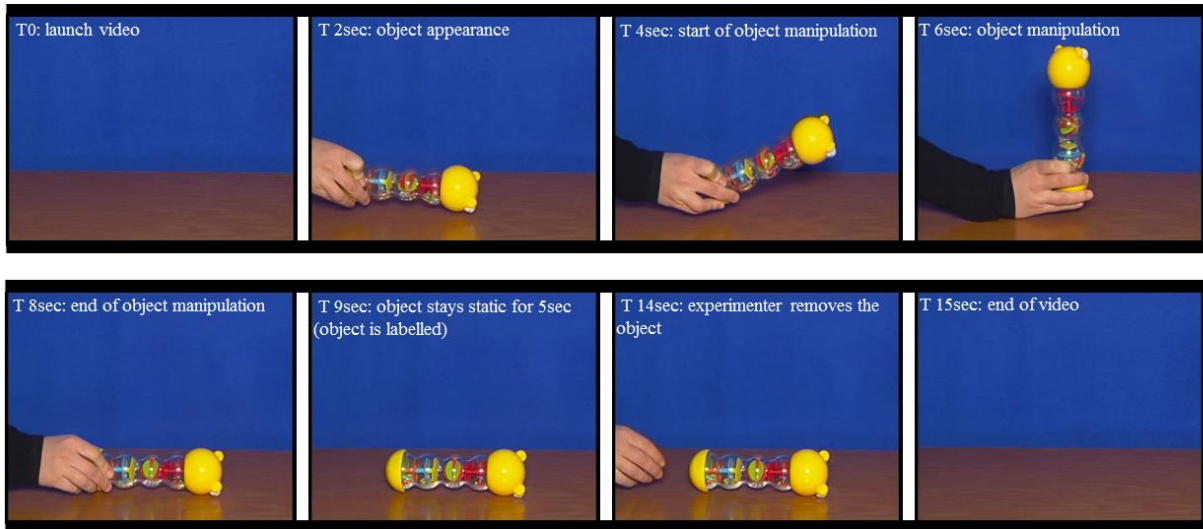


Figure 29 - Illustration of a video sequence used for the learning phase. During the video clip, the novel object was manually displayed, briefly manipulated by the hand of an experimenter, and labelled three times in a row in an ecological utterance while remaining static on the table (e.g., “Regarde! Ça c’est un *rivou*, un *rivou*, un *rivou*”; meaning “Look! This is a *rivou*, a *rivou*, a *rivou*”). The object was removed after being labelled. Video duration: 15s.

2.4.2 Testing stimuli

During the retention session, stimuli consisted of static color photographs of the three objects, simultaneously displayed on the screen. So, testing consisted of 3-Alternative Forced-Choice (3AFC) trials. For each trial, participants heard audio recordings of the same female speaker, asking to touch the image corresponding to the newly learned pseudo-words. The utterances used simple syntactic constructions with a definite determiner preceding the pseudo- words, which are typical of child-adult interactive play in French. The audio recording was played 2100ms after trial onset. One out of three pre-recorded utterances was selected for each child (e.g., “Il est où le *rivou*?”; meaning “Where is the *rivou*?”).

2.5 Apparatus

The experiment was run using a Windows Surface Pro 4 (display size 12.3”, display resolution 2736 x 1824 (5MP)) and generated from a self-developed program under Python software. The touch-screen was placed at arm-distance, facing the child using an articulated mount securely attached on a table.

2.6 Procedure

During the experiment, children sat on a booster seat next to their parents or on their parent’s lap, facing the touch-screen placed at a 70 cm distance to facilitate pointing [Figure 30]. Participants

provided answers by touching on the screen. Parents were instructed to avoid interactions with their children but could encourage them to respond if necessary. To avoid any bias, the investigator stood behind the participant. For the adult group, subjects were instructed to test and evaluate an application designed for children.



Figure 30 – Photograph of a child participant performing the experiment

The study began with three warm-up trials [Figure 31]. During this familiarization training phase, each familiar object was introduced to the participants in short video sequences of 15s duration. During a follow-up 3AFC testing phase, color photographs of the three familiar objects were presented simultaneously on the screen in a triangular configuration, each familiar object serving as target once. As some pilot experiments found that some children could demonstrate a strong bias toward a specific position or object irrespective of the label pronounced, each of these 3 warm-up test trials was correct-answer-blocked, in the sense that only a correct answer could launch the following trial. This warm-up session was immediately followed by the novel word-object pairs learning phase [Figure 31 & Figure 32]. Each novel object was introduced to the participants a variable number of times. This learning phase followed the same procedure as previously described for familiar items.

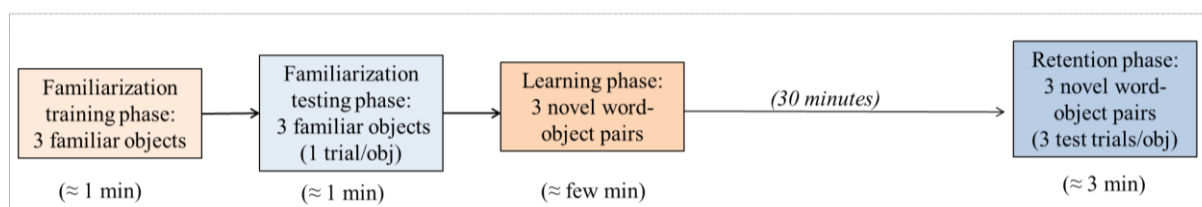


Figure 31 - Schematic of the learning and testing sequences

After learning, children were allowed to play in the experimental room for a 30-min period while their parents completed the IFDC, the French productive vocabulary checklist. During the corresponding period, adult participants were asked to fill in a survey about their general opinion on the use of touch-screens in preschools. Then, they completed 5 different neuropsychological

tests (MOCA, “*Fluence verbale*”, fNART, “*Codes*”, and “*Figures*”) to keep them busy for the entire time interval.

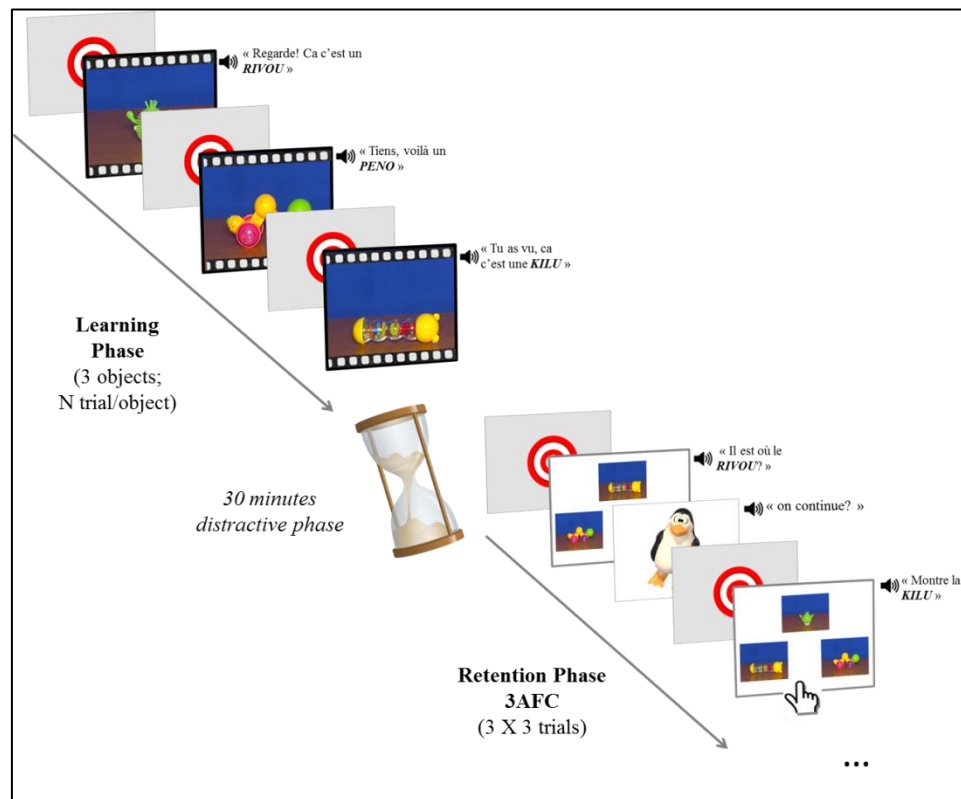


Figure 32 - Illustration of the learning and testing procedures for novel objects. Participants were presented with 3 novel objects paired to arbitrary chosen labels. Participants initiated each learning trial themselves by pressing a red button on the center of the touch screen. During learning, each object appeared in a short video sequence of 15s duration. After a 30-min distractive period, retention was tested using a 3 alternative forced-choice (3AFC) procedure consisting of 9 trials. No feedback was provided but to keep participants focused and to encourage them to continue, each test trial was followed by a 3s GIF animation showing a moving penguin accompanied by a stimulating sentence such as “on continue?” (meaning: “should we continue?”).

Retention test trials for novel objects implemented the same 3 AFC task procedure as with the familiar warm-up trials, except that each novel object served as target three times in an interspersed fashion (testing session = 9 trials) and any answer could trigger the following trial (trials were not correct-answer-blocked in this case) [Figure 31 & Figure 32]. This repeated testing procedure is largely employed in the literature, especially in ostensive naming paradigms (e.g. Bion et al., 2013c; Gurteen, Horne, & Erjavec, 2011a; Hollich et al., 2000; Houston-Price et al., 2005; L. Smith & Yu, 2008; Wojcik, 2017; Woodward, Markman, & Fitzsimmons, 1994a). For both familiar and novel objects’ test trials, the position of the target object was pseudo-randomized with the constraint that the target could not appear at the same screen position for more than two consecutive trials. There were no time constraints for responding but to avoid impulsive responses to be recorded as answers, a minimum delay of 500ms after label onset was required to record the participants’ touching response. Moreover, the verbal instruction was

repeated after a 6s interval if no answer was given. Participants had no feedback on their responses but a neutral GIF animation encouraged them to continue [Figure 32].

2.7 Analyses

2.7.1 *Accuracy during testing*

Tactile touching responses during the retention test trials were automatically recorded and analyzed as a measure of retention. For each age group, the level of performance, calculated as the proportion of hits (i.e. trials for which each participant correctly identified the referent), was compared to levels expected by chance, i.e. 33% (the three newly learned objects appeared simultaneously on the screen at each test trial) in binomial tests. In some cases, one-tailed univariate *t*-tests against chance (33%) were additionally conducted to enable comparison with studies that only use this statistical measurement. Note that this statistical analysis gives a less refined overview of the results as it considers the mean performance instead of individual performance. For each age group, the effect size (Cohen's *D*) of the performance was also calculated. This quantitative measure of the amplitude of the effect, evaluating the strength of the statistical claim, represents a more refined interpretation of the results. The magnitude of the effect sizes given by the values of the Cohen's *D*, suggested by Cohen (1988) and later expanded by Sawilowsky (2009) are the following: very small, $0 \leq d \leq 0.01$; small, $0.01 \leq d \leq 0.2$; medium, $0.2 \leq d \leq 0.5$; large, $0.5 \leq d \leq 0.8$; very large, $0.8 \leq d \leq 1.2$ and huge, $d \geq 1.2$.

2.7.2 *Differences between groups*

Next, to assess whether there were differences in performance between groups, generalized linear mixed-models (GLMM) were performed, including “*age group*” (or “*learning condition*”) as the fixed effect, repeated measures on individuals as a random effect and the hit rate (i.e. the dependent variable) as binomial data (0 or 1). This model compares the performance of the different groups between each other while preserving the data of each test trial of each subject and considering repeated measurements (each participant underwent multiple test trials) as random effects. A GLMM model is the one that best fits binomial data.

Our device also recorded the time participants took to response after label onset (i.e. delay between label onset and tactile response). One-way ANOVAs were used to assess mean time differences between groups. Follow-up Tuckey post-hoc tests indicated where the significant differences originated from.

2.7.3 *Effect of the language level*

For the younger age groups, we collected their receptive (only for 18-month-olds) and productive vocabulary scores (for 18-month-olds and 24-month-olds) on the basis of the IFDC checklist. In order to examine whether there was a correlation between the mean performance score of participants during testing and their level of language in reception and/or production, we conducted linear regressions that kept the continuous gradient of the IFDC scores (i.e. from 0 to 100%). Next, we assigned five language-level categories according to the IFDC scores. Precisely, we allocated the categories as following:

- ***Pre-vocabulary spurt***: $0 \leq \text{IFDC production score (\%)} \leq 20$
- ***Begin vocabulary spurt***: $20 \leq \text{IFDC production score (\%)} \leq 40$
- ***Vocabulary spurt***: $40 \leq \text{IFDC production (\%)} \leq 60$
- ***Advanced vocabulary spurt***: $60 \leq \text{IFDC production score (\%)} \leq 80$
- ***Post-vocabulary spurt***: $80 \leq \text{IFDC production score (\%)} \leq 100$

One-way ANOVAs were used to assess whether there were differences in accuracy between language-level categories. Follow-up Tuckey post-hoc tests were used to determine the origin(s) of the difference(s).

2.7.4 *Effect of attentiveness*

The experiments were videotaped (camera centered on the participants' eyes), which enabled us to measure for each participant his/her score of attentiveness during learning. Using the *VPVideoLab* software developed by researchers of the LPNC research center of Grenoble, two coders reported the number of times children looked away from the screen during the presentations of the videos. In addition, for each participant, they also measured the total looking time outside of the screen, transcribed in percentage (% distractibility). From this value a score of attentiveness could be established. Linear regressions were conducted in order to determine whether accuracy and attentiveness during learning were correlated.

3 FIRST EXPERIMENT. *Ostensive labeling: effect of the number of presentations on word retention*

3.1 Introduction

One claim at the heart of the M4 project is that the memory strength increases roughly linearly with the number of presentations during the initial encoding phase (#claim N°3). As detailed in the beginning on this manuscript, it is assumed that a strengthening process of activated neural networks across repeatedly presented stimuli would result in highly selective neurons (Masquelier & Thorpe, 2007). The connections between the recruited synapses would be reinforced and stabilized across the repeated exposures whereas unreinforced synaptic connections would degrade. It has been claimed that such a biological mechanism would enable long-term memories to remain intact in the absence of reactivation during the intervening period (Larzabal, Bacon-macé, Muratot, & Thorpe, 2017; Thorpe, 2011). At a behavioral level, this claim posits that the more a pattern is repeated during learning, the better would an individual be at remembering it after a delay. In adults, it has indeed been evidenced that performance increased with the number of repetitions of images briefly presented in a stream of hundreds or thousands of visual stimuli displayed in very rapid succession (RSVP, i.e. rapid serial visual presentation) (Thunell & Thorpe, 2019a). Interestingly, the authors found that even only two presentations were sufficient to enable recalling above chance levels after a brief delay of minutes (Thunell & Thorpe, 2019a).

In this experiment, we aimed to know if such a rule would also apply to an early-developing brain system. One might hypothesize that infants and young children are perhaps similarly reliant on repetition to acquire and store knowledge about their surrounding environment. A repetition-based learning mechanism may even be more crucial for infants and young children since their neural system matures and undergoes heavy structural changes during the first years of life, such as dramatic synaptic pruning as reviewed in the previous chapter. In the context of word learning, similar connectionist and associative Hebbian theories have indeed been proposed to support retention of receptive vocabularies (McMurray et al 2012). As detailed earlier, those theories posit that over the course of appearance of bimodal sensory inputs in a temporal contiguity, the connection weights that are originally random and weak may either become strengthened (if recruited) or pruned (if spurious connections) (McMurray et al 2012).

Moreover, considering repetition as a key component of memory formation during childhood also makes sense regarding the developmental segregation between the onset of semantic memories (basically build upon repetitive learning events) and episodic memories (unitary personal events).

As reviewed previously, some authors suggested that declarative memory is primarily semantic, whereas long-lasting episodic memory presumably only emerges after the second year of life, as reflected by the phenomenon of infantile amnesia (Newcombe, 2015). The underlying question is how many exposures to the sensory inputs a young child would require for forming a memory that would survive after a delay. This question, as well as the claim that memory is positively correlated to the number of exposures during encoding, have never been directly tested in children with complex cross-modal stimuli, such as word-object pairs. In the context of ostensive labeling, from the overall literature reviewed above, it can be suggested that the number of exposures necessary to induce comprehension on immediate test trials tends to decrease during development (e.g. Gurteen et al, 2011; Schafer & Plunkett, 1998; Woodward et al, 1994; Bion et al, 2013). Nonetheless, since most of those studies investigated word learning immediately after learning, it is not clear whether the immediate understanding of a word's meaning would have translated into durable memory traces. It might be that the number of exposures required to induce immediate word comprehension may not be sufficient to promote a longer retention.

In this first experiment, we aimed to investigate developmental retention abilities after a 30-min delay, according to a varying number of exposures to the cross-sensory inputs during learning. Specifically, we exposed 18-month-olds, 24-month-olds, 4-year-olds, and adults (control group) to three word-object pairs that were ostensively presented 1, 3 or 5 times each. Retention was measured using a 3AFC procedure after a 30-min distractive period. Children (and adult) participants were kept busy during the entire distractive period (playing, drawing, eating, filling-in a survey, etc.) to minimize the possibility that they would have internally rehearsed the knowledge they just acquired. It is assumed that young children are not able to actively maintain information in their working-memory as adults can do. First of all, because the prefrontal cortex is the latest brain region to mature and is still in a profound maturational stage in children under 5 years of age (see chapter 1 for more details), and second because our young participants were not aware about the following test, and thus had no obvious reason to maintain the information active for a near future use. Therefore, we assume that if retention is observed after a 30-min distractive period, consolidation processes should have already begun during this intervening period.

We hypothesized that children's accuracy to remember the names of objects after this time delay may be positively correlated to the number of times they saw those associations during learning. We also expected differences between age groups, namely the older age groups should perform better than the younger ones. We had no serious expectation about the minimal number of presentation that would be necessary to induce word retention after this 30-min delay, but we

supposed that the older age groups (4-y-olds and adults) would require fewer exposures than the younger age groups (18- and 24-month-olds) to demonstrate retention.

3.2 Methods

3.2.1 Participants

Participants consisted of 26 eighteen-month-olds, 36 twenty-four-month-olds, 28 four-year-olds and 20 adults. Data from 7 eighteen-month-olds were not included in the analyses due to fussiness (n=5), technical problems (n=1), or parental support (n=1). Data from 15 twenty-four-month-olds could not be included in the analyses due to fussiness (n=10), technical problems (n=2), object bias (i.e. selection of the same object on all test trials, n=1), side bias (i.e. selection of the same position of the screen on all test trials, n=1) and failure to engage in the task (n=1). Finally, data from 8 four-year-olds also had to be excluded from the analyses, due to technical problems (i.e. the sound system did not work; n=4), suspected audition troubles (n=1), side bias (n=1) and fussiness (n=2). Details about the final sample are represented in the [Table 1](#).

3.2.2 Procedure

The methodology corresponded to the description provided in the section “Materials and Methods” with the precision that each object appeared a different number of times, interspersed. Specifically, one object appeared only once (i.e. participants watched a unique 15s video-clip during which the object was briefly manipulated and labeled 3 times in a row), another appeared three times and the last one five times interspersed. In total, the learning phase consisted of 9 video-clips presented in a random order. The number of presentations (1, 3 or 5) was randomly attributed to each object. In this within-subjects procedure, participants were their own controls.

Testing occurred after a 30-min distractive period and consisted of 9 trials (3 test trials/object) as described in the “Materials and Methods” section.

3.3 Results

3.3.1 Main results

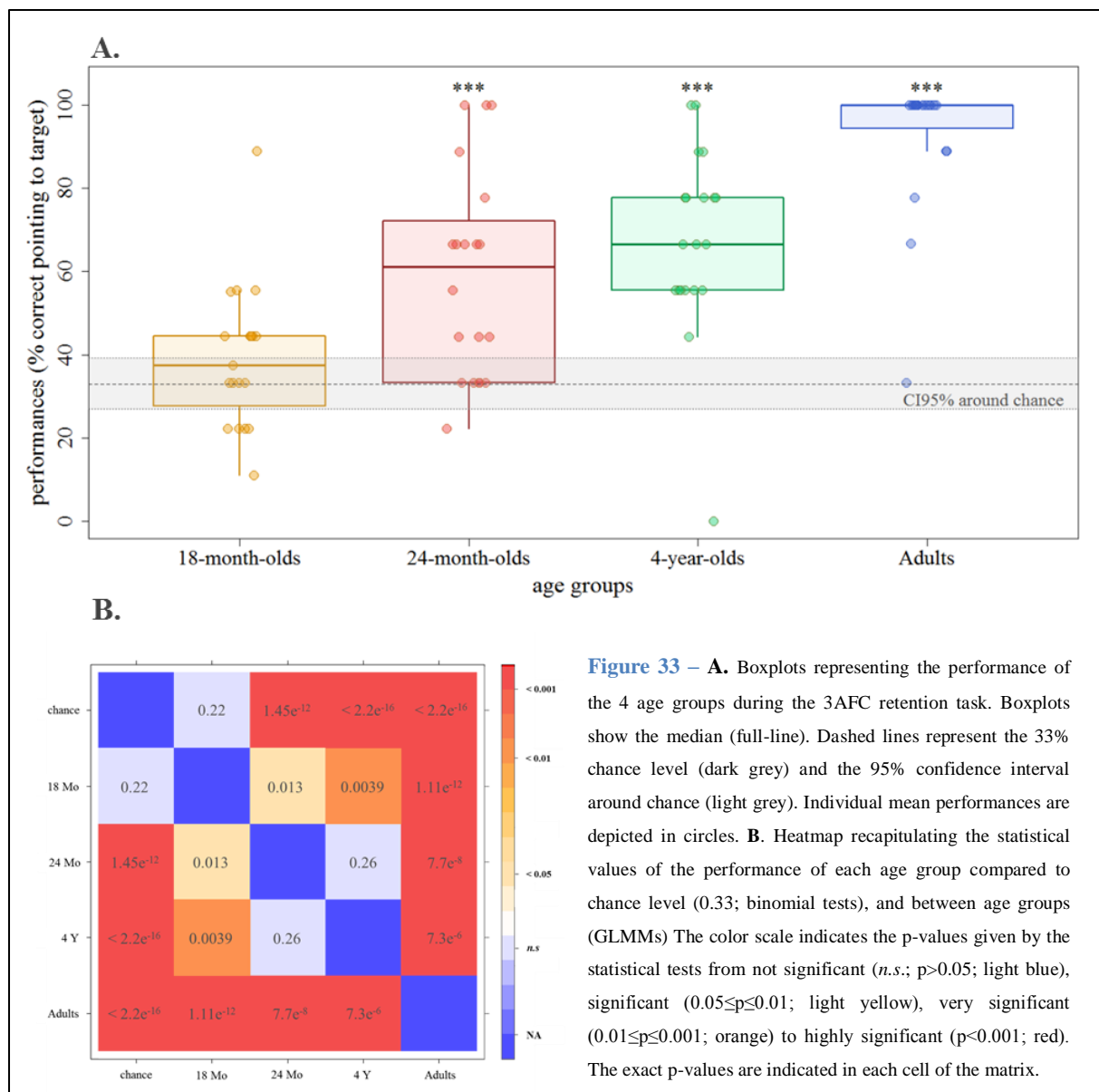
In order to determine whether the number of presentations of cross-modal sensory inputs during learning affects later retention, participants was taught 3 novel word-object pairs to which they were exposed a varying number of times. Each subject saw one object only once, another one three times and the last one five times, interspersed.

Table 1 – Details about the composition and results of each age group.

Age groups	18-Mo	24-Mo	4-Y	Adults
N subjects	19	20	20	20
N females	11	10	9	10
Mean age (\pm SD)	18.23 Mo (\pm 0.98)	24.03 Mo (\pm 0.62)	4.24 Y (\pm 0.31)	31.21 Y (\pm 11.13)
N subjects with older siblings at home	9	7	9	-
N subjects exposed to another language at home	5	7	1	3
Mean overall accuracy (% correct trials) (\pm SD)	39.4 (\pm 17.46)	58.9 (\pm 25)	67.2 (\pm 22.3)	92.8 (\pm 16.6)
<i>P</i> (exact binomial tests)	0.22	$1.45e^{-12}$	$<2.2e^{-16}$	$<2.2e^{-16}$
Cohen's <i>D</i>	0.37	1.04	1.53	3.6
Mean RT (sec) (\pm SD)	6.65 (\pm 3.9)	4.28 (\pm 2.13)	2.63 (\pm 0.76)	1.65 (\pm 0.75)
Mean attentiveness during learning (%) (\pm SD)	88.13 (\pm 6.99)	93.31 (\pm 5.72)	98.69 (\pm 0.59)	-
IFDC score (voc in reception) (%) (\pm SD)	81.9 (\pm 11.7)	-	-	-
IFDC score (voc in production) (%) (\pm SD)	24.8 (\pm 20.1)	69.4 (\pm 25.6)		
Neuropsych-tests scores (%) (\pm SD)	-	-	-	80.23 (\pm 7.91)
Mean Accuracy 1 REP (% correct trials) [CI95%]	29.8 [18.4-43.4]	46.7 [33.7-60.0]	53.3 [39.9-66.3]	88.3 [77.4-95.2]
<i>P</i> 1 REP (binomial tests)	0.67	0.028	0.0013	$<2.2e^{-16}$
Mean Accuracy 3 REP (% correct trials) [CI95%]	41.1 [28.1-55.0]	61.7 [48.2-73.9]	70.0 [56.8-81.1]	93.3 [83.8-98.1]
<i>P</i> 3 REP (binomial tests)	0.20	$6.26e^{-06}$	$6.26e^{-09}$	$<2.2e^{-16}$
Mean Accuracy 5 REP (% correct trials) [CI95%]	47.4 [34.0-61.0]	68.3 [55.0-79.7]	78.3 [65.8-87.9]	96.6 [88.5-99.6]
<i>P</i> 5 REP (binomial tests)	0.02	$4.04e^{-08}$	$7.65e^{-13}$	$<2.2e^{-16}$
Mean RT 1 REP (sec) (\pm SD)	7.48 (\pm 8.84)	4.7 (\pm 3.72)	3.29 (\pm 2.12)	1.87 (\pm 1.45)
Mean RT 3 REP (sec) (\pm SD)	7.63 (\pm 5.53)	4.44 (\pm 2.89)	2.5 (\pm 0.83)	1.96 (\pm 0.91)
Mean RT 5 REP (sec) (\pm SD)	5.95 (\pm 4.65)	3.35 (\pm 1.61)	2.76 (\pm 0.98)	1.57 (\pm 0.54)

Overall results

Overall results demonstrate that 18-month-olds did not choose the targets greater than chance levels ($M= 39.4\%$, $CI_{95\%} = [30.3-45.3]$, exact binomial $p=0.37$) [Table 1 & Figure 33, A]. In contrast, 24-month-olds performed significantly above chance level on this task ($M= 58.9\%$, $CI_{95\%} = [51.3-66.1]$, exact binomial $p<0.001^{***}$) with a very large effect size [Table 1 & Figure 33, A]. Children of 4-years of age similarly chose the target object significantly more often than would have been expected by chance ($M= 67.2\%$, $CI_{95\%} = [59.8-74]$, exact binomial $p<0.001^{***}$) with a huge effect size [Table 1 & Figure 33, A]. As expected, adults performed highly above chance level ($M= 92.8\%$, $CI_{95\%} = [74-86.5]$, exact binomial $p<0.001^{***}$) again with a huge effect size [Table 1 & Figure 33, A]. To summarize, our results show that all age groups except the 18-month-old age group successfully recognized the newly learned word-object pairs after a 30-min delay.



Our apparatus also recorded the time participants took to respond after they heard the label (i.e. time elapsed between label onset and tactile response; Response Time (RT)). Logically, if an individual can recognize a stimulus well, he/she might be expected to respond faster when encountering this stimulus again compared to a poorly retained one. As suggested by the [Figure 34](#), for all age groups except for the 18-month-old age group, there seems to be a correlation between the performance of participants and the time they needed to respond [[Figure 34](#)]. Specifically, the more accurate were the participants, the faster they seemed to respond during testing. However, simple linear regressions did not establish any significant correlation between the hit rate and the RT of 18-month-olds ($F(1,17)=0.69$, $R^2=0.04$, $p=0.4$). Despite the strong trends observable on the graphs, significant correlations could neither be identified for the 24-month-old group ($F(1,18)=3.16$, $R^2=0.15$, $p=0.09$) nor for the adult group ($F(1,18)=0.87$, $R^2=0.04$,

$p=0.36$). A significant correlation was only established for the participants of the 4-year-old group ($F(1,18)=6.89$, $R^2=0.27$, $p=0.017^*$).

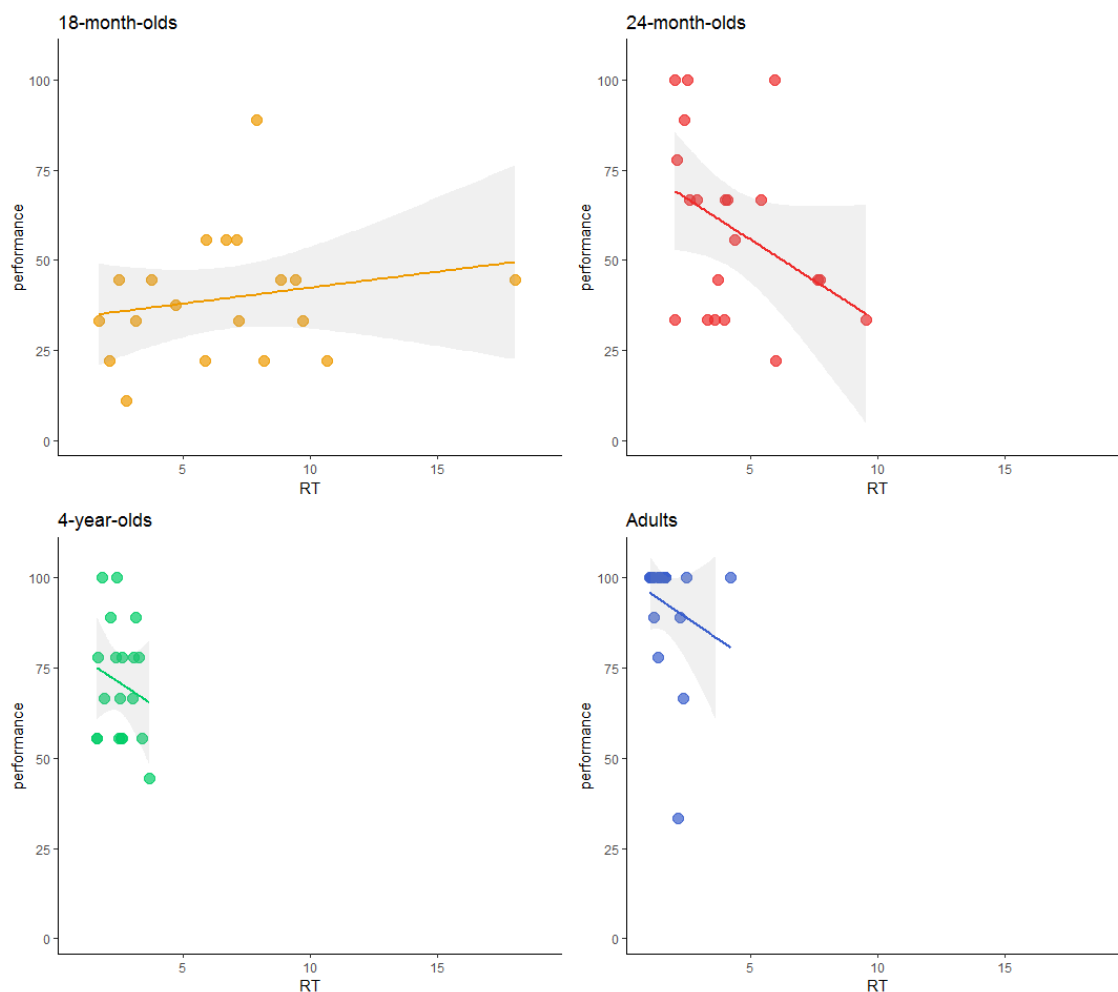


Figure 34 – Graphs showing the mean performance of each individual during testing according to his/her Response Time (RT). The four age groups are represented (18-month-olds (yellow); 24-month-olds (red); 4-year-olds (green); adults (blue)). Each filled circle represents an individual. Lines represent the linear regressions and the grey shadows indicate the 95% confident intervals around mean.

Next, we aimed to determine if the performance and the Response Time significantly differed between groups.

First, a generalized linear mixed model (GLMM) was conducted to investigate whether there were significant differences in performance between age groups. The heatmap of the [Figure 33 \(B.\)](#) summarizes the p-values revealed by the model, and uses a color scale to better visualize the significant differences. Overall, the model indicates that adults significantly outperformed the three child groups. No significant difference was measured between 24-month-olds and 4-year-olds. Finally, the model shows that 18-month-olds performed significantly worse than the older age groups. Further analyses were conducted in order to better understand why 18-month-olds had

overall poorer retention scores than 24-month-olds. These analyses are reported in the “*complementary analyses*” below.

Second, a one-way ANOVA revealed significant RT differences between age groups ($F(3,75)=18.13$, $p=6.03e^{-09}$). Follow-up Tuckey post-hoc tests indicated that the mean Response Time of the 18-month-old participants was significantly longer than for 24-month-olds (adjusted $p=0.008^{**}$), 4-year-olds (adjusted $p<0.001^{***}$) and adult participants (adjusted $p<0.001^{***}$). Twenty-four-month-olds required significantly more time to respond than adults (adjusted $p=0.002^{**}$) but not more than 4-year-olds (adjusted $p=0.1$). Finally, the overall RT of the 4-year-old group did not significantly differ from that of the adult group (adjusted $p=0.5$).

Minimal number of presentations to induce retention

Next, in order to examine the minimal number of exposures that induced a significant retention, binomial tests for each age group and for each pair separately were conducted (see [Table 1](#) for the detailed statistical values). Interestingly, results showed that the pair presented only once triggered significant retention scores in 24-month-olds, 4-year-olds and adults [[Table 1 & Figure 35](#)]. Eighteen-month-olds performed greater than chance levels only for the pair presented five times [[Table 1 & Figure 35](#)].

Effect of the number of presentations on memory

We also investigated whether the performance of each age group was correlated to the number of presentations of the pairings during learning [[Table 1 & Figure 35](#)]. At first glance, and across the different age groups, the ability of participants to recognize the associated name after a 30-min delay does appear to improve and become faster when the object is presented more times during learning. For each age group separately, a GLMM was conducted in order to examine if the performance improved significantly with the number of presentations. For those GLMMs, the “*number of presentations*” corresponded to the fixed effect (1, 3 and 5). For the 18-month-old group, the model revealed a slightly significant increase of the performance for the pair seen 5 times compared to the pair seen only once (GLMM, $Z\text{-value}=1.91$, $p=0.05^{*}$). No significant differences were established between the conditions 1-3 presentations (GLMM, $Z\text{-value}=1.25$, $p=0.2$) and 3-5 presentations (GLMM, $Z\text{-value}=0.67$, $p=0.5$). Similarly, the model indicated that 24-month-olds performed greater for the pair seen 5 times compared to the pair seen only once (GLMM, $Z\text{-value}=2.58$, $p=0.01^{**}$) but no significant differences were established by the model between the conditions 1-3 presentations (GLMM, $Z\text{-value}=1.8$, $p=0.07$) and 3-5 presentations (GLMM, $Z\text{-value}=0.8$, $p=0.42$). For the 4-year-old group, the model revealed significant

differences in the performance between the conditions 1-3 presentations (GLMM, Z -value=2.0, $p=0.04^*$) and 1-5 presentations (GLMM, Z -value=3.01, $p=0.003^{**}$), but not between the conditions 3-5 presentations (GLMM, Z -value=1.11, $p=0.26$). For the adult group, significant differences were only established between the conditions 1-5 presentations (GLMMs, Z -value₁₋₃=1.23, $p=0.2$; Z -value₃₋₅=1.02, $p=0.3$; Z -value₁₋₅=2.02, $p=0.04^*$).

Overall, in accordance with our expectations, repeated presentations of the word-object pairs positively influenced retention in all age groups, especially when the pairs were presented five times compared to a unique presentation. Only in the 4-year-old group was there clear evidence for an increase of performance between the pair seen only once and the one seen three times.

In addition, for each age group, we examined if the number of presentations influenced the RT [Table 1 & Figure 35]. Again, except for the 4-year-old group, the number of presentations did appear to result in faster responses during testing. However, none of these tendencies reached significant levels, indicating in the end that we cannot conclude that participants really responded faster for the pairs seen a higher number of times.

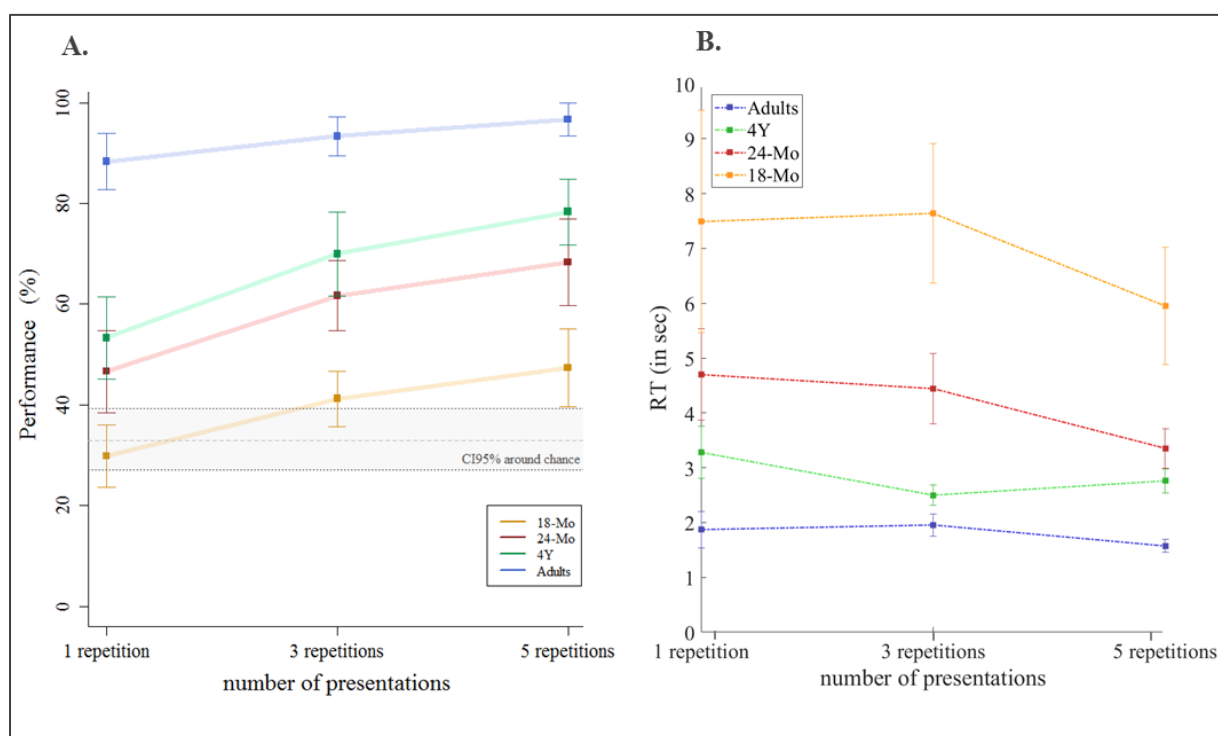


Figure 35 – Graphics representing the mean performance (% of correct touching responses, \pm SEM) (A) and the mean Response Time (RT, \pm SEM) (B) of each age group according to the number of presentations of the pairs during learning (1, 3 and 5 presentations). On graphical A dashed lines represent the 33% chance level (dark grey) and the 95% confidence interval around chance (light grey).

3.3.2 Complementary analyses

Language level

To go further, we sought to determine the source of the difference in overall performance between the 18-month group and the 24-month group. First, we hypothesized that the language level may play an important role in the ability of children to learn and memorize new words after very few exposures.

We found a significant difference in the mean scores of verbal *production* between 18-month-olds ($M=24.8$, $SD=20.1$) and 24-month-olds ($M=69.4$, $SD=25.6$; $t=-6.07$, $p=5.75e-7^{***}$) [Figure 36].

Table 2 – Summary of the number of subjects and mean accuracy of the two younger age groups (18- and 24-month-olds) according to the five language-level categories.

		IFDC production score (%)					
		0	20	40	60	80	100
		Pre-vocabulary spurt	Begin-vocabulary spurt	Vocabulary spurt	Advanced vocabulary spurt	Post-vocabulary spurt	
18 months	N subjects	11	5	1	2	0	
	Mean accuracy (% correct responses)	33.33	44.44	88.88	38.8	-	
24 months	N subjects	0	3	5	3	9	
	Mean accuracy (% correct responses)	-	48.1	68.8	33.33	69.4	

When assigning the verbal production score of each participant to its corresponding language-level category (see “Materials and methods”), we observed that none of the 18-month-old participants reached the ‘post-vocabulary spurt’ category and that most of them were in the “pre-vocabulary spurt” category by the day of the experiment (11 out of 19 subjects) [Table 2]. In contrast, most 24-month-old participants reached the “post-vocabulary spurt” category (9 out of 20) and none of them had a production rate inferior to 20 words (i.e. “pre-vocabulary spurt” category).

In sum, by the day of the experiment, most of the 18-month-old subjects we recruited were pre-vocabulary spurts whereas most of the 24-month-olds yet underwent the vocabulary spurt phenomenon, which is in accordance with the literature.

We were not able to evaluate whether there was also a discrepancy in the lexical *receptive* scores between the two age groups since the IFDC checklist does not include this measurement for 24-month-olds.

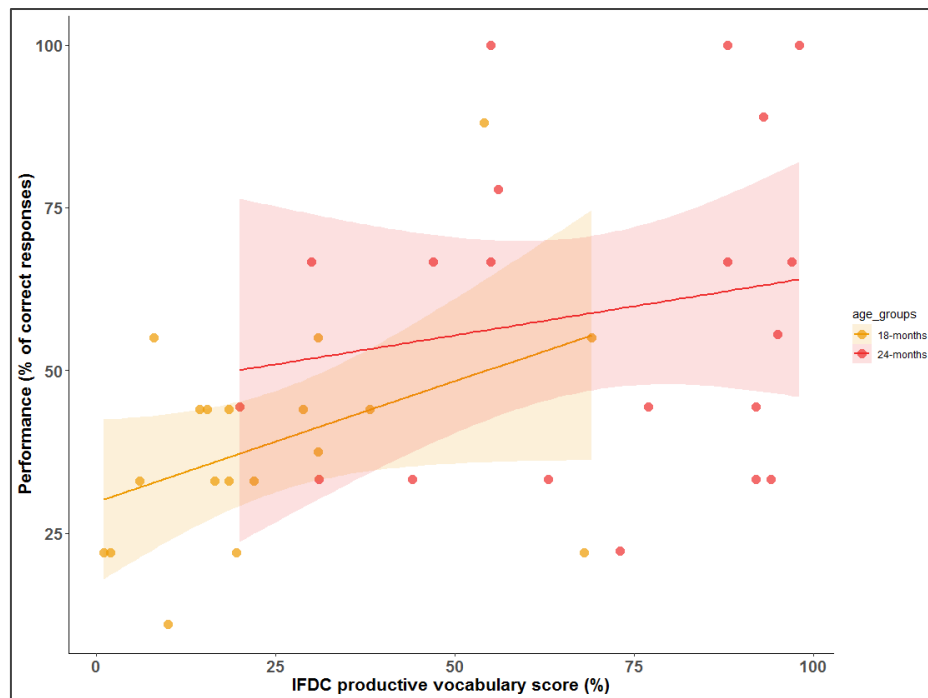


Figure 36 – Graphical representing the mean performance of each individual during testing according to his/her IFDC productive vocabulary score (i.e. a 100-word French checklist). Each individual is depicted in a filled circle (18-month-olds (yellow); 24-month-olds (red)). Lines represent the linear regression calculated by the model for each age group separately (18-month-olds (yellow); 24-month-olds (red)) and the transparent curves indicate the 95% confident intervals around mean.

Next, in order to assess whether receptive and/or productive scores influenced word learning, linear regressions were conducted for the 18- and 24-month-old groups separately (continuous gradient of IFDC scores). For the 18-month-old group, there was nearly a significant correlation between the verbal **production score** of children participants and their mean performance to recognize word-object pairs ($F(1,17)=3.87$, $R^2=0.19$, $p=0.06$) [Figure 36]. However, no significant correlation was established between their **receptive score** and their mean performance during testing ($F(1,17)=0.7$, $R^2=0.03$, $p=0.41$). Note that 18-month-olds' receptive scores were on average very high ($M=81.96\%$, $SD=11.7$) and close to the maximal score attainable (100%) which may explain the absence of correlation. For the 24-month-old group, there was no significant correlation between the verbal **production score** and the mean retention score of participants during testing ($F(1,18)=0.62$, $R^2=0.03$, $p=0.44$) [Figure 36].

When analyzing the data according to the five language-level categories, we found a significant effect of the language level on performance in the 18-month group (AOV, $F(3,15)=5.7$, $p=0.008^{**}$) [Table 2]. Tuckey post-hoc tests indicated that participants from the “vocabulary spurt” category performed significantly better than those from the “pre-vocabulary spurt” category (adjusted $p=0.005^{**}$), the “begin vocabulary spurt” category (adjusted $p=0.02^{*}$) and curiously the “advanced vocabulary spurt” category (adjusted $p=0.03^{*}$). Note however that the vocabulary spurt group consisted of a single subject (an outlier) and the advanced vocabulary

spurt group of only two subjects; thus these results cannot give rise to objective interpretations. For the 24-month group, no significant differences in performance could be established between the subjects allocated to the various language-level categories (AOV, $F(3,16)=1.96$, $p=0.16$).

Attentiveness during learning

We also hypothesized that the difference in performance between 18- and 24-month-olds could arise from differences in their capability to remain concentrated during learning. Specifically, we suspected that 18-month-olds were more distracted during learning than 24-month-olds, which could have affected their ability to retain the associations. To investigate this hypothesis, we measured the percentage of time participants looked away from the screen during learning (analysis of the eye movements). We were then able to calculate a score of attentiveness during learning. We found that 18-month-olds as a group ($M=88.13\%$; $SD=6.99$) were indeed significantly less attentive than 24-month-olds ($M=93.31\%$; $SD=5.72$) (Wilcoxon, $W=185.5$, $p=0.029^*$) [Figure 36].

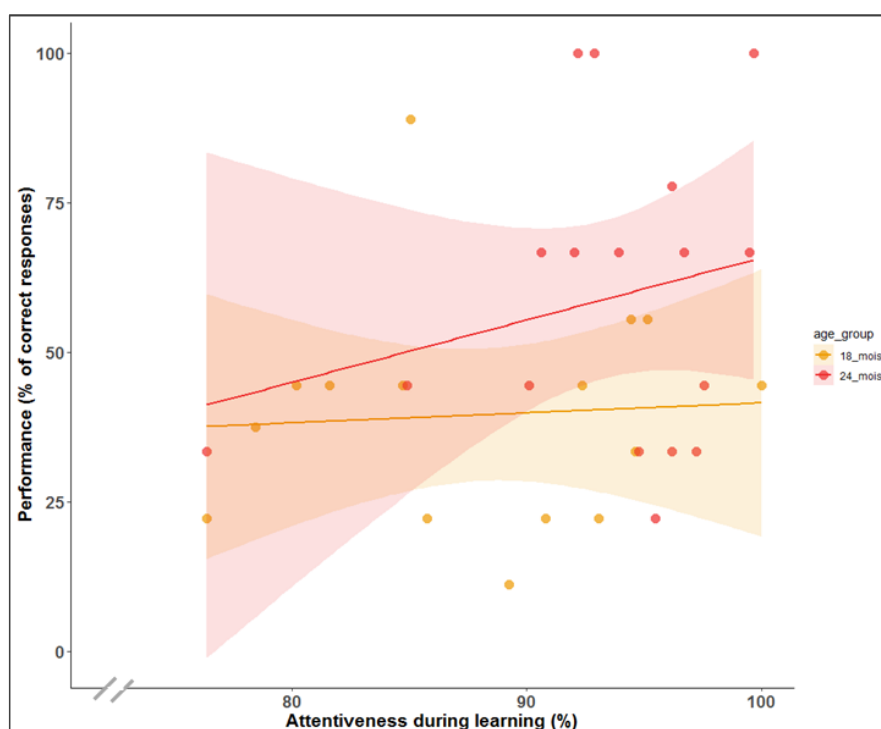


Figure 37 - Graphical representing the mean performance of each individual during testing according to his/her score of attentiveness during learning (in %). Each individual is depicted in a filled circle (18-month-olds (yellow); 24-month-olds (red)). Lines represent the linear regression calculated by the model for each age group separately (18-month-olds (yellow); 24-month-olds (red)) and the transparent curves indicate the 95% confident intervals around mean.

Next, we investigated whether differences in attentiveness affected retention within each age group. Although higher the scores of attentiveness seemed to be associated with higher accuracy in the 24-month-olds, linear regressions did not reveal significant correlations between

participants' scores of attentiveness during learning and their following performance at remembering the names of the objects during testing ($F_{18-Mo}(1,13)=0.05$, $R^2=0.004$, $p=0.82$; $F_{24-Mo}(1,15)=0.87$, $R^2=0.055$, $p=0.36$) [Figure 37]. Note that the overall scores of attentiveness of both samples were very high ($\geq 88\%$ in mean) which may explain the absence of significant differences. Note that participants who were clearly not concentrated during learning were not included in the final sample (they were categorized as “fussiness” participants in the methods) as it is commonly done in the literature.

Inter-individual variability within age groups

Within each age group, our overall results show a high inter-subject variability in performance during testing [Figure 33]. As demonstrated above, for the 18- and 24-month-old groups independently, this variability does not arise from differences in attention during learning, neither from differences in language skills. Linear models that included these two variables concomitantly (“*attentiveness*” and “*language level*” (continuous variables)) did not establish any significant interactions between those two variables ($F_{18-Mo}(3,11)=0.9$, $R^2=0.2$, $p=0.46$; $F_{24-Mo}(3,13)=0.2$, $R^2=0.05$, $p=0.84$).

Moreover, in no age group did we find a gender effect or an age effect. Living with older siblings (i.e. only siblings a few years older were considered) [Table 1] did not appear to influence word learning (GLMM, $Z_{18-Mo}=-0.7$, $p=0.46$; $Z_{24-Mo}=0.69$, $p=0.48$; $Z_{4Y}=0.79$, $p=0.43$). In 24-month-olds, we found a significant effect of being regularly exposed to a second language at home* (GLMM, $Z_{24-Mo}=2.07$, $p=0.038^*$). In adults, there was no correlation between the scores obtained on the neuropsychological tests performed during the distractive period [Table 1] and the retention rates (GLMM, $Z=1.77$, $p=0.07$). In addition, within the adult group we did not find an age effect ($R^2=0.01$, $p=0.67$), and nor was there an effect of the socio-professional status (“*students*” vs “*workers*”). Overall, the various variables we explored did not allow us to accurately determine the source(s) of the inter-individual variability observed within each age group. We only found that being regularly exposed to a second language seems to positively influence word learning in 2-year-olds.

**Note that participants were all native French speakers and that French had to be the dominant language to be included in our study. Before the experiment, we gathered information from the caregivers, in particular the presence of older siblings (and their age) and the exposure rate to another language (e.g. a nanny speaking another language, regular contact with members of the family speaking another language, regular exposure to TV clips in another language, etc.).*

Children with an estimated exposure rate of at least 30% (but less than 50%) to another language constituted the subgroup “second language” of the “language” variable.

3.4 Discussion

In this first experiment, we aimed to investigate whether memory for novel word-object pairs could be correlated to the number of times participants encountered the sensory inputs during encoding. We also attempted to determine the minimal number of exposures participants would require to exhibit retention after a 30-min delay, across development. To address these questions, 18-month-olds, 24-month-olds, 4-year-olds and adults underwent a learning procedure that consisted of 3 novel word-object pairs: one being presented only once, another three times and the last one five times, in an interspersed fashion. Participants’ knowledge about the names of the newly introduced objects was tested after a 30-min distractive period using a 3-alternatives forced-choice task.

First of all, it is important to remind the reader that this study is the first to have exposed participants from such a large developmental scale (18-months to 4-years of age, plus adults) to the same protocol. This enabled us to make fair comparisons between age groups. Indeed, in the literature, most studies focused on only one (or two closed) age groups (e.g. Bion et al., 2013; Gurteen, Horne, & Erjavec, 2011; Horst & Samuelson, 2008; Kucker & Samuelson, 2012; Woodward et al., 1994, etc.). Moreover, given the large diversity of the methodologies used (e.g. forced-choice tasks vs. habituation switch paradigms; looking while listening vs. preferential reaching; real objects vs. static images vs. video clips, etc.) and the parameters manipulated across studies (e.g. number of word-object pairs to learn; number of presentation of each pair; delay between learning and testing, etc.), it is very difficult to impartially compare the results obtained by those isolated studies between each other. Here, we tried to enlighten for the first time, large developmental changes in children’s ability to retain novel information, such as object names, on the basis of a protocol that was equivalent for all.

Ostensive labeling: an influent strategy to promote word retention?

Our overall results demonstrate that all age groups except the 18-month-old group successfully completed the task. Eighteen-month-olds performed on average worse and required more time to respond than the three other age groups. Not surprisingly, adults were on average the most accurate and the fastest to respond during testing. Interestingly, and contrary to our expectations, 4-year-olds did not show better nor faster retention scores than participants two years younger than them. This result is particularly interesting given that in the context of inferential reasoning,

4-year-olds are generally reported capable of learning and remembering the name of a novel object over a long time period (e.g. Carey & Bartlett, 1978; Markson & Bloom, 1997; Waxman & Booth, 2000), whereas in contrast, 2-year-olds commonly fail to recognize the associations on immediate or relatively short-term (~5min) test trials (Horst & Samuelson, 2008). Here, we show that in the context of ostensive naming, 4-year-olds performed in a similar way than 2-year-olds, not better, not faster. In other words, 2-year-olds performed as great as 4-year-olds, indicating that ostensive naming of objects is an influential way to acquire new vocabularies at 2 years of age.

The underlying question is what accounted for these results and for the fact that we highlight a gap between the ability of children aged 18 months and 2 years to remember object names but not between children of 2 and 4 years old? Our complementary analyses provided evidence that 18-month-olds had poorer vocabulary production rates than 2-year-olds and were also significantly more distracted during learning. Even if we could not establish a correlation between language level neither attentiveness during learning and accuracy during testing within each age group separately, our findings strongly suggest that the difference between 18- and 24-month-olds arose from a part from those cognitive variables. Children may need to have undergone the vocabulary spurt phenomenon to rapidly acquire and incorporate new words in their lexicon. As a matter of fact, researchers assumed that lexical growth plays a critical role in the encoding process of word learning (Mervis & Bertrand, 1994). It is also largely admitted that attention positively influences encoding and that this cognitive skill is mediated by the prefrontal cortex, a structure that matures only progressively during development (see Chapter 1 for a review). In line with this idea, we observed that younger children had also generally more difficulty in understanding the instructions (e.g. “touch the object”, “press the button”, etc.). Finally, our apparatus (a touch-screen) was perhaps less adapted for a population of 18 months of age, who often tried to play with it in another way than the purpose we originally intended. Taken together, it is not surprising that the 18-month-olds from our experiment were on average more distracted during learning than the 24-month-olds. Consequently, the failure of the 18-month-olds in demonstrating retention after a 30-min delay may primarily come from a default in their encoding abilities.

In contrast, 2-year-olds reached an attentiveness score of almost 100%, exactly as 4-year-olds. Hence, the comparable retention abilities of 2 and 4 years old may partly be attributable to similar concentration skills. It might also be explained by the fact that most individuals from both age groups were post-vocabulary spurt participants. It might be that a minimal verbal production rate is required for the encoding process but perhaps not for the following consolidation process. For all that, in the first place, we assume that the absence of difference could arise from the learning method itself. As reviewed in Chapter 3, 2-year-olds were not able to retain the names of objects

they fast-mapped by exclusion 5-minutes prior to the test, except if the objects were additionally explicitly labeled a few times by an experimenter holding them or pointing at them (Expe 2, Horst & Samuelson, 2008), or if their attention was drawn toward the referents by illuminating the target and by covering the familiar competitors (Axelsson et al., 2012). Here, our current finding provides strong evidence that ostensive labeling represents a powerful word learning mechanism in 2-year-olds. Since we did not establish any significant difference in the ability of 2- and 4-year-olds to remember object names in the context of ostensive labeling, it can be proposed that older children rely on more refined word learning strategies to rapidly and efficiently increase their vocabularies, including, for example logical exclusion (Bion et al., 2013). This hypothesis will be tested in Study N°3. A last hypothesis is that memory formation of ostensively named objects is supported by brain structures (see Chapter 1), whose maturational stage is not different enough between the ages of 2 and 4 years to trigger a noticeable improvement in word retention following very few exposures. As reviewed earlier (see Chapter 3), it is not clear if semantic knowledge recruits the hippocampal system or the cortices surrounding the hippocampus (entorhinal, perirhinal cortices) or perhaps even only neocortical structures. Although the hippocampus is undoubtedly more mature at 4 years of age than at 2 years, it is possible that the 30min delay implemented in our design does not allow differences to be established. It would be interesting to increase the latency between learning and testing (ideally 24 hours) to see whether in this case 4-year-olds would outperform 2-year-olds.

To conclude on this first part of the discussion, we showed that from 2 years of age, ostensive labeling is an efficient word learning strategy to induce word retention. We demonstrated that very few exposures to objects and their corresponding labels were sufficient to promote recognition after 30 minutes, which is a much longer latency than that used in most previous studies. Since it is unlikely that during the 30min break our young participants internally rehearsed the associations they just acquired, consolidation processes should have begun during the distractive phase. In other words, the results we report here cannot be simply explained by a precocious ability to mentally rehearse the information. This suggests that the synaptic strengthening process would already be operant after very few exposures and very early in life.

Effect of the number of presentations on word retention

In this study, we also addressed the question of whether an individual's ability to recall information after a delay is correlated to the number of times he/she previously encountered that information. This claim has recently been demonstrated in adults using visual stimuli displayed in very rapid succession (Thunell & Thorpe, 2019a). Here, we aimed to know if this claim could also

be fulfilled in a young population, whose brain systems still undergo heavy developmental changes.

In this study, in accordance with our expectations, we showed that the object seen five times was significantly better remembered by all age groups than the object seen only once. Four-year-olds showed better retention scores for the object seen five times compared to the one seen three times, and this latter was also better remembered than the one seen only once.

In adults, in typical psychophysics studies on memory, accuracy is not only measured by a percentage of correct responses, it is also commonly determined by the participants' reaction time. Naturally, the better the subjects remember a stimulus, the faster they respond to this stimulus. Nonetheless, this measure is rarely employed in studies involving young children for the simple reason that this population requires long time ranges to respond and that most protocols are not equipped to record this variable properly (e.g. a live experimenter interacting with the child; a story book as learning support, etc.). Here, one advantage of using a such stringent home-made software coupled with a touch-screen apparatus, is that it precisely and automatically recorded the time that elapsed between the label onsets and the touching responses. Although, we could not establish significant effects, it seemed that participants (in general) responded faster to the most encountered associations.

One might argue that the effect of the number of presentations on word retention that we report in this study can simply be explained by an effect of familiarity. Indeed, the more an object was seen during learning, the more familiar it became for the participants and the more often it could have been chosen by the participants irrespective of the label requests. Hence, participants may have simply better recognized the pair seen five times just because they developed a higher sense of familiarity toward this object. To rigorously rule out this possibility, a control experiment implementing a between-subjects condition may be required. Nonetheless, our within-subjects design has also its advantage because each subject is his own control which wouldn't have been the case with a between-subjects design. Moreover, this issue would truly be acceptable if participants would have failed in the two other conditions, which was not the case here. Since all age groups (except the 18-month-old age group) also successfully recognized the associations encountered only 1 and 3 times, this suspicion can be easily discarded. For the 18-month-olds, there is indeed an uncertainty about this potential confounding factor that should be elucidated in future research.

Minimal number of presentations

This experiment was also designed to investigate the minimal number of times an individual should encounter a cross-modal sensory input to remember it after a delay during development. In a previous study carried out with adults, the authors found that only two presentations of an image briefly flashed among a stream of hundreds or thousands of other images (RSVP) were sufficient to enable recalling after a brief delay of minutes (Thunell & Thorpe, 2019a). This study supports the idea that the human brain can form a memory after a minimal number of exposures to the sensory stimulus. During early childhood, in the context of word learning, the overall literature suggests that the number of times young children should be exposed to the novel sensory inputs to demonstrate retention after a minimal delay decreases with age (see Chapter 3). For instance, Bion et al. showed that 18-month-olds successfully recognized two word-object pairs, which were ostensibly taught to them four times each just before the test (Bion et al., 2013), whereas younger children required at least a dozen of presentations to demonstrate immediate word comprehension (e.g. Gurteen et al., 2011; Schafer & Plunkett, 1998). Thus, from the overall literature, it can be claimed that a repetition-dependent learning process seems to be less vital from 18 months of life, which coincides with the onset of the vocabulary spurt phenomenon. Nonetheless, these studies tested their participants immediately after learning. Little is known about the number of presentations young children would require to manifest retention after a delay.

Here, we implemented a 30-min distractive period during which children were not re-exposed to the three word-object associations. We showed that only the object presented five times during learning allowed significant retention in 18-month-olds. Since most of our 18-month-olds were pre-vocabulary spurt participants, our finding fortifies the idea that a repetition-based learning is required for children younger than the age 1.5 year to maintain a memory after a certain latency and that the vocabulary burst is perhaps an initiator phenomenon for a rapid vocabulary increase. Nonetheless, as already discussed above, we cannot exclude the possibility that 18-month-olds chose the object seen five times significantly greater than chance just because it was the most familiar item. We can only conclude that a unique or even three repetitions of cross-modal inputs were not enough to induce a memory trace in 18-month-olds.

On the other hand, our results indicate that 2- and 4-year olds, as well as adults, performed above chance levels even for the object presented only once. Nevertheless, this result does not prove that a single presentation was sufficient to trigger a significant retention in these age groups. Alternatively, participants may have only well-remembered the names of the objects seen three and five times and simply used a deduction strategy to logically infer the sole item they did not properly remember the allocated name of from the un-retained label. In other words, they may

simply have applied a “referent selection” strategy. As reviewed in Chapter 3, this skill has been shown to emerge between 18- and 24-months of age, specifically after the vocabulary spurt phenomenon (Bion et al., 2013; Halberda, 2003; Mather & Plunkett, 2011; Mather & Plunkett, 2009; Mervis et al. 1994). Thus, on the basis of this study, we cannot conclude about the minimal number of presentations necessary for a memory of novel object names to survive after a 30-min delay. We can only attest that in the context of ostensive labeling, 3 presentations were enough to form a memory trace in children as young as 2 years of age.

In order to disentangle whether participants from 2 years of age can truly memorize the name of an object seen only once, or if they simply use a “referent selection” by exclusion strategy to disambiguate this type of situation, we developed the two following studies. The second study replicates the current one with the specificity that the three novel word-object pairs were all presented a unique time to the participants. This study should objectively assess whether children as young as 2 years of age can learn and remember the names of objects encountered only once. In a third study, we tested the reasoning abilities of 2-year-olds, 4-year-olds and adults to fast-map words to their corresponding referents by exclusion. To that aim, we used a rigorous methodological approach (i.e. the same touch-screen apparatus and the same stimuli). Then we also tested participants’ memory abilities after an equivalent 30-min distractive period.

4 SECOND EXPERIMENT. *24-month-olds and above remember novel object names after a single learning event*

This experiment constitutes an article currently in revision in the journal Cognition

4.1 Introduction

In the previous study, we attempted to examine the minimal number of exposures an early-developing brain system would require to form and maintain after a delay a memory trace between both a novel label and its referent. We showed that retention was positively correlated to the number of presentations participants encountered during learning, and we also found that from 2 years of age, a single presentation triggered significant performance rates. Nonetheless, this study did not allow us to conclude that 2-year-olds effectively remembered the name of the object they saw only once. The alternative option is that they only remembered the names of the two other objects (seen 3 and 5 times) and that they logically inferred the identity of the third item without actually remembering its name. If so, a single learning trial may not be sufficient to induce a memory trace in an immature brain, and in that case, a repetition-based process may still

be the required mechanism for a memory involving complex bi-modal sensory inputs to be formed and to last in time. At first glance, this hypothesis seems quite plausible given that 24-month-olds in previous studies failed to demonstrate retention of object names encountered only once after a five-minute delay (Horst & Samuelson, 2008) or even immediately after learning (Bion, Borovsky, & Fernald, 2013). Nonetheless, in these studies learning consisted of “referent selection” trials, which is a radically different word learning strategy. These studies demonstrated that 2-year-olds cannot form a memory of an object name from a unique “referent selection” trial. However, authors showed that 2-year-olds could retain the mappings for 5 minutes, provided that the objects could be inferred by logical exclusion and in addition deliberately labeled several times by an experimenter holding them up (Expe 2, Horst & Samuelson, 2008). This last finding suggests the aptness of ostensively naming the objects to foster 24-month-olds pairing recall after a minimum delay. Sadly, in the literature, this word learning strategy alone has received much less attention in children above 2 years of age in comparison to typical “fast-mapping” tasks.

It has been claimed that the onset of the disambiguation ability allows children to rapidly acquire new words and increase their lexical repertoire (e.g. Spiegel & Halberda 2011). However, even if ostensive naming has been much less investigated in children above 2 years of age, such children can presumably continue to acquire vocabularies through unambiguous learning situations.

Moreover, since 2-year-olds were shown able to solve the mapping problems but no longer to remember the object names, it has been proposed that the referential ambiguity represents a situation-time problem that must be solved within the context of a single inferential event but that building long lasting linkages would only be solved over developmental time as children encounter multiple presentations of the elements to be mapped (Mcmurray et al., 2012, plus see Chapter 3 of our review) . In other words, word learning from ambiguous situations would be a slow process (Bion et al., 2013).

In the light of these questions, does it mean that forming long lasting memories between sound patterns (words) and visual stimuli (objects) is necessarily a slow, repetition-dependent mechanism? Given the inspiring findings we obtained in our previous study and given the presumed influential efficiency of the ostensive naming learning procedure, we hypothesize that this word learning strategy may be more efficient than “referent selections” to induce actual word learning in 2-year-olds. If so, we hypothesize that whilst a one-trial “referent selection” learning event has been shown unlikely to yield retention at that age, in the context of ostensive labeling, a one-trial learning event should induce word retention after a 30-min delay in 2-year-olds.

4.2 Methods

4.2.1 Participants

Twenty-three 24-month-old children (14 females; M : 24 months 2 days, SD : 20.13 days; range: 22 months 17 days - 25 months 6 days) with a mean productive vocabulary score of 73.5% on the 100-word IFDC checklist (range: 10-99%) completed this study. Data from 4 additional children were not included in the analyses due to fussiness ($n = 3$), and failure to engage in the task ($n = 1$). Twenty 4-year-old children (8 females, M : 4 years 2 months, SD : 3.96 months; range: 3 years 9 months - 4 years 11 months) were also included. One additional 4-year-old child was excluded from the analyses due to a systematic response bias (selection of the image on top of the layout on all test trials). A control group composed of 20 adult participants (13 females, M : 31 years 4 months, SD : 10.12 years; range: 21 years 11 months - 55 years 9 months) who were naïve to the purpose of the study, was used for comparison with the children.

4.2.2 Procedure

The stimuli and overall procedure were identical to the descriptions made in the section “Materials and Methods”. In this second experiment, each of the three objects appeared only once. Hence, learning consisted of three trials (i.e. participants watched one video-clip per object). Testing occurred after a 30-min distractive period and consisted of 9 trials (3 test-trials per object interspersed) exactly as described in the “Materials and Methods” section.

4.3 Results

4.3.1 Mains results

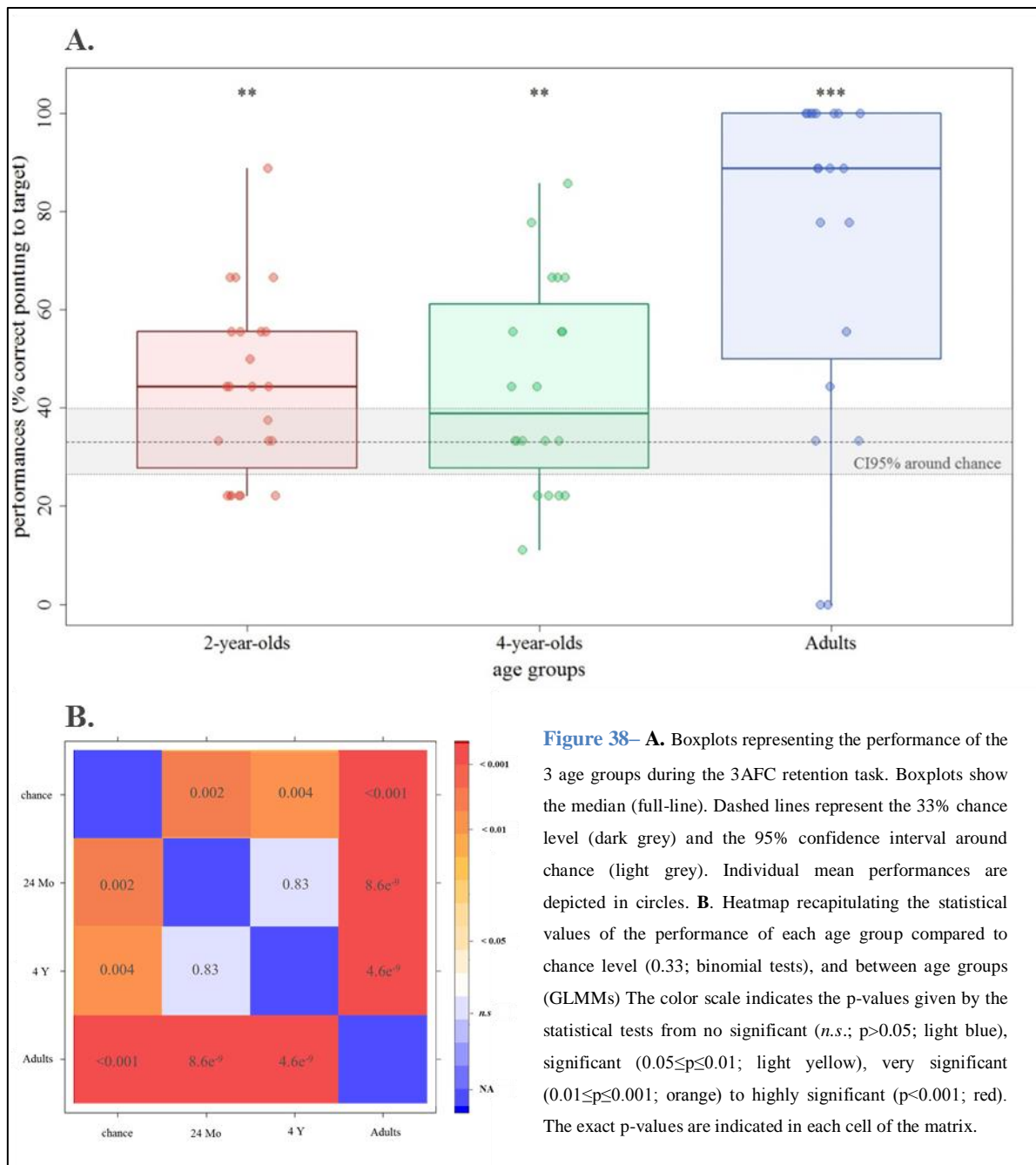
In order to determine whether children were able to form a reliable memory trace of randomly paired word-object associations after a single exposure to each object, we tested performance in a retention test session performed 30 min after the learning session. Touching responses on the screen during the retention test trials were automatically recorded and analyzed as a measure of retention. Among all 63 subjects, two 24-month-olds and one 4-year-old failed to complete the task (8 trials out of 9, $N_{24-Mo} = 1$; 4 trials out of 9, $N_{24-Mo} = 1$; 7 trials out of 9, $N_{4Y} = 1$), but as each novel object appeared as target at least once, we did not exclude their data from the analyses. For each age group, the hit rate, calculated as the proportion of trials for which each child correctly identifies the referent, was compared to levels expected by chance, (i.e. 33%) in binomial tests [Table 3].

Table 3 - Details about the composition and results of each age group.

Age groups	24-Mo	4-Y	Adults
N subjects	23	20	20
N females	14	8	13
Mean age (\pm SD)	24.14 Mo (\pm 0.68)	4.19 Y (\pm 0.33)	31.34 Y (\pm 10.12)
N subjects with older siblings at home	11	5	-
N subjects exposed to another language at home	2	6	8
Mean overall accuracy (% correct trials) (\pm SD)	43.9 (\pm 18.28)	44.3 (\pm 20.99)	73.9 (\pm 33.87)
<i>P</i> (exact binomial tests)	0.002	0.004	<0.0001
Cohen's <i>D</i>	0.60	0.54	1.21
Mean RT (sec) (\pm SD)	5.98 (\pm 5.33)	3.14 (\pm 1.65)	2.65 (\pm 2.03)
Mean attentiveness during learning (%) (\pm SD)	96.67 (\pm 7.43)	96.9 (\pm 3.30)	-
IFDC voc in production) (%) (\pm SD)	73.5 (\pm 26.3)		
Neuropsych-tests scores (%) (\pm SD)	-	-	79.46 (\pm 7.40)

Accuracy during testing

The results show that 24-month-olds performed significantly above chance ($M= 43.9\%$, $SD= 18.28$, $CI_{95\%} = [36.81-50.93]$), exact binomial $p=0.002^{**}$ with a medium effect size (Cohen's $d=0.60$) [Table 3 & Figure 38]. Children of 4-years of age similarly chose the target object significantly more often than would be expected by chance ($M= 44.3\%$, $SD= 20.99$, $CI_{95\%} = [36.41-51.44]$), exact binomial $p=0.004^{**}$, with a medium effect size (Cohen's $d=0.54$). As expected, adults performed well above chance level ($M= 73.89\%$, $SD=33.87$, $CI_{95\%} = [66.83-80.14]$), exact binomial $p<2.2e^{-16}$ ***, with a large effect size (Cohen's $d=1.21$) confirming the feasibility of this task [Table 3 & Figure 38]. Similarly, using one-tailed univariate t -tests against chance (33%) for each age group independently, the mean performance of each age group were significantly above chance level indicating successful retention (2-year-olds: $t(22)=2.77$, $p=0.009^{**}$; 4-y-olds: $t(19)=2.4$, $p=0.02^{*}$; adults: $t(19)=5.4$, $p=3.3e^{-5}$ ***). Altogether, our results show that participants in the three age groups successfully remembered the novel word-object pairs for which they only had a single learning experience.



We also examined the Response Time (RT) participants required to make their choices. For each age group, we aimed to know whether there was a correlation between accuracy and RT during testing. As in the previous study a trend can be observed: the more accurate a participant was, the faster he/she tended to respond [Figure 39]. Simple linear regressions were conducted in order to determine whether these tendencies were significant. A significant effect could only be established in the adult group ($F(1,18)=4.46$, $R^2=0.20$, $p=0.049^*$).

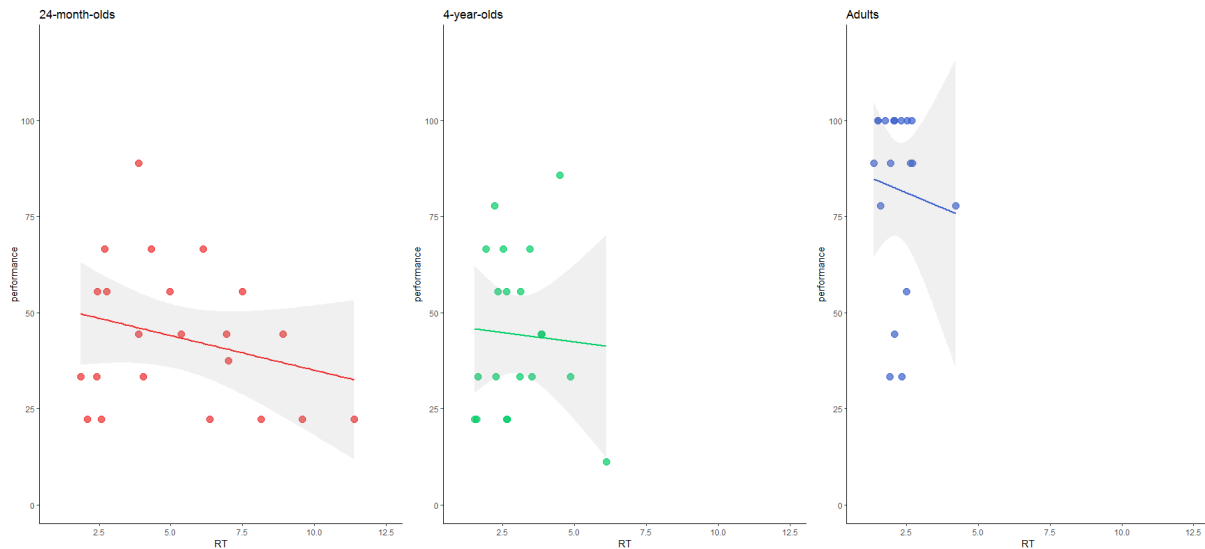


Figure 39 - Graphical representing the mean performance of each individual during testing according to his/her mean Response Time (RT). The three age groups are represented (24-month-olds (red); 4-year-olds (green); adults (blue)). Each filled circle represents an individual. Lines represent the linear regressions and the grey shadows indicate the 95% confident intervals around mean.

Between-groups differences

Next, to assess whether performance improved with age a generalized linear mixed-model (GLMM) was conducted, including *age* as fixed effect (2-year-olds vs. 4-year-olds vs. adults). The model indicates that adults significantly outperformed 2- and 4-year-olds (GLMM, Z -value=4.75, $p<0.001$) but no significant difference was observed between the 2- and 4-year-old groups (GLMM, Z -value=0.2, $p>0.05$) [Figure 38, B.].

Finally, we investigated whether differences in RT could be evidenced between age groups [Figure 38]. A one-way ANOVA revealed significant RT differences between age groups ($F(2,60)=13.8$, $p=1.17e^{-5}$). Follow-up Tukey post-hoc test indicated that all age groups differed between each other in their mean time to respond except between 4-year-olds and adults (adjusted $p=0.64$).

In sum, our results demonstrate that 2- and 4-year-olds performed similarly and significantly above chance level, although 4-year-olds were in mean significantly faster than the 2-year-olds. Adults exhibited the highest retention rate and were in mean as fast as the 4-year-olds to make their touching choices.

Inter-studies comparison

A related aim of the current experiment was to find out if participants from our previous study could have truly learned the name of the object seen once or simply applied an elimination process to solve the task. To address this question from another angle, we compared the current

results with the results obtained for the word-object pair seen only once by the participants of the previous experiment. In other words, we sought to determine whether current participants were as accurate at remembering three word-object pairs to which they were exposed only once as the participants from the previous study who had only one pair to recognize. In none of the three age groups were we able to establish significant differences, although 4-year-olds and adults from the previous experiment seemed to have outperformed those from the current experiment (**2-Y-olds**: $M_{1pair}=46.6\%$, $SD_{1pair}=36.51$, $M_{3pairs}=43.9\%$, $SD_{3pairs}=18.29$, $t=-0.31$, $p=0.7$; **4-Y-olds**: $M_{1pair}=53.3\%$, $SD_{1pair}=36.21$, $M_{3pairs}=44.3\%$, $SD_{3pairs}=20.98$, $t=-0.96$, $p=0.3$; **Adults**: $M_{1pair}=88.3\%$, $SD_{1pair}=24.83$, $M_{3pairs}=73.9\%$, $SD_{3pairs}=33.87$, $t=-1.54$, $p=0.1$).

4.3.2 Complementary analyses

Two-year-olds: language level

To go further, we aimed to determine whether children with larger vocabularies showed better retention than children who speak less (thus who have poorer vocabularies). No significant linear relationship between participants' verbal production score and performance was found ($R^2=0.05$, $p=0.32$) [Figure 40]. Note that most participants were advanced or post-vocabulary spurt children by the day of the experiment (i.e. 16 out of the 23 subjects had an IFDC score superior to 60%) which explains the absence of correlation.

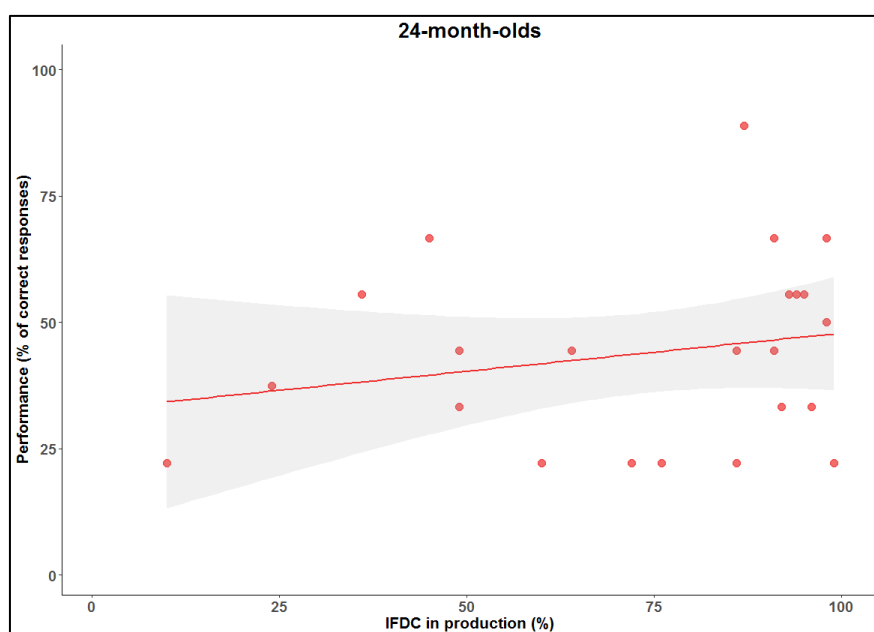


Figure 40 – Graphical showing the mean performance of each 24-month-old individual during testing according to his/her IFDC score (% verbal production). Each individual is depicted in a red filled circle. The line represents the linear regression calculated by the model and the shadow indicates the 95% confident interval around mean.

Two- and 4-year-olds: attentiveness during learning

As in the previous study, in order to assess whether attentional resources during encoding benefited memory formation, we measured the percentage of time 2- and 4-year olds looked away from the screen during learning which then enabled us to calculate their score of attentiveness. As it can be seen, participants in both age groups were on average very attentive in this short experiment (remember that the learning phase consisted of three 15-sec video clips, i.e. less than 1 minute in total; and note also that any child who was too distracted during learning was not included in the analyses). The mean scores of attentiveness were comprised between 90-100% except for one 2-year-old [Table 3 & Figure 41]. There was no significant difference between the two groups. There were no significant correlations between performance and attentiveness in either of the two groups ($F_{2Y}(1,12)=0.3$, $R^2=0.02$, $p=0.59$; $F_{4Y}(1,13)=1.95$, $R^2=0.13$, $p=0.18$), which can be easily explained by the overall high level of attentiveness of the participants.

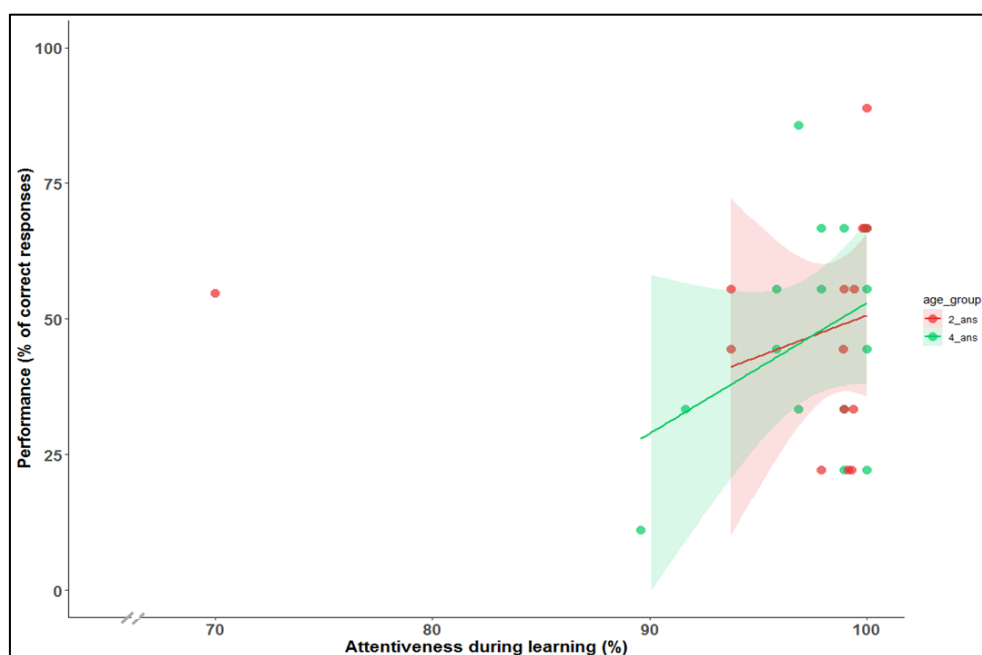


Figure 41 - Graphical representing the mean performance of each participant during testing according to his/her score of attentiveness during learning (in %). Each subject is depicted in a filled circle (2-year-olds (red); 4-year-olds (green)). Lines represent the linear regressions of each age group and the transparent curves indicate the 95% confident intervals around mean.

Inter-individual variability within age groups

Finally, we attempted to highlight the source(s) of the inter-participant variability. As shown above, this variability does not appear to arise from a difference in attention during learning, nor from a difference in verbal production skills in 2-year-olds who were predominantly advanced or post-vocabulary spurt participants. For the 2-year-olds from this experiment, living with older siblings at home also had no significant influence on word learning ($t=1.44$, $p=0.17$). In 4-year-

olds, we were also unable to establish any significant effect of being regularly exposed to another language ($t=0.97$, $p=0.34$). In adults, we found a nearly significant effect of age: in this experiment, mean performance had a strong tendency to diminish with age, and almost 20% of the inter-individual variability could be explained by this *age* variable ($F(1,18)=4$, $R^2=0.18$, $p=0.06$) [Figure 42]. Note however that this tendency is essentially due to few “old” individuals who did not succeed at testing.

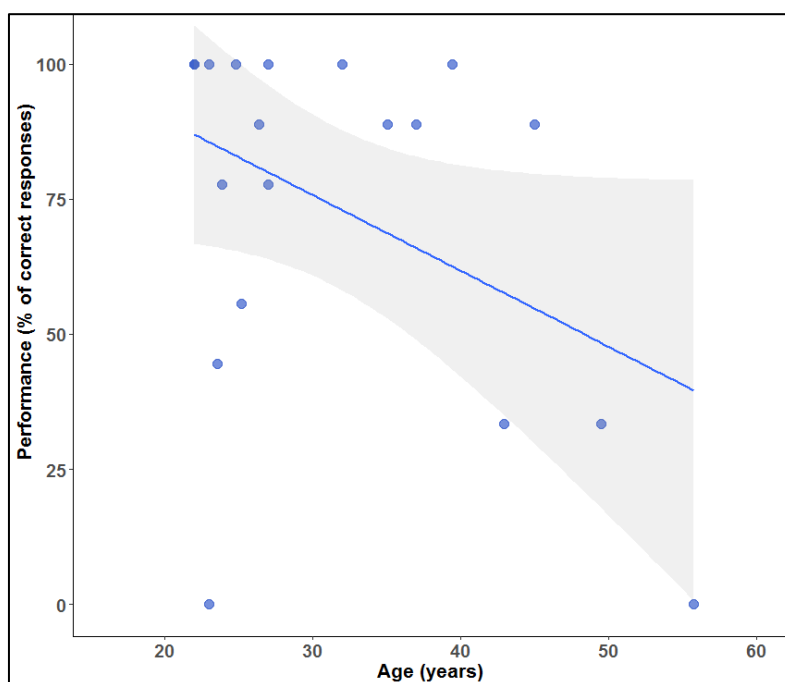


Figure 42 - Graphical showing the mean performance of each adult participant during testing according to his/her age (in years). Each participant is depicted in a blue filled circle. The line represents the linear regression for and the transparent curve indicates the 95% confident interval around mean.

4.4 Discussion

This study asked whether adults and children of 2- and 4-years of age can remember the names of newly learned objects after a 30-minute delay, following a single learning experience. To this purpose, participants were ostensibly taught three novel word-object pairs using short video sequences. More precisely, for each pair, participants watched a 15s video sequence during which the object was manually displayed, briefly manipulated and explicitly labeled 3 times in a row in an ecologically valid sentence before disappearing. Retention was measured using a 3-alternative forced-choice task paradigm, which offered a straightforward measure of whether the name of the object has been retained.

Using this design, our results demonstrate successful retention in children as young as 2 years of age, supporting evidence that a unique learning event involving ostensive naming is sufficient for

a memory between both an object and its associated label to be formed and to survive a delay. We also found that overall, 2- and 4-year old children performed equally well, although 4-year-olds were significantly faster. These two age groups were highly and similarly attentive during learning and this component did not appear to affect later retention. Moreover, there was no effect of productive vocabulary in 2-year-olds, although most had already undergone the vocabulary spurt before being tested. As one would expect, adults performed significantly better than children but were not faster than 4-year-olds.

Overall, our findings suggest that under adequate age-fitted learning strategies, an extremely rapid and successful word learning, similar to that previously described in preschoolers, can already occur in younger children. Consequently, our results highlight the high efficiency of the ostensive naming paradigm if compared to inferential reasoning learning methods (“referent selection” trials) in 2-year-olds.

Ostensive naming vs. “referent selection” by exclusion in 2-year-olds

Previous research showed that when 2-year-olds were facing a learning procedure that involved mapping a novel word to its referent by logically excluding familiar objects (i.e. “referent selection”), they failed to retrieve the fast-mapped words 5 minutes after learning except if the targets were additionally deliberately taught to the children by an experimenter holding them up and pointing at them (Horst & Samuelson, 2008). While word learning from ambiguous situations like “referent selections” have been modeled as being slow processes in 2-year-olds (Bion et al., 2013; McMurray et al., 2012), our finding demonstrates that word learning from unambiguous situations provided with explicit naming is rapid and efficient in children of that age. We demonstrate here that from 2 years of age, a single ostensive learning event is sufficient to form and maintain in memory the association between both a novel word and a novel object, providing the first evidence against a slow, repetition-based word learning mechanism.

Together with previous studies (Axelsson et al., 2012; Horst & Samuelson, 2008), it validates the idea that ostensive naming is an influential learning procedure to promote word retention at that age. Nonetheless, the efficiency of this word learning strategy may evolve during development. One may have expected, for instance, that the ability to encode and retrieve information would have increased with age. On the contrary, the ostensive naming procedure employed in the current experiment did not reveal any significant difference between 2- and 4-year-olds’ capacity to remember object names, although 4-year-olds did have shorter response times during the test trials compared to 2-year-olds. To explain this intriguing finding, the same explanations as those provided in the discussion of the previous experiment can be advanced. Among them, the

possibility that 4-year-olds rely on more refined strategies such as disambiguating language-related situations is the most likely. Indeed, although significant levels were not reached, it seemed that the 4-year-old participants from the previous study were better at recognizing one pair seen once than the participants from the current experiment who had three pairs seen once to retain. Such a difference was not apparent for the 2-year-old group. Consequently, we propose that the 4-year-olds from the previous experiment had a greater ease to use a deductive process to solve the task than 2-year-olds who perhaps only relied on their memory.

In order to examine whether ostensive labeling is effectively more efficient to induce word learning than inferential reasoning at 2 years of age and whether there is a shift in efficiency in these word learning strategies throughout development, we designed the experiment N°4.

Underlying mechanisms

Our work also provides insights about the gradual increase in efficiency of the ostensive naming procedure throughout development. Specifically, it suggests that word learning via ostensive naming gradually shifts from a slow repetition-based process at ages 1-1.5 years (e.g. Gurteen et al., 2011; Schafer & Plunkett, 1998; Woodward et al., 1994) to an extremely rapid and efficient process around two years of age.

This finding raises questions about the putative underlying mechanisms. We hypothesize that this fast learning phenomenon induced by ostensive naming might be attributable to brain maturational processes (peak of synaptic density in the dentate gyrus of the hippocampus reached around two years of age (e.g. Jabès & Nelson, 2015; Seress, 2001)) and/or to general improvements in attention and language acquisition since most 2-year-olds have already undergone the vocabulary spurt and are provided with substantial large receptive vocabularies (see Fenson et al., 1994; Kuhl, 2004). As we have shown, most of the 2-year-olds who participated in this experiment were advanced or post-vocabulary spurt children, and exhibited very high scores of attentiveness during learning, which allowed them to exceed the scores of 18-month-olds of our previous experiment who required at least five exposures to the pairs to demonstrate retention after the same time interval. Note that language and attentional resources are also brain maturational-dependent cognitive functions; therefore improvements of these skills are logically also direct consequences of the maturation of the underlying brain areas and circuitries.

Apart from the brain maturation hypothesis, the finding that declarative memories can be formed and survive without the need of repetition during early childhood is very intriguing given that

repetition has been shown to be a central mechanism of memory formation and maintenance as evidenced by Hebbian and STDP (Spike-Time-Dependent-Plasticity) theories. Remember that those theories are in favor of a strengthening process of activated neural networks across repeatedly presented stimuli, resulting in highly selective neurons and in the reinforcement of the recruited synapses (Masquelier & Thorpe, 2007). Earlier, we hypothesized that children may even be more reliant on repetition, to prevent the synaptic connections responsible for the treatment of the novel sensory input from being pruned. Moreover, it has been suggested that declarative memory is at first essentially build upon repetitive learning events, thus that memory would primarily be semantic (Newcombe, 2015). We also reviewed earlier that in the context of word learning, similar associative Hebbian theories have indeed been proposed to support retention of receptive vocabularies (McMurray et al., 2012). So, what accounted for the results obtained here?

What is repetition?

As a first food for thought, one should address the fundamental following question: what does repetition mean and how do we measure it? Or in other words, what does an event mean and how do we define it in a timeframe? Does repetition imply a succession in time of comparable situations or does it correspond to the number of times one encounters similar sensory inputs? Do we, for example, consider that a 6sec continuous presentation of a visual stimulus is different from three brief presentations of 2sec scattered in time? And if so, should we talk about a single event in the first case and three events in the second? Obviously, repetition (or an event) is subjective, reliant on the context and very difficult to define. An event has a time component with both a beginning and an end. Here the whole learning sequence occurred only once for each object. Nevertheless, during the 15s video clip, participants will almost certainly have switched their eye-gazes many times to explore the whole scene (e.g. they looked at the whole object, then at the experimenter's arm, then focused their attention on the manipulation act, then perhaps focused on a specific detail about the object, etc.). Thus, it might be that the 15s learning event was actually constituted of multiple brief visual repetitions and that each of them contributed to the strengthening process. It would be interesting to investigate whether a shorter visual exposure to each object would have yield the same result. For all that, what we demonstrate here is that word learning during early childhood clearly does not need to be distributed in time to induce memory. However, repetition (in the sense of multiple exposures to the information spaced in time, or simply multiple internal reactivations of the information) may be required for the memory trace to be maintained over longer delays. Further research is required to investigate whether recall can survive with longer delays, including overnight, without any rehearsal in young children. A recent study demonstrated that 30-month-old children were able to retain the names of

objects presented four times each, over a week (Wojcik, 2017). But the participants performed an encoding test immediately after learning which itself induced a rehearsal of the information and thus reinforced the mappings (see Roediger & Butler, 2011 for a review). To look at this question, in another experiment (experiment N°6), we examined the retention skills of some of the 4-year-olds who participated in this experiment after a month delay.

Ecological validity of the methodology

Finally, we would like to emphasize that in the context of ostensive naming, the discrepancy between our finding and the findings of previous studies may arise from the ecological validity of our set up. In the literature, the majority of the studies that employed ostensive labeling as a learning method, used either static images displayed on a screen (e.g. Bion et al., 2013; Schafer & Plunkett, 1998; Schwab & Lew-Williams, 2016; Wojcik, 2017; Zosh et al., 2013) or static objects presented on a tray (e.g. Gurteen et al., 2011), coupled with auditory labels pronounced either in isolation (e.g. Schafer & Plunkett, 1998; Werker et al., 1998) or embedded in a structurally identical phrase (the same carrier frame was used for each stimulus) (e.g. Zosh et al., 2013). In a few studies, word labels were spoken in various carrier phrases (Bion et al., 2013; Wojcik, 2017).

However, in real-life, auditory labels are rarely or never pronounced in isolation, and a child generally manipulates the object that his/her parent refers to, or sees the adult manipulating the targeted object (e.g. “look, I will use this blender to mix your soup”). Previous studies have indeed revealed the importance of providing functional or causal information about the object being labeled to foster later retention (Booth, 2009a; Booth, McGregor, et al., 2008; Markson & Bloom, 1997a). Therefore, the impressive results that we report here can also be explained by the superiority of the ecological nature of our task (although we employed a touch-screen apparatus) compared to participants from previous ostensive naming studies who necessitated more exposures during learning to demonstrate retention.

Overall, we designed an experiment that is a trade-off between (i) a rigorous control of the parameters known to influence learning (especially the external cues) and (ii) an ecologically valid ostensive naming procedure that somehow mimics real-life experiences. On one hand, all participants encountered the exact same learning conditions (e.g. intonation of the speaker’s voice, duration of the learning trials, similar access to the objects’ functions, etc.) that could only be controlled by a computer-based program. On the other hand, the ecological nature of our task relies in the use of video sequences that allowed an access to the potential functions of the objects since they were manipulated. We also recorded several different carrier sentences that used children-directed speech. Finally, our stimuli were 3D objects resembling toys, which may be

more attractive for children than meaningless abstract forms. Consequently, our findings also highlight the extreme importance of the methodological choices when designing an experiment.

4.5 Conclusion

To conclude, we showed that forming a memory trace for associated cross-modal sensory inputs following a single learning experience is possible in children as young as age 2 years. To the best of our knowledge, it is the first time that a demonstration of a declarative memory that is not reliant on repetition has been seen so early in human development. We discussed the possibility that repetition could however be provided by the participants themselves during the 15s video clip. As a perspective work related to this issue, it would be interesting to explore whether the memory trace could survive longer delays without rehearsal. Finally, given the absence of increased performance in 4-year-olds, we proposed that the word learning strategies may evolve during development and that ostensive labeling may not be the sole efficient method to induce retention in older children.

5 THIRD EXPERIMENT. *“fast-mapping”: an influent strategy to promote word retention in 4-year-olds*

5.1 Introduction

In the first experiment, we demonstrated that children of 2-years of age but not of 18-months of age were able to learn and retain object names after a 30-min delay following very few exposures. There was however an uncertainty about the actual retention of the pair presented only once. Children may have effectively stored the name of this object seen once and properly retrieved it during testing. Alternatively, they may have simply applied an elimination process to solve the task. The second experiment showed that 2-year-olds are in fact capable of retaining the names of objects to which they had a single learning experience. It suggests that participants from our first experiment may indeed have remembered the name of the object that was presented only once. Nevertheless, this finding does not completely eliminate the alternative option that children utilized a deductive strategy. In the literature, it has been demonstrated that around 18-months of age children start to disambiguate language-related situations. For example, a previous study measured 14-, 16- and 17-month-olds’ preferential looking behaviors when presented a familiar and a novel object and asked to look at the *[label]* (Halberda, 2003). Results revealed that only 17-month-olds increased looking to the novel object in response to hearing the novel label,

suggesting that the word-learning principles that drive fast-mapping (see Chapter 3) are not operative before this age (Golinkoff et al., 1992; Markman & Wachtel, 1988; Mervis & Bertrand, 1994). In another preferential looking paradigm, Bion et al. showed that 24-month-olds but not 18-month-olds manifested a reliable preference for the novel object on disambiguation trials (Bion et al., 2013). The authors proposed that the emergence of the disambiguation ability coincides with the vocabulary spurt phenomenon (Bion et al., 2013). If so, 2-year-olds from our first experiment - who were predominantly post-vocabulary spurt children - could just as well have used this elimination process. In order to verify if the findings supported by those studies would also apply to our touch-screen apparatus, we confronted 2-year-olds, 4-year-olds and adults with ambiguous word-related situations. Specifically, learning consisted of mapping novel labels to unfamiliar objects presented among known distractors.

The second purpose of the current experiment was to investigate whether this learning strategy can induce retention after a 30-min distractive period. Remember that previous studies showed that 24-month-olds were able to fast-map labels to novel objects but not to recall the associations after a 5-min break (Horst & Samuelson, 2008) nor immediately after learning (Bion et al., 2013). Consequently, the *dynamic associative word learning model* proposed that the referential ambiguity is not isomorphic to learning at that age (McMurray et al., 2012). It may only constitute an online process for immediate communicative requirements without the need for young children to remember the linkages. The authors claimed that building long lasting linkages would only be solved across multiple presentations of the fast-mapped elements (McMurray et al. 2012). Here, we wondered whether multiple “referent selection” trials would enable 2-year-olds to retain the associations after a delay. And if so, what would be the minimal number of presentations required to induce retention? To address these questions, participants were exposed to a varying number of repeats of the ambiguous situations. As in our first experiment, one pair appeared only once (i.e. participants had to fast-map that pair only once), another one three times and the last one five times, in an interspersed fashion.

Since not only 2-year-olds but also 4-year-olds and adults were recruited for this experiment, we should also shed light on the suitability of this inferential learning method in promoting word learning later in development. Previous studies showed that actual learning from a single “referent selection” trial is nearly operative at 30-months of age (Bion et al., 2013) and clearly established in 3- and 4-year-olds (Carey & Bartlett, 1978; Markson & Bloom, 1997; Vlach & Sandhofer, 2012; Waxman & Booth, 2000; Zosh et al., 2013). Thus, in contrast our previous ostensive naming paradigm here we expect 4-year-olds to outperform 2-year-olds.

5.2 Methods

5.2.1 *Participants*

Participants consisted of 29 twenty-four-month-olds, 26 four-year-olds and 21 adults (plus 3 eighteen-month-old pilot participants). Data from 6 twenty-four-month-olds could not be included in the analyses due to fussiness (n=1), technical problems (n=2), failure to engage in the task (n=1), and age limit exceeding (n=2). Finally, data from 1 four-year-old was discarded from the analyses, due to systematic response bias (selection of the same object on all test trials). Details about the final sample are represented in the [Table 4](#).

5.2.2 *Stimuli*

Stimuli consisted of images of three unknown objects and three familiar objects. The objects were randomly chosen but in accordance with the judgment of the caregiver (i.e. familiar objects had to be truly familiar and novel objects truly unknown to the participant). Caregivers also verified that none of the 12 pseudo-words could evoke anything to their child. Three pseudo-words were then randomly chosen by the experimenter and assigned to the three novel objects. Objects and pseudo-words were the same as those utilized in the two previous experiments.

5.2.3 *Procedure*

In this experiment, learning consisted of disambiguating word-related situations. Specifically, participants underwent multiple 3 alternatives forced-choice trials during which they had to map labels onto their referents (i.e. “referent selection” trials). During these trials, the image of a novel object, as well as two images of familiar objects were simultaneously displayed on the screen [[Figure 43](#)].

Exactly 1500ms after trial onset, a pre-recorded voice asked participants to touch an image: for 4 trials, a familiar item was requested, and for 9 trials the novel object was requested (the requests for familiar and novel objects were interspersed). This procedure was essential to ensure that children’s choices were not simply based on novelty preference. For each trial, the label was embedded in a child-directed carrier phrase randomly chosen among three different pre-recorded carrier sentences: “touche le/la [*label*]”; “montre le/la [*label*]”; “il/elle est où le/la [*label*]”. In this experiment, the label was pronounced only once (rather than three times in a row like in the previous ostensive naming paradigms). There were no time constraints for responding but the whole instruction was automatically repeated every 6sec if the participant did not make his/her choice during this time interval. In order to avoid impulsive responses to be recorded as answers,

a minimum delay of 500ms after label onset was required to record the participants' touching response. Moreover, for participants not to learn mistaken associations, those trials were correct-answer-blocked, in the sense that only a correct response could launch the following trial. Participants had no feedback on their responses but a neutral GIF animation encouraged them to continue [Figure 43]. A video camera mounted on a tripod was situated behind the participant's shoulder and recorded the touching responses in order to analyze a posteriori the number of correct "referent selections". Since trials were correct-answer-blocked, two coders analyzed the video recordings retrospectively and indicated as incorrect each trial for which the participants did not touch the target first. The intercoder-reliability was superior to 95%.

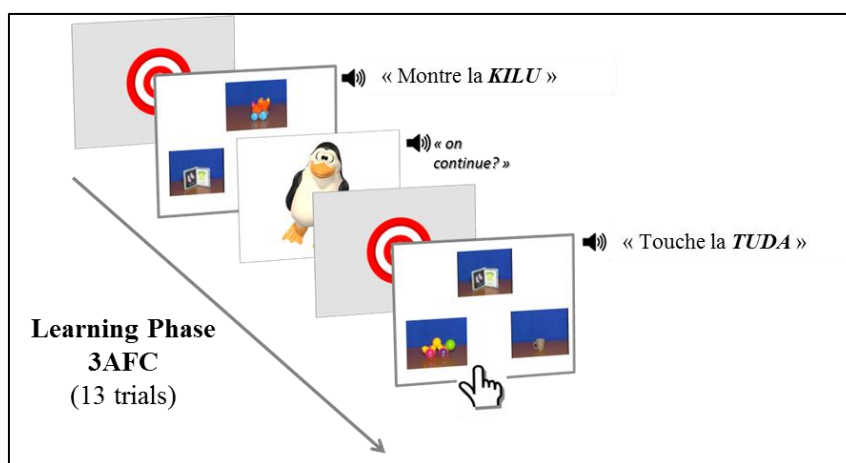


Figure 43 – Illustration of the learning procedure employed in the « fast-mapping » experiment. Participants initiated themselves each “referent selection” trial by pressing a red button on the center of the touch screen. Participants were then presented with 2 familiar objects (randomly chosen among three) and one novel object. Participants were asked to touch either a familiar object (4 trials) or the novel object when hearing an unknown label (9 trials). One novel object served as target on 1 trial, another novel object on 3 trials and the last one on 5 trials, interspersed. No feedback was provided but to keep participants focused and to encourage them to continue, each “referent selection” trial was followed by a 3s GIF animation showing a moving penguin accompanied by a stimulating sentence such as “on continue?” (meaning “should we continue?”).

The experiment began with three warm-up trials. During this familiarization phase, only familiar objects were involved. Each familiar object was targeted once. This familiarization phase was immediately followed by the learning phase per se. One of the three novel objects (N°1) was the target on one trial, another one (N°2) on 3 trials and the last one (N°3) on 5 trials. Overall, the learning phase consisted of 13 “referent selection” trials (1 trial involving object N°1, 3 trials involving object N°2, 5 trials involving object N°3 and 4 trials involving familiar objects) that appeared in an interspersed fashion.

After learning, children participants played in the experimental room for a 30-min period. For the same time period, adult participants completed the neuro-psychological tests and were asked to fill in a survey about their general opinion on the use of touch-screens in preschools.

Testing consisted of 3AFC “referent retention” trials that involved the three novel objects. Novel objects served as target three times in an interspersed manner (testing session = 9 trials) and any answer could trigger the following trial (trials were not correct-answer-blocked in this case).

5.3 Results

5.3.1 Main results

Table 4 - Details about the composition and results of each age group.

Age groups	2-Y	4-Y	Adults
N subjects	23	25	21
N females	13	10	12
Mean age (±SD)	24.63 Mo (±0.74)	4.31 Y (±0.39)	34.14 Y (±14.42)
N subjects with older siblings at home	14	7	-
N subjects exposed to another language at home	4	5	3
IFDC score (voc in production) (%) (±SD)	72 (±22.9)	-	-
Neuropsych-tests scores (%) (±SD)	-	-	78.79 (±6.41)
REFERENT SELECTION			
Mean overall accuracy (% correct trials) [95% CI]	75.69 [69-81.4]	100 (±0)	100 (±0)
P (exact binomial tests)	<2.2e ⁻¹⁶	<2.2e ⁻¹⁶	<2.2e ⁻¹⁶
Cohen’s D	2.55	-	-
Mean overall RT (sec) (±SD)	6.76 (±6.2)	2.10 (±0.54)	2.23 (±1.24)
Mean RT 1 REP (sec) (±SD)	6.97 (±9.5)	2.52 (±0.71)	2.89 (±1.75)
Mean RT 3 REP (sec) (±SD)	6.61 (±4.48)	2.01 (±0.36)	2.01 (±0.8)
Mean RT 5 REP (sec) (±SD)	6.73 (±4.22)	1.89 (±0.35)	1.95 (±0.97)
REFERENT RETENTION			
Mean overall accuracy (% correct trials) (±SD)	40.05 (±17.3)	63.55 (±29)	81.48 (±29.2)
P (exact binomial tests)	0.06	<2.2e ⁻¹⁶	<2.2e ⁻¹⁶
Cohen’s D	0.39	1.04	1.65
Mean RT (sec) (±SD)	4.10 (±2.9)	2.42 (±0.69)	2.02 (±0.65)
Mean Accuracy 1 REP (% correct trials) [95% CI]	38.09 [26.1-51.2]	58.6 [46.7-69.9]	80.95 [69.1-89.7]
P 1 REP (binomial tests)	0.42	6.06e ⁻⁶	6.8e ⁻¹⁵
Mean Accuracy 3 REP (% correct trials) [95% CI]	31.74 [20.6-44.7]	62.6 [50.7-73.6]	84.13 [72.7-92.1]
P 3 REP (binomial tests)	0.89	1.9e ⁻⁷	<2.2e ⁻¹⁶
Mean Accuracy 5 REP (% correct trials) [95% CI]	46.03 [33.3-59.1]	69.3 [57.6-79.5]	79.36 [67.3-88.5]
P 5 REP (binomial tests)	0.03	2.02e ⁻¹⁰	5.2e ⁻¹⁴
Mean RT 1 REP (sec) (±SD)	3.6 (±2.69)	2.49 (±0.78)	2.22 (±0.45)
Mean RT 3 REP (sec) (±SD)	4.51 (±3.48)	2.46 (±0.64)	2.04 (±0.77)
Mean RT 5 REP (sec) (±SD)	4.2 (±2.53)	2.32 (±0.65)	1.79 (±0.62)

In order to determine if during development children are able to disambiguate word-related contexts by mapping new words to their referents by elimination, and whether they can retain the mappings for a latency of 30-min, 2-year-olds, 4-year-olds and adults were firstly exposed to

“referent selection” trials (i.e. learning phase) and then to “referent retention” trials (i.e. testing phase).

Referent selection

The three age groups selected the target objects significantly above chance level, with huge effect sizes (Cohen’s $d > 2$) [Table 4 & Figure 45]. There were no significant differences between age groups in their capability to select the novel objects by exclusion (GLMM, $Z=0.015$, $p=0.9$). However, 2-year-olds were significantly slower to complete the “referent selection” trials than the two other age groups (AOV, $F(2,189)=13.6$, $p=3.9e^{-13}***$; Tukey post-hoc, adjusted $p_{(2Y-4Y)}<0.001***$, adjusted $p_{(2Y-Ad)}<0.001***$) [Figure 44]. In contrast, four-year-olds and adults were equally fast in their mean time to select the targets (adjusted $p=0.98$) [Figure 44].

When analyzing the mean Response Time (RT) according to the number of presentations of the pairs during learning, we found that 4-year-olds and adults responded on average significantly faster for the pair they fast-mapped five times compared to the pair they fast-mapped only once (4-year-olds: AOV, $F(2,66)=9.94$, $p=0.00017***$; Tukey post-hoc; adj $p_{(1-5)}=0.00018***$; adults: AOV; $F(2,55)=3.38$, $p=0.04*$; Tukey post-hoc; adj $p_{(1-5)}=0.05*$). In contrast, 2-year-olds were equally fast irrespective of the number of times they had to disambiguate the situations [Figure 44]. Finally, we found that 2-year-olds were on average significantly slower to respond than the two other age groups whatever the number of presentations, and that 4-year-olds and adults had similar RTs (AOVs, 1 PRESENTATION: $F(2,51)=3.44$, $p=0.039*$, Tuckey post-hoc, adj $p_{(2Y-4Y)}=0.05*$, $p_{(2Y-Ad)}=0.1$, $p_{(4Y-Ad)}=0.97$; 3 PRESENTATIONS: $F(2,66)=23.3$, $p=2.21e^{-8}***$, Tuckey post-hoc, adj $p_{(2Y-4Y)}<0.001***$, $p_{(2Y-Ad)}<0.001***$, $p_{(4Y-Ad)}=0.99$; 5 PRESENTATIONS: $F(2,66)=28.18$, $p=1.43e^{-9}***$, Tuckey post-hoc, adj $p_{(2Y-4Y)}<0.001***$, $p_{(2Y-Ad)}<0.001***$, $p_{(4Y-Ad)}=0.99$).

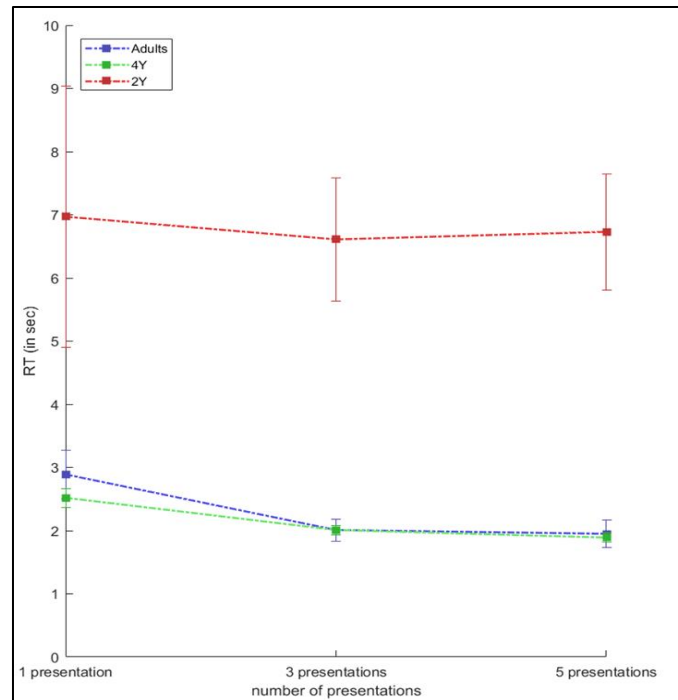


Figure 44 - Graph representing the mean Response Time (RT, \pm SEM) of each age group to select the target objects according to the number of presentations of the word-object pairs during learning (1, 3 and 5 presentations).

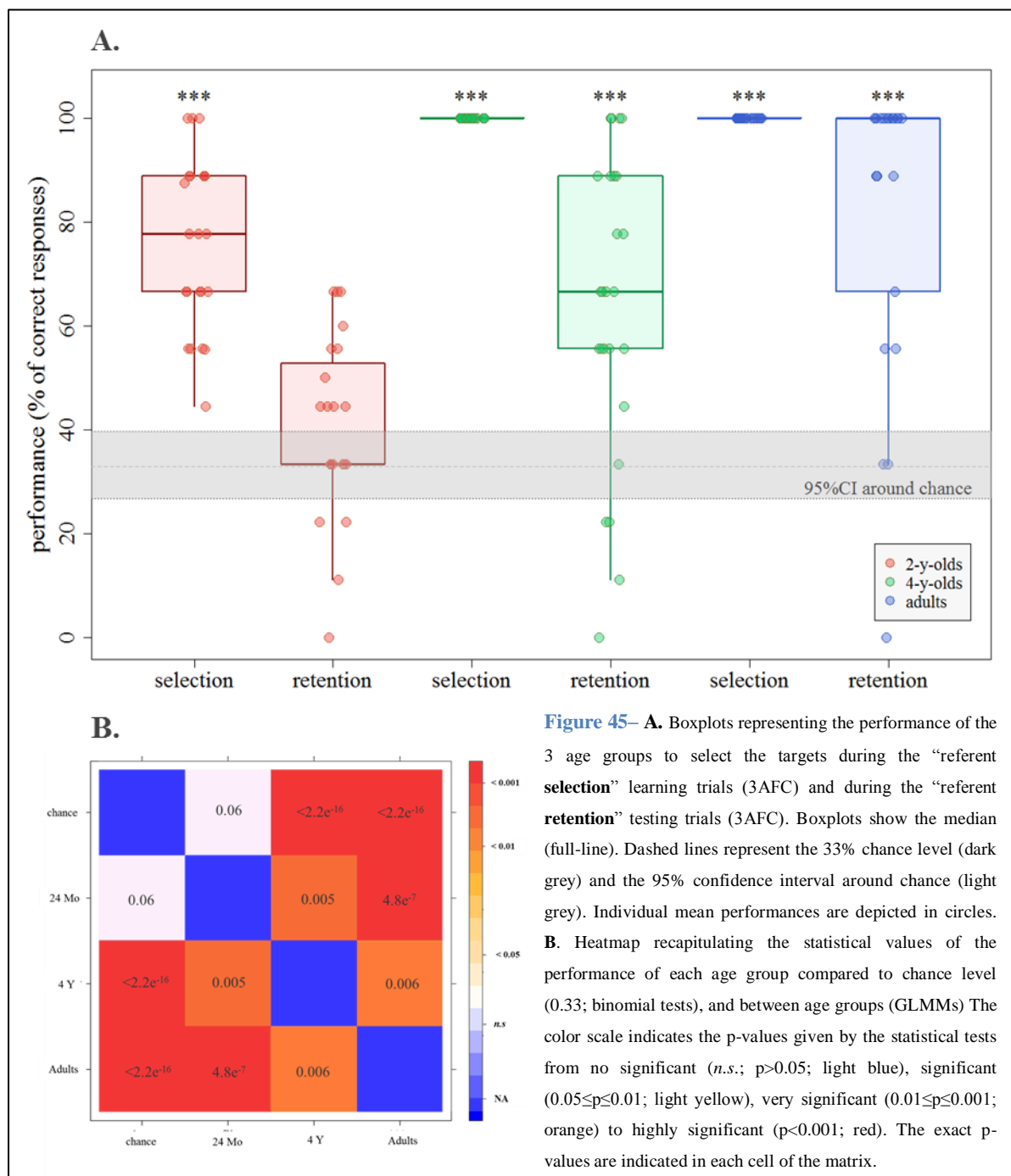
In sum, the three age groups successfully achieved the “referent selection” learning phase, with the particularity that 4-year-olds and adults did not make a single error and required the same time delay to respond. Two-year olds were in contrast significantly slower to respond, and this, whatever the number of times they disambiguated the same word-related problems. We started to include 18-month-olds as well, but the data from three pilot participants revealed that they all failed to solve this mapping phase ($M=29.16\%$).

Referent retention

The memory of the three age groups was tested after a 30-min distractive period. Among all 69 subjects, two 2-year-olds failed to complete the testing phase (6 trials out of 9, $N = 1$; 5 trials out of 9, $N = 1$), but as each novel object appeared as target at least once, we did not exclude their data from the analyses.

Results show that 2-year-olds did not perform significantly above chance ($M=40.05\%$, $SD=17$, exact binomial $p=0.06$) [Table 4 & Figure 45]. In contrast, children of 4-years of age and adults often recognized the fast-mapped associations (4-year-olds: $M=63.55$, $SD=29$, very large effect size ($d=1.04$), exact binomial $p<0.001^{***}$; adults: $M=81.48$, $SD=29.2$, huge effect size ($d=1.65$), exact binomial $p<0.001^{***}$) [Table 4 & Figure 45]. When using one-tailed univariate t -tests against chance (for comparison ease with studies from the literature that only use this statistical

tool), the same degrees of significance were found (2-year-olds: $t(22)=1.95$, $p=0.065$; 4-year-olds: $t(24)=5.27$, $p=2.07e^{-5***}$; adults: $t(20)=7.59$, $p=2.58e^{-7***}$).



Then, a generalized linear mixed-model was conducted in order to examine whether differences in performance could be established between age groups. The model indicated that 4-year-olds outperformed 2-year-olds (GLMM, $Z=2.76$, $p=0.005^{**}$) [Figure 45, B.]. Adults similarly outperformed 2-year-olds (GLMM, $Z=5.03$, $p=4.8e^{-7***}$) but also 4-year-olds (GLMM, $Z=2.75$, $p=0.006^{**}$) [Figure 45, B.]. Note that when using a one-way ANOVA (which considers the

mean accuracy of each individual rather than the accuracy of each single trial of each individual and which doesn't take into account the repeated measurements on individuals), significant differences were only established between 2- and 4-year-olds ($F(1,66)=14.35$, $p=6.7e^{-6}$; Tukey post-hoc, $\text{adj } p_{(2Y-4Y)}=0.0067^{**}$) and between 2-year-olds and adults ($\text{adj } p_{(2Y-Ad)}<0.001^{***}$). The mean accuracy of 4-year-olds was not significantly different from that of adults ($\text{adj } p_{(Ad-4Y)}=0.06$) with this statistical tool.

Moreover, 2-year-olds were significantly slower to complete the “referent retention” trials than 4-year-olds and adults (AOV, $F(2,203)=26.4$, $p=6.6e^{-11}$, Tukey post-hoc, adjusted $p_{(2Y-4Y)}<0.001^{***}$, adjusted $p_{(2Y-Ad)}<0.001^{***}$). On the other hand, four-year-olds responded as fast as adults (adjusted $p=0.37$) [Table 4].

In sum, although the three age groups were similarly successful at disambiguating the contexts during learning, only 4-year-olds and adults successfully remembered the associations after a 30-min delay. Furthermore, 4-year-olds and adults required an equivalent RT during testing that was significantly below the time 2-year-olds needed to make their touching responses.

Minimal number of presentations to induce retention

Next, we aimed to know if the number of times each object was fast-mapped by exclusion during learning influenced the subsequent retention of the associations. Specifically, we first attempted to determine whether a minimum number of presentations was necessary to promote retention, especially for 2-year-olds who failed on the whole to solve the “referent retention” task.

Our results reveal that when 2-year-olds fast-mapped a label onto its referent only once ($M=38.1$, exact binomial $p=0.42$) or even three times ($M=31.7$, exact binomial $p=0.89$), they were unlikely to recognize the associations after the 30min break [Table 4 & Figure 46]. However, five “referent selection” trials involving the same target induced retention in 2-year-olds ($M=46.03$, exact binomial $p=0.03^{*}$). In contrast, 4-year-olds and adults demonstrated significant retention regardless of the number of times the objects were selected during the learning phase. In other words, for these two groups, a significant level of performance was reached even for the object that was fast-mapped only once during the previous “referent selection” trials [Table 4 & Figure 46].

Results according to the varying number of times each pair was fast-mapped during learning

Second, we aimed to know whether retention was positively correlated to the number of times participants fast-mapped each label onto its corresponding referent. Generalized linear mixed-model were conducted and did not establish any significant relation between the *accuracy* and the *number of presentations* in any of the three age groups (2-year-olds: $Z_{1-3}=0.75$, $p=0.5$; $Z_{1-5}=0.9$, $p=0.4$; $Z_{3-5}=1.64$, $p=0.09$; 4-year-olds: $Z_{1-3}=0.59$, $p=0.5$; $Z_{1-5}=1.6$, $p=0.1$; $Z_{3-5}=1.02$, $p=0.3$; adults: $Z_{1-3}=0.64$, $p=0.5$; $Z_{1-5}=0.32$, $p=0.8$; $Z_{3-5}=-0.96$, $p=0.3$) [Table 4 & Figure 46].

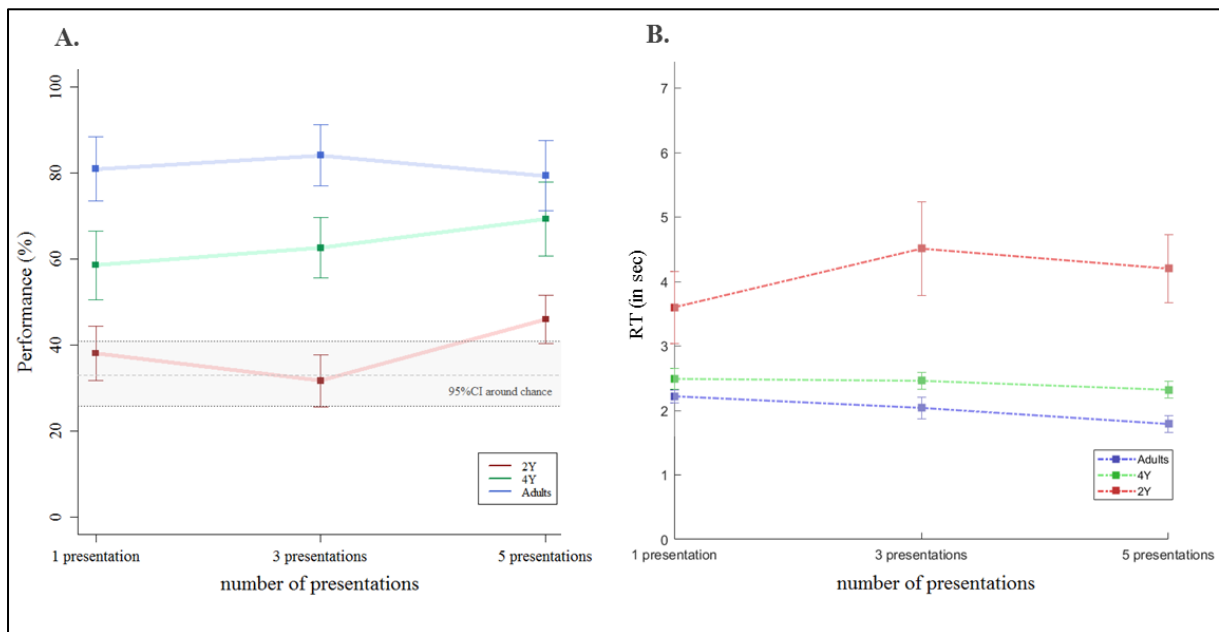


Figure 46 - Graphs representing the mean performance (% of correct touching responses, \pm SEM) (A) and the mean Response Time (RT, \pm SEM) (B) of each age group according to the number of selections of each object during learning (1, 3 and 5 “referent selection” trials). On the graph A dashed lines represent the 33% chance level (light grey) and the 95% confidence interval around chance (dark grey).

Moreover, rehearsing the mappings during learning had no effect on the response time participants required to make their choices during testing; i.e. participants did not respond faster for the pair mapped 3 or 5 times compared to the pair mapped only once (AOVs, 2-year-olds: $F(2,72)=0.57$, $p=0.56$; 4-year-olds: $F(2,72)=0.45$, $p=0.64$; adults: $F(2,60)=2.43$, $p=0.09$) [Table 4 & Figure 46]. Nonetheless, 4-year-olds and adults were significantly faster to respond than 2-year-olds whatever the number of presentations of the pairings encountered during learning (AOVs, 1 PRESENTATION: $F(2,66)=4.47$, $p=0.015^*$, Tuckey post-hoc, adj $p_{(2Y-4Y)}=0.05^*$, $p_{(2Y-Ad)}=0.018^*$, $p_{(4Y-Ad)}=0.84$; 3 PRESENTATIONS: $F(2,66)=9.07$, $p=0.0003^{***}$, Tuckey post-hoc, adj $p_{(2Y-4Y)}=0.003^{**}$, $p_{(2Y-Ad)}=0.0006^{***}$, $p_{(4Y-Ad)}=0.77$; 5 PRESENTATIONS: $F(2,65)=14.86$, $p=4.84e^{-6}^{***}$, Tuckey post-hoc, adj $p_{(2Y-4Y)}=0.0002^{***}$, $p_{(2Y-Ad)}<0.001^{***}$, $p_{(4Y-Ad)}=0.49$) [Figure 46].

5.3.2 Complementary analyses

Two-year-olds: language level

We conducted linear models in order to assess if 2-year-olds' verbal production skills influenced their capability to select the novel object during the “referent selection” trials and to remember the associations during the following retention trials. Despite the noticeable trend [Figure 47], no significant correlations between *IFDC* productive scores and “referent selection” skills ($F(1,19)=1.53$, $R^2=0.07$, $p=0.2$) neither between *IFDC* productive scores and “referent retention” skills ($F(1,20)=0.4$, $R^2=0.02$, $p=0.5$) could be established. Note that this absence of significant correlation is probably due to the fact that most 2-year-olds participants were already advanced ($N=4$) or post-vocabulary spurt ($N=12$) children. Only three participants were either pre-vocabulary spurt children ($N=1$) or at the onset of this phenomenon ($N=2$). The three pilot participants of 18-months of age had verbal production scores of less than 10% ($M=7\%$).

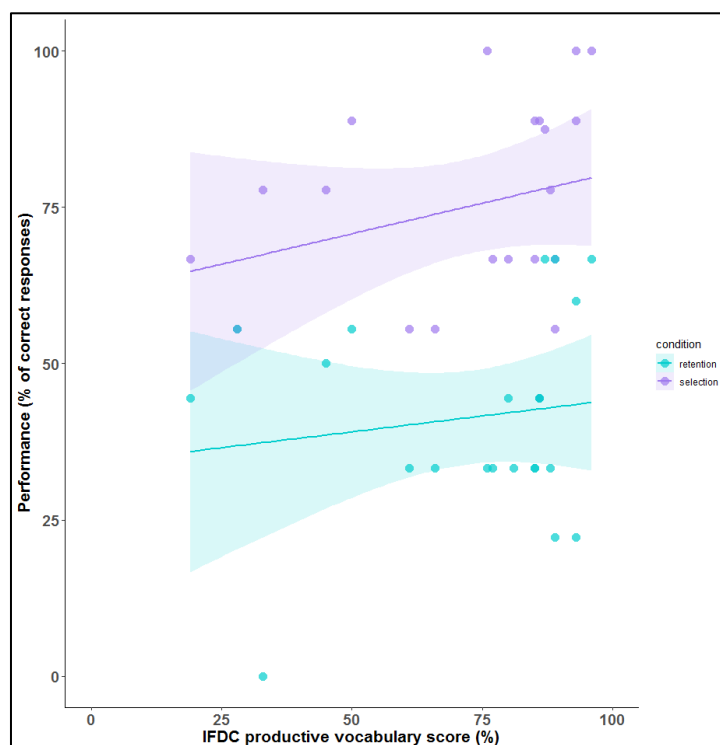


Figure 47 - Graph showing the mean performance of 2-year-old participants to select the referents by exclusion during learning (“selection”, violet) and to remember the names of the referents during testing (“retention”, turquoise-blue) according to their IFDC productive scores (%). Each individual is depicted in a filled circle. The lines represent the linear regressions calculated by the models and the shadows indicate the 95% confident interval around mean.

Two-year-olds: correlation between “referent selection” and “referent retention”?

To foster our comprehension about the mechanisms that drive learning through ambiguous contexts, we attempted to know if the child’s ability to store and remember the linkages after a delay is directly linked to their capacity to disambiguate the word-referent situations during

learning. Since 4-year-olds and adults were 100% accurate during learning, we could not address this question to these 2 age groups. However, we wanted to explore if the 2-year-olds who exhibited the highest fast-mapping accuracies were those that were best at recalling the names of these objects during the following test. As it can clearly be noticed on the graph, there was no significant correlation between 2-year-olds' ability to infer words to their referents by exclusion and the retention of these associations ($F(1,20)=0.059$, $R^2=0.0027$, $p=0.8$) [Figure 48]. Note that this result may be explained by the fact that most 2-year-olds very well completed the “referent selection” learning phase (mean performance > 50%) [Figure 48].

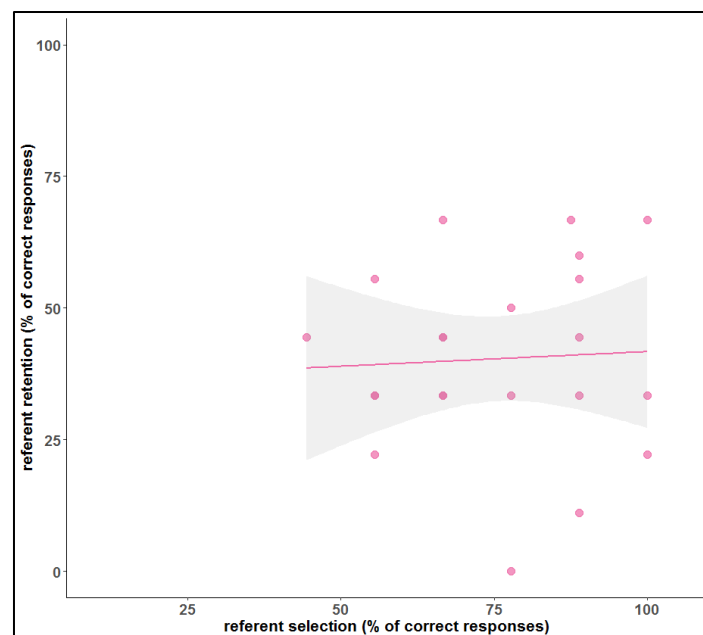


Figure 48 - Graph illustrating 2-year-olds' performance during testing (% of correct responses) according to their ability at disambiguating word-referent contexts during learning (% of correct responses). Each individual is depicted in a pink filled circle. The line represents the linear regression calculated by the model and the shadow indicates the 95% confident interval around mean.

Two-year-olds: looking behaviors during learning

Furthermore, we also investigated the source of the difference in performance during testing between 2- and 4-year-olds. Specifically, we hypothesized that the difference in performance might arise from differences in attending at the images during learning. Previously, we showed that 2-year-olds required significantly more time to make their touching responses during learning. Despite this, is this RT increase due to an increased time spending to look at the target or primarily at the distractors?

To address this question, five 2-year-old and five 4-year-old participants were videotaped from the front (rather than from the back). This allowed us to record their eye movements and thus to measure a posteriori their preferential looking behaviors toward each object during learning (see Zosh et al., 2013 for similar procedure). Precisely, we used the VPVideoLab software developed

by the LPNC research center, to measure the exact time spent looking at each image during each “referent selection” trial. In total, 85 “referent selection” trials were analyzed (42 involving 2-year-old participants and 43 involving 4-year-olds participants).

We found that 2-year-olds spent on average 55% of their total looking time attending at the target, against 60.7% for the 4-year-olds. We also established that the 2-year-olds looked on average 1 second longer at the target images than the 4-year-olds, which is significant ($M_{2Y}=3.77$, $SD_{2Y}=2.4$; $M_{4Y}=2.71$, $SD_{4Y}=0.9$; $t(52)=2.6$, $p=0.01^*$) [Figure 49]. Furthermore, our analyses revealed that they spent significantly more time looking at the distractors relative to the 4-year-olds; specifically 2-year-olds looked almost twice as long as the 4-year-olds at the distractors before touching the screen ($M_{2Y}=3.03$, $SD_{2Y}=2.69$; $M_{4Y}=1.75$, $SD_{4Y}=0.88$; $t(50)=2.9$, $p=0.005^{**}$) [Figure 49]. However, 2- and 4-year-olds made on average a comparable number of visual saccades during the “referent selection” trials ($M_{2Y}=4.38$, $SD_{2Y}=4.4$; $M_{4Y}=4.09$, $SD_{4Y}=2.03$; $t(57)=0.38$, $p=0.7$) [Figure 49].

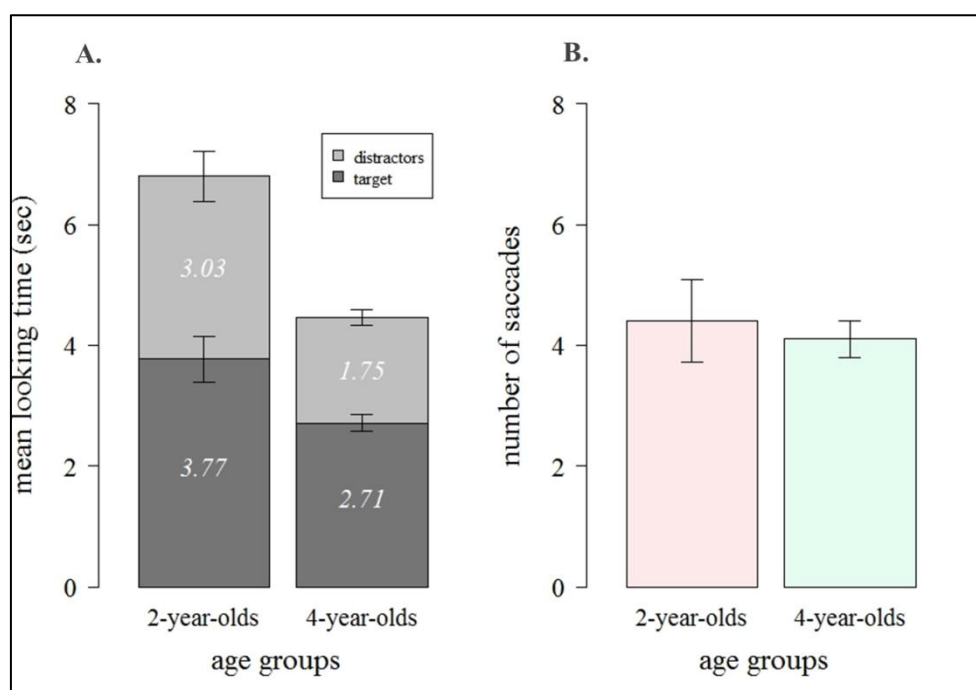


Figure 49 – A. Barplots indicating the mean looking time (in sec, \pm SEM) 2- and 4-year-old participants spent at looking at the distractors and at the targets during the “referent selection” trials. B. Barplot representing the mean (\pm SEM) number of visual saccades executed by the two age groups during the “referent selection” trials.

Consequently, the difference in RT between the two age groups stands essentially in the fact that 2-year-olds spent on average more time attending at an image before switching their eye-gaze irrespective of the image type (target or distractor), although the highest difference between the two groups arose with the looking time toward the distractive images.

To go further, in both age groups, we looked for significant correlation between the mean looking time at the target during learning and the performance during testing. Intriguingly, the less time the 4-year-olds spent looking at the targets during learning, the better they seemed to have remembered the names of these objects, as attested by the multiple R-squared and the near significant p-value (2-year-olds: $F(1,3)=0.14$, $R^2=0.05$, $p=0.7$; 4-year-olds: $F(1,3)=8.05$, $R^2=0.72$, $p=0.06$). Similarly, we did not find any correlation between the mean looking time at the distractors during learning and the performance during testing, or between the number of visual saccades made during learning and memory performance.

Taken together, these findings demonstrate that **the amount of time participants spent looking at the novel object during the selection phase does not influence the later retention** of the name of that object. Hence, it can be strongly suggested that **the difference in retention between 2- and 4-year-olds does not arise from the differences in their looking behaviors during learning**.

Adults: the effect of age

In children, none of the various variables that could have influenced the storage of the names of the objects (e.g. living with older siblings, being exposed to another language, etc.) appeared to produce a significant effect. However, we found a significant effect of age in adults. Specifically, older adults appear significantly worse at remembering the names of objects acquired through an exclusion process ($F(1,19)=8$, $R^2=0.30$, $p=0.01^*$) [Figure 49].

In adults, we also found a significant effect of the socio-professional status because *students* were significantly better at remembering the object names than *workers* ($M_{students}=96.9$, $SD_{students}=5.1$; $M_{workers}=64.4$, $SD_{workers}=35.4$; $t(9)=2.87$, $p=0.017^*$). Even though we took a special care in recruiting several young workers and some old students, one could argue that these variables are confounding variables since the students were on average younger than workers ($M(age)_{students}=23y$; $SD(age)_{students}=3.44$; $M(age)_{workers}=46.3y$; $SD(age)_{workers}=11.43$). Nonetheless, no significant interaction could be established between both *age* and *socio-professional status* variables (ANCOVA, $F(3,17)=2.9$, $R^2=0.34$, $p_{[socio-prof-status]:[age]}=0.7$).

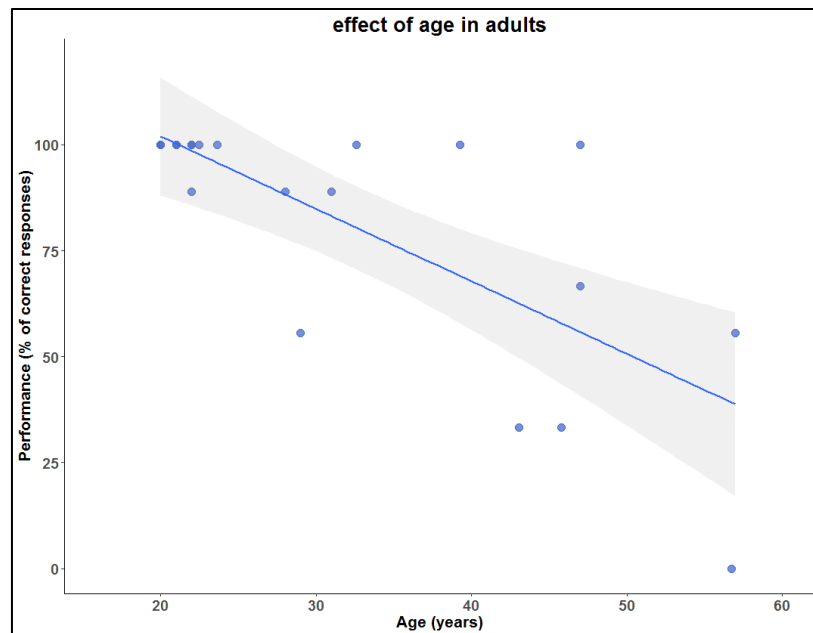


Figure 50 - Graph showing the mean performance of each adult participant during testing according to his/her age (in years). Each participant is depicted in a blue filled circle. The line represents the linear regression and the transparent curve indicates the 95% confident interval around mean.

Finally, we did not find any effect of the neuropsychological scores of adult participants on their performance during testing.

5.4 Discussion

This study explored the effect of another common word learning strategy on the storage of novel object names during early childhood and adulthood.

Referent selection

Taken together, our results demonstrate that **from 2-years of age**, children are capable of **using an inferential strategy to fast-map a novel word to the unique unfamiliar object of a scene**, which is consistent with the findings of the literature (e.g. Bion et al., 2013; Horst & Samuelson, 2008). Moreover, here we showed that **4-year-olds succeeded on the referential selection trials** in a way that was **perfectly comparable to that of adults**, both in terms of **accuracy and speed**. In contrast, 2-year-olds were on average slower to disambiguate the contexts. As an explanation of this result, we evidenced that 2-year-olds spent significantly more time looking at the novel object but also substantially more time looking at the familiar items compared to 4-year-olds. A perspective to this work would be to continue the recruitment of pre-vocabulary spurt children (i.e. 18-month-olds and/or 2-year-olds with low verbal production rates) in order to specify the exact role of the language level on solving inferential word-related paradigms. This would either

support or refute Bion's assumption that the emergence of the disambiguation ability coincides with the vocabulary spurt phenomenon (Bion et al., 2013).

Referent retention

We also wanted to know if being such actively involved in the learning procedure would result in an actual storage of the information. In the current study, we showed that 2-year-olds failed to recognize the fast-mapped associations after a break of 30-min, which is in accordance with Horst and Samuelson's pioneering work in 2-year-olds after a minimal delay (Horst & Samuelson, 2008). This finding contradicts the claim that at that age, "fast-mapping" allows a child to rapidly acquire new words and increase his/her lexical repertoire (Spiegel & Halberda, 2011). Despite this, we did find that when **2-year-old** participants **fast-mapped at least five times** the same word to its associated referent during a learning session, **actual retention** was observed after a 30-min delay. This finding is in line with a previous word learning model, claiming that the referential ambiguity is a situation-time problem that is solved within the context of a single inferential event but **building long lasting linkages** would only be solved over developmental time as children encounter **multiple presentations of the elements to be mapped** (McMurray et al 2012). Here, we provide the first evidence that 2-year-olds indeed required multiple inferential events to form and maintain in memory the associational link between a word and its object.

On the other hand, we demonstrated that this learning method was very efficient in 4-year-olds. Indeed, we found that they answered as fast as adults during testing and were almost as accurate as them to remember the associations after a 30-min delay, even after a single "referent selection" event. This study also enabled us to make direct comparisons between age groups with a strong age difference despite using exactly the same protocol, something that has never been done before. Hence, we provide a robust proof that **learning via inferential reasoning is much more powerful at 4 years of age than at 2 years of age**. This finding, along with the findings of previous studies showing that 3.5- and 4-year-olds could remember the name of a fast-mapped object after a delay up to a month (although the interpretation of these results is controversial, see Chapter 3), assumes that the **developmental shift from simply disambiguating language-related situations to actual retention of that newly acquired knowledge occurs between 3 and 4 years of age** (e.g. Markson & Bloom 1997, Vlach & Sandhofer, 2012; Waxman & Booth, 2000).

Underlying mechanisms

Undeniably, a central underlying question is what accounted for the success of the 4-year-olds to remember the names of the fast-mapped objects and concurrently the lack of success of the 2-year-olds on this same task? What are the putative underlying mechanisms? Are object names learned via ostensive labeling processed differently from object names acquired via an inferential strategy? According to our opinion, three main hypotheses could explain the discrepancy.

First, the difference may be attributable to **the size of the lexical repertoire** of 2- and 4-year-olds and to their **respective past experience with language**. Here we did not find evidence that verbal production skills influenced 2-year-olds' ability to select the referents and to remember the mappings. As an explanation, most 2-year-olds are already post-vocabulary spurt children. Unfortunately, the receptive vocabulary couldn't be evaluated since no specific questionnaires for children of 2-years of age are available in the French language. Four-year-old participants completed an ELO test (i.e. Evaluation du Language Oral) but since no equivalent test exists for 2-year-olds, we could not make rigorous comparisons between these two age groups based on the language level. Nevertheless, it is reasonable to suppose that 4-year-olds are endowed with a larger lexical repertoire than 2-year-olds, for the simple reason that they acquired a larger experience with language since birth (see Fenson et al., 1994). As a consequence, and thanks to their higher expertise with language, it may have been easier for 4-year-olds to rapidly tag the familiar items as distractors and to rapidly discard them during the "referent selection" trials, because they already encountered numerous instances of objects belonging to those subordinate categories in the past. A model indeed suggested that successful fast-mapping is partially driven by a well-developed representation of the object categories prior to the "referent selections" (Mayor & Plunkett, 2010). In contrast, we found that the 2-year-olds spent significantly more time looking at the distractors, which could be explained by the fact that they required more time to identify the subordinate category to which they belong to in order to tag them as distractors. Ultimately, during the "referent selection" trials, **2-year-olds were perhaps also learning information from the familiar objects** (e.g. learning a new cup shape, a new car brand, etc.). If so, their neural system was presumably encoding and storing information about the three images and not only about the novel one. Such multiple learning could have **interfered with the consolidation process of the visual features of the novel object**. On the other hand, 4-year-olds are probably already provided with more concept-neurons (or "category-neurons") (Quiroga et al., 2005), and thus do not require a profound treatment of the visual properties of the familiar items.

Our second hypothesis is related to the **attentional resources**. We postulated that the 2-year-olds were perhaps less attentive than the 4-year-olds when learning the new information, which would

not have allowed them to encode the information as efficiently as the 4-year-olds. Along with this idea, we suggest that the **attention allocated to each object** during the “referent selection” trials may play a key role in the encoding process. Together, we predicted that if the 4-year-olds spent on average more time looking at the novel objects during learning than the 2-year-olds, they would have had more chances to encode the visual features of these objects. On the contrary, we found that the 2-year-olds were those who spent the longer time attending at the novel objects. Thus, they should have been the most likely to encode the visual properties of the targets, but conversely they were those who failed the most during testing. Moreover, the number of visual saccades was similar for the two age groups, ruling out the possibility that 2-year-olds were more distracted during learning (each time a participant looked away from the screen was counted as one saccade). In sum, there was **no evident diminished attentional resource in 2-year-old participants**, which allows us to claim that the difference in performance during testing cannot be attributed to a lack of attention during learning of these participants. Thus, we can relatively serenely put this hypothesis aside.

Consequently and lastly, we hypothesized that the impaired performance of 2-year-olds may not arise from a default in encoding but rather from a **default in the consolidation process** of the information. If so, the discrepancy in accuracy between 2- and 4-year-old children may be caused by a significant **differential in the maturational stage of the brain structures responsible for a successful “fast-mapping”** (in the sense of remembering the names of objects acquired through an inferential strategy). Remember that, in our first study using ostensive labeling as a word learning method, we speculated that the absence of difference in accuracy between 2- and 4-year-olds was presumably due to an absence of significant enhancement of maturation of the underlying brain structures. At least, this was not measurable at a behavioral level following a 30min delay.

Hence, this would mean that these two word learning methods are processed differently and do not rely on the same brain structures. Specifically, learning words via **ostensive labeling** would be supported by **brain areas that are already functional early in life** (i.e. around two years of age regarding our results), whereas learning words via an **inferential method** would be mediated by **brain areas that mature only later on** (i.e. around 4 years of age regarding our results). But which brain regions could it be? As reviewed in Chapter 3, in the overall literature there is a huge debate about whether both word learning strategies are dependent on the hippocampus or not.

In typical “fast-mapping” studies carried out on healthy adults or amnesic patients (Coutanche & Thompson-Schill, 2014; Greve et al., 2014; Sharon et al., 2011), the authors directly compared the effect of “ostensive labeling” vs “referent selection” on word retention and tried to localize the

underlying brain areas using fMRI. As reviewed previously, in those studies participants were explicitly told to remember the names of the objects taught via ostensive labeling (e.g. “remember the *mangosteen*”) whereas they were not for the associations acquired through an inferential reasoning (e.g. “is the *numbat*’s tail pointing up?”). So, in those paradigms, “ostensive labeling” is defined as an **explicit learning process**, while on the contrary, “referent selection” refers to an **implicit (or incidental) learning process**. Some of the authors found that explicit word learning was mainly hippocampus-dependent while **implicit word learning** was mostly reliant on **extrahippocampal neocortical regions** (Atir-sharon et al., 2015; Sharon et al., 2011). These neocortical regions encompassed essentially lateral and anterior temporal lobes, the inferior prefrontal cortex and anterior parahippocampal structures (Atir-sharon et al., 2015). However, other authors found contradicting results and claimed that the **hippocampus was admittedly not essential for the disambiguation task** (“referent selection”), but was **necessary for the maintenance of the arbitrary associations** (Warren & Duff, 2014; Sakhon et al. 2018).

These findings, in addition to being contradictory, **cannot be directly transposed to a child population** for the simple reason that children, whatever the paradigm, always learned the new associations implicitly. Indeed, in the context of ostensive labeling, in contrast to the paradigms developed for adults and amnesic patients, children were not instructed to remember the associations, neither told to pay a special attention on what the stimuli because of a following “test”. Consequently, the **dichotomy between explicit versus implicit word learning**, that is typically made in adult paradigms, does not apply in children setups. So, for children, what makes these two word learning strategies different? In the context of **ostensive labeling**, children **passively** observe an experimenter (or an adult in everyday life) performing actions with the object to learn while pointing at it and simultaneously labeling it. In the context of **inferential learning**, children **actively** disambiguate the situations by eliminating distractors. There, learning consists in a more active process that may result in a deeper processing of the information. Indeed, the benefits of “depth of processing” on retention of new information have been demonstrated and largely studied in human memory (Craik & Tulving, 1975). Moreover, an active disambiguation of the information appeals to **executive functions** (see Robinson & Pascalis, 2004 for a discussion). Given that executive functions are essentially mediated by prefrontal regions such as the **dorsolateral prefrontal cortex**, which are the latest parts of the brain to mature, it can be assumed that only 4-year-olds remembered the fast-mapped associations because their prefrontal cortices are significantly better developed (though they do not reach adult levels of maturation) than those of 2-year-olds.

Nonetheless, this assumption does not exclude a **probable involvement of the hippocampus** in the consolidation processes of the associations following both learning methods. As reviewed previously, the hippocampus is known to play **a critical role in binding and linking facts and details together** (Eichenbaum et al., 1996; McClelland, 2013; McClelland et al., 1995; Squire & Zola-morgan, 1991). However, the slow and prolonged maturation of some components of the hippocampal system (see Chapter 1), especially the dentate gyrus, prevents the hippocampus from being fully functional early in life (for example, consider the phenomenon of infantile amnesia). Consequently, it can be proposed that the immature hippocampal system that is in place at 2 years of age would be **sufficient to support the retention of associations acquired via ostensive labeling but not via an inferential method**. The **maturational processes that occur between 2 and 4 years of life** (e.g. increased connectivity within the hippocampal structure and in relation with extra-hippocampal structures, diminished pruning events, etc. (Jabès & Nelson, 2015)) may confer on the hippocampus a **maturational status that is sufficient to promote the retention of fast-mapped words around 3 or 4-years of age**. In other words, only a more mature hippocampal system (though adult levels of maturation are not reached at 3-4 years of age) may be required to support long-term retention of associations acquired via a disambiguation process. This would fit with Warren & Duff's claim that the **hippocampus is not essential for the online disambiguation process (since 2 year olds are able to disambiguate the associations) but required for the maintenance of the knowledge in memory** (Warren & Duff, 2014). Another argument in favor of this hypothesis is that the age of 4 years is generally considered as the offset of the infantile amnesia phenomenon, thus the possibility for children to retain information for their entire lives. This suggests that most hippocampal functions are already operative from that age and should be able to support retention of information acquired by an inferential process.

Finally, we cannot either rule out the alternative option that **extrahippocampal structures** were responsible for the **binding and consolidation processes of fast-mapped words**, as reported by Sharon et al. in healthy adults and few amnesic patients (see Chapter 3)(Sharon et al., 2011). Consequently, not only the hippocampus, but also the **anterior temporal lobe** could be involved in linking associative knowledge together. It might be that word learning through inferential reasoning was rather supported by these extrahippocampal structures that may reach a satisfactory level of maturation around 4 years of age but not at 2 years of age.

Moreover, word learning is by definition semantic. As reviewed and discussed previously, there is still uncertainty about the neural correlates of semantic information. The dominant view is that semantic knowledge would be less dependent on the hippocampus but would rather recruit the cortices surrounding the hippocampus (perirhinal and entorhinal cortices), as well as neocortical

structures (temporal and temporo-parietal regions. It can also be speculated that the hippocampus is not involved at all in the whole encoding and consolidation process of word-object pairs. According to this viewpoint, word learning would occur outside of the hippocampus and would mainly depend on the perirhinal and entorhinal cortices along with neocortical regions, structures that also display progressive maturational stages across development.

Obviously, the only way to reconcile these various hypotheses about the neural correlates of word-object pairs acquired through an inferential process across development would be to use neuro-imagery techniques, such as fMRI. Functional MRI studies require an awake state of the participants and for the reasons reported earlier are rarely conducted on young children. To date, no fMRI study has investigated the brain regions involved in the consolidation process of fast-mapped words in young children.

5.5 Conclusion

This study was firstly designed to investigate whether 18-month-olds, 2-year-olds, 4-year-olds and adults would be capable of inferring an unknown word to its referent by logically excluding known items, using our touch-screen apparatus. Except for the 18-month-old group (work in progress), children as young as 2 years of age successfully solved the task, confirming our hypothesis that the 2-year-old participants who underwent our first study may have used this strategy for correctly identifying the word-object pair to which they were exposed only once.

The main contribution of this study is that it allowed us to directly compare the ability of 2 and 4-year-olds and adults to retain the information after a delay, since they were all exposed to the same material to learn (plus same environment, same experimenter, same apparatus). We showed that 4-year-olds and adults outperformed 2-year-olds on the retention test, and demonstrated that this learning strategy was not efficient at age 2 years to durably remember the associations except when the children were facing the disambiguation problems many times (at least 5 times). We provided three potential explanations for this developmental shift in retention of fast-mapped words between 2 and 4 years of age. Among them, children's background with language and the maturational stage of the underlying brain structures are the most likely factors.

CHAPTER V.

ONGOING EXPERIMENTS...

1 FOURTH EXPERIMENT. *Ostensive labeling versus “fast-mapping”: which strategy is the most efficient to promote word retention during childhood?*

1.1 Introduction

Combining the results of our first and third studies, we demonstrated that two completely different learning strategies were equivalently effective to induce a successful retention of novel word-object pairs in 4-year-olds and adults, whereas only one of the two triggered significant retention in 2-year-olds. This latter consisted of passively watching short videos displaying isolated novel objects that were manipulated by an experimenter’s arm and subsequently labeled three times in a row in an ecologically valid sentence. This learning paradigm, defined as *ostensive labeling*, mimics in a certain way, the everyday experiences of a young child observing an adult showing him a novel object and labeling it. Ostensive labeling is one of the three main word learning strategies described in the literature (see Chapter 3) and corresponds to the most unambiguous learning situation a child can encounter in his daily life. On the other hand, the other word learning strategy we focused on in this thesis corresponds to the ability to disambiguate language-related situations by logically excluding known objects when hearing a completely novel word. In this *referent selection* learning strategy, the child is required to actively solve the disambiguation problem by fast-mapping the unknown word to the sole unknown object of the scene. This learning method did not induce retention in 2-year-olds after a 30min delay.

One of the objectives of this thesis is to highlight the learning principles and optimal conditions required to induce a long-lasting memory of complex sensory stimuli such as word-object pairs during development. For example, we would like to know which of these two central word learning methods is the most efficient to promote a durable retention of the names of novel objects in young children and whether there is a change in efficiency of those methods during early childhood. At first glance, it would be tempting to simply compare the effects on word retention

that we reported in our first and third studies, and to conclude that an ostensive naming method overcomes an inferential learning at 2-years of age and that this pattern of efficiency evolves during the 3rd or 4th year of life. However, are we allowed to do so? The answer is clearly no, for few reasons that we attempted to summarize in the following figure [Figure 51].

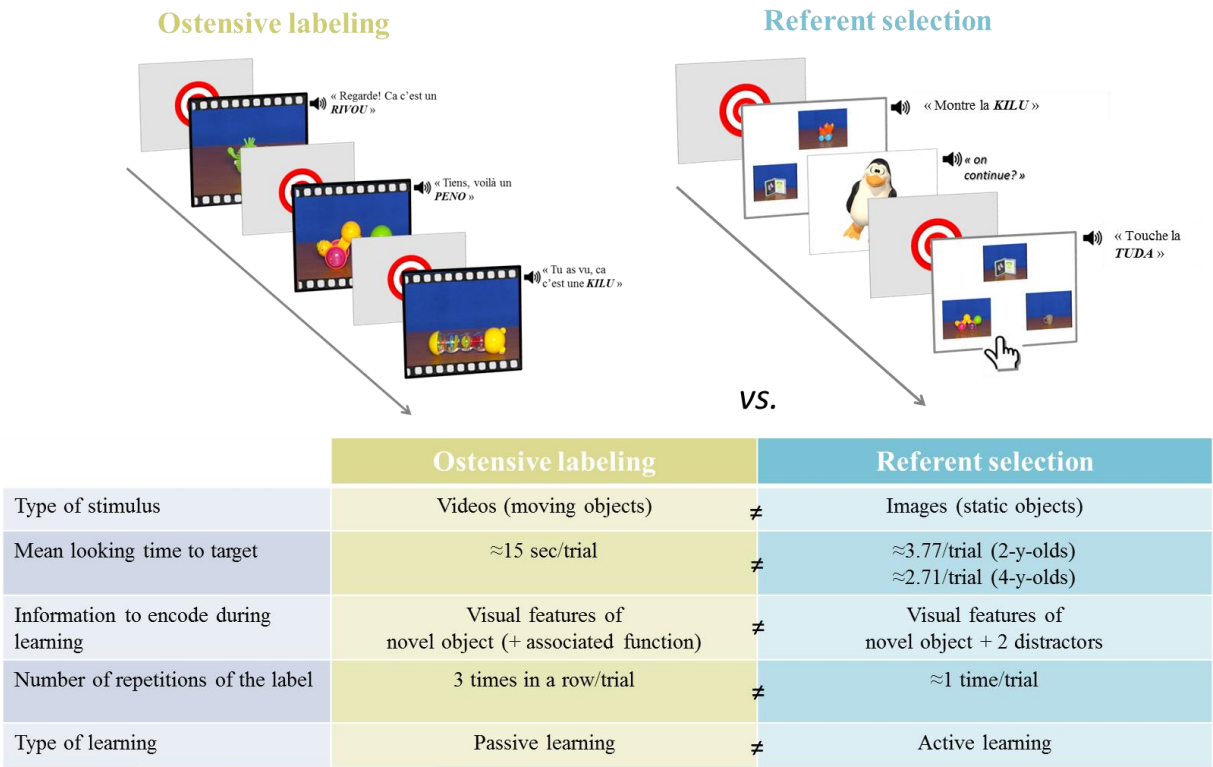


Figure 51 – Graphical summarizing the main experimental differences of the two learning methods utilized in this thesis.

The apparent higher efficiency of the *ostensive labeling* method (1st study) over the *referent selection* method (3rd study) reported in 2-year-olds may simply come result major differences in the experimental setups. As summarized in **Figure 51**, children who underwent the ostensive labeling experiment had access to causal information about the novel objects (i.e. functionality of the objects provided in the video clips) which is known to enhance learning (Booth, 2009a; Booth, McGregor, & Rohlfing, 2008; Markson & Bloom, 1997a). Moreover, the time participants spent looking at each novel object was 15sec in the ostensive labeling condition (or a little bit less if participants were distracted) against 3.77sec on average for 2-year-olds and 2.71sec on average for 4-year-olds in the “referent selection” paradigm, as measured previously (see complementary results of the 3rd experiment). Consequently, children who underwent the ostensive labeling paradigm had on average four to five times longer to encode the visual properties of the novel objects compared to children who participated in the “referent selection” experiment. Moreover, children from the “referent selection” paradigm were not only integrating information about the novel objects but also about the familiar objects. Furthermore, they had to mentally and actively

discard them to solve the task, which appeals to executive functions that were not solicited in the ostensive labeling trials. Lastly, the difference in performance may simply be attributed to the fact that the novel objects were labeled three times in the row during the ostensive labeling trials against just once in the referent selection trials (though the instruction was repeated after a 6sec delay if no answer was provided).

Taken together, all these parameters may have contributed for the significant performance of 2-year-olds in the ostensive labeling paradigm and concurrently for their lack of success in the “referent selection” paradigm. If so, one could argue that the optimal learning conditions were simply gathered in the first experiment but not in the third one. Given these numerous experimental differences, it would not be fair to conclude on the difference in accuracy just on the basis of the learning strategies themselves.

This is why we developed a fourth study. In the current and ongoing experiment, we attempted to standardize our ostensive labeling design on the basis of the “referent selection” setup, so that ultimately the two learning conditions are only differentiated by: an *unambiguous passive* learning versus an *ambiguous active* learning. The current protocol should allow us to objectively compare the effect of these two learning strategies on the memory formation of object names. Furthermore, it should provide answers to the following questions:

Did 2-year-olds from our first study take advantage of the causal information about the objects provided in the video clips? Did they take advantage of the three successive repetitions of the labels and of the longer duration of the learning events to successfully encode the sensory inputs?

In this more stringent ostensive labeling paradigm, if 2-year-olds will fail to remember the names of the objects, this will indubitably mean that participants from our first ostensive labeling study took advantage of the parameters mentioned above to form their memories. On the contrary, if the 2-year-olds do succeed in remembering the associations, it will allow us to argue that the difference in accuracy between our 1st and 3rd studies was not due to differences in the experimental setups but well and truly to the learning strategies themselves (ambiguous active learning vs. unambiguous passive learning).

So, to standardize the settings, novel objects will not be presented through video clips anymore but will consist of the same static images used in the “referent selection” experiment (images of the same size). In this way, participants will not have access to additional causal information about the novel objects. Furthermore, novel objects will be displayed against two scrambled images of familiar objects (see Atir-sharon et al., 2015 for similar methodology). In this way, the triangular configuration of the learning trials implemented in the “referent selection” paradigm

will be kept, with the distinction that only the image of the unknown object will be meaningful. In this way, the non-ambiguity aspect that characterizes the ostensive labeling method will be maintained. Moreover, we will adopt the mean looking time to target measured in the previous “referent selection” experiment to the current protocol (plus an estimated percentage of looking time to the scrambled images, see following methods) and the label will be pronounced only once (rather than 3 times in a row). Finally, as in the “referent selection” paradigm, participants will be taught three novel word-object pairs, one of which will be presented once, another three times and the last five times, in an interspersed fashion.

This study is in progress and so only preliminary data will be shown.

1.2 Methods

1.2.1 *Participants*

Participants consist of 18 twenty-four-month-olds and 8 four-year-olds. Data from 4 twenty-four-month-olds could not be included in the analyses due to fussiness ($n=3$), systematic response bias (i.e. selection of the same object on all test trials, $n=1$). Note that this experiment is ongoing and that we need to recruit more participants in each age group. We also intend to include adult participants. Details about the current sample are represented in [Table 5](#).

1.2.2 *Stimuli*

Stimuli consisted of images of three unknown objects associated with their randomly chosen pseudo-words and three scrambled images of familiar objects (i.e. distractors). The novel objects were chosen prior to the experiment in accordance with the assessment of the caregiver about the novelty aspect of the stimuli.

The scrambled images were generated with a self-developed program under Python software which consists of shuffling the pixels (definition: 720x576) of the original image of a familiar object, so that the resulting image contains the luminance and color as the original one, but with the constraint that it is meaningless [[Figure 52](#)]. Ultimately, the same amount of information is conveyed to the retina of the participants but no profound treatment is allowed.

Three familiar objects were also selected prior to the experiment and served for the warm-up trials.

1.2.3 Procedure

The learning procedure implements the one used in the previous “referent selection” study, except that participants are not required to select an image by excluding images of familiar items. Participants launch each trial by pressing a red button situated on the center of the screen [Figure 52]. Then, they are simply asked to passively look at static images displayed on the screen (the image of a novel object is displayed against two scrambled images) while a tape-recorded voice is labeling the novel object in an ecologically valid sentence [Figure 52]. The scrambled images are meaningless (i.e. distractors) and only the image of the novel object can make sense.

The duration of each trial corresponds to the mean time participants of the previous “referent selection” study spent looking at the novel object, plus half of the mean time they spent looking at the familiar items. We presumed that since the distractive images are currently meaningless, participants should spend almost half the time looking at them than in the previous experiment that involved meaningful images of familiar objects. Consequently, here the duration of each trial is 5.27sec ($3.77\text{sec} + (3.03/2) \text{ sec}$) for the 2-year-old participants and to 3.58sec ($2.71\text{sec} + (1.75/2) \text{ sec}$) for 4-year-olds.

Exactly 1500ms (2-year-olds) or 1000ms (4-year-olds and adults) after trial onset, a recorded voice labels the novel object using a child-directed carrier phrase randomly chosen among three options (e.g. “regarde ca c’est un/une [label]”; “tu as vu, ca c’est un/une [label]”; “tiens, voilà un/une [label]”). The label is pronounced only once. At the end of each trial, the same neutral GIF animation as the one used in the former experiment, encourages participants to continue with another trial [Figure 52].

In total, participants undergo 9 trials since the objects to learn are presented either 1, 3 or 5 times in a pseudo-random order.

As in the previous studies, the experiment begins with three warm-up trials involving familiar objects. During this familiarization phase, each familiar object (chosen with the caregiver prior to the experiment) is displayed on the screen against two scrambled images. Each familiar object appears only once on the screen.

After learning, the children played in the experimental room for a 30-min period. For the same time period, adult participants complete the neuro-psychological tests and are asked to fill in a survey about their general opinion on the use of touch-screens in preschools.

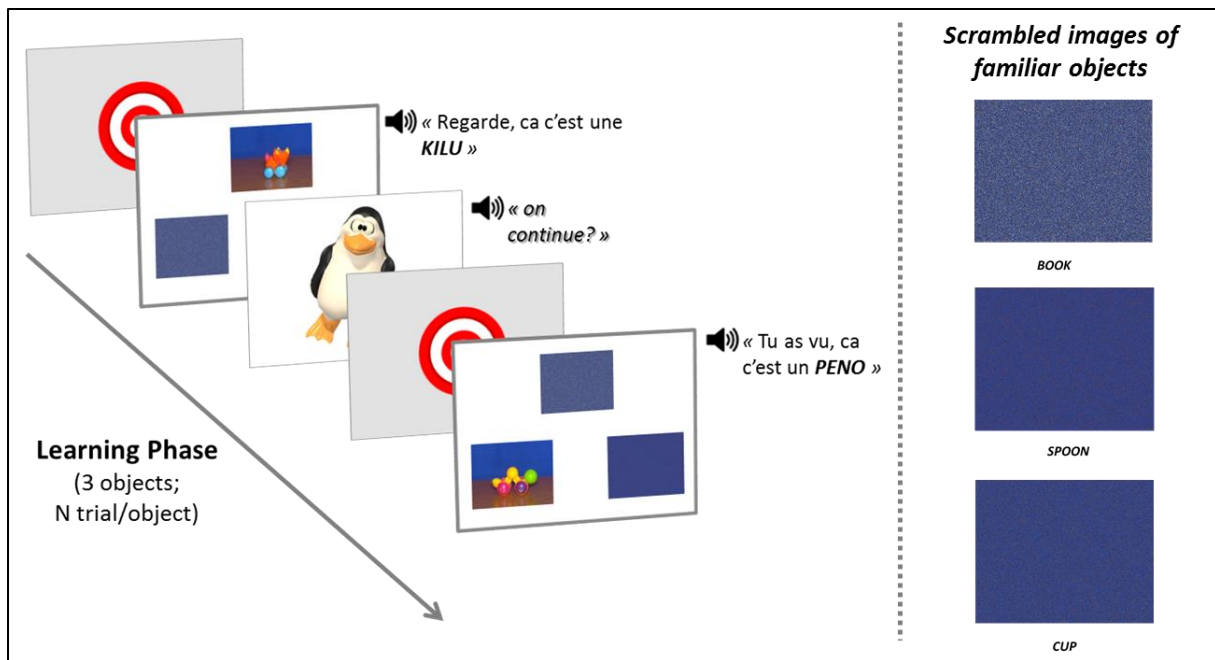


Figure 52 - Illustration of the learning procedure (left). Participants initiate themselves each trial by pressing a red button on the center of the touch screen. Participants are then presented with 2 scrambled images of familiar objects (randomly chosen among three, see enlarged examples on the right) and one novel object. A tape-recorded voice labels the object using an ecologically valid sentence (e.g. “regarde, ca c’est une [label]). The label is uttered once per trial. Trial duration is 5.27sec for 2-year-old participants and 3.58sec for 4-year-olds and adult participants. One of the three novel objects appears only once, another one on 3 trials and the last one on 5 trials, interspersed. A 3s GIF animation showing a moving penguin accompanied by a stimulating sentence encourages participants to continue.

Testing consists of 3AFC trials involving the three novel objects, exactly as performed in the former experiments. Novel objects serve as target three times in an interspersed manner (testing session = 9 trials) and any answer can trigger the following trial.

1.3 Results

Preliminary results indicate that the 2-year-olds performed significantly above chance ($M=53.36\%$, $SD=21.13$; exact binomial $p=9.87e^{-6***}$) with a very large effect size (Cohen’s $d=0.88$) [Table 5 & Figure 53]. The 4-year-olds also remembered the associations significantly above the level expected by chance ($M=77.7\%$, $SD=20.57$, exact binomial $p=8.56e^{-15***}$) with a huge effect size (Cohen’s $d=2.16$). Moreover, we found that 4-year-olds outperformed 2-year-olds (GLMM, $Z=2.7$, $p=0.0068^{**}$).

Next, we compared the preliminary results of the two age groups with the performance obtained by the participants of the same ages from our 1st and 3rd studies.

Table 5 - Details about the composition and results of each age group.

Age groups	24-Mo	4-Y	Adults
N subjects	14	8	
N females	8	4	
Mean age (\pmSD)	24.70 Mo (\pm 0.72)	3.9 Y (\pm 0.26)	
N subjects with older siblings at home	4	3	-
N subjects exposed to another language at home	3	3	
IFDC voc in production (%) (\pmSD)	76 (\pm 18.02)	-	-
Mean overall accuracy (% correct trials) (\pmSD)	53.36 (\pm 21.13)	77.7 (\pm 20.57)	
P (exact binomial tests)	9.87e-6	8.56e-15	
Cohen's D	0.88	2.16	
Mean RT (sec) (\pmSD)	6.13 (\pm 4.15)	2.45 (\pm 0.72)	
Mean Accuracy 1 REP (% correct trials) [95% CI]	45.24 [30-61]	83.33 [63-95]	
P 1 REP (binomial tests)	0.1	5.5e-7	
Mean Accuracy 3 REP (% correct trials) [95% CI]	51.78 [35-67]	75 [53-90]	
P 3 REP (binomial tests)	0.01	3.08e-5	
Mean Accuracy 5 REP (% correct trials) [95% CI]	63.09 [47-77]	75 [53-90]	
P 5 REP (binomial tests)	8.06e-5	3.08e-5	
Mean RT 1 REP (sec) (\pmSD)	6.39 (\pm 4.81)	2.53 (\pm 0.60)	
Mean RT 3 REP (sec) (\pmSD)	5.61(\pm 3.19)	2.56 (\pm 0.68)	
Mean RT 5 REP (sec) (\pmSD)	6.41(\pm 4.54)	2.27 (\pm 0.89)	

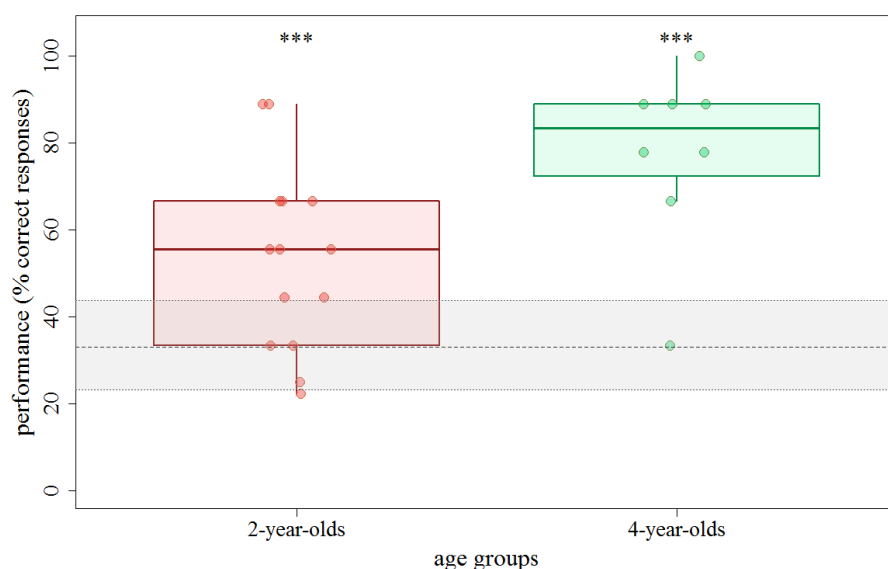


Figure 53 - Boxplots representing the performance of the 2- and 4-year-olds during the 3AFC retention task. Boxplots show the median (full-line). Dashed lines represent the 33% chance level (dark grey) and the 95% confidence interval around chance (light grey). Each individual mean performance is depicted in a circle.

2-Year-olds: Interestingly, no significant differences in accuracy were established between the 2-year-old participants who underwent this experiment and those who underwent the previous ostensive labeling condition (GLMM, $Z=0.77$, $p=0.44$) [Figure 54]. This result indicates that 2-year-old participants performed as well in the previous ostensive labeling study as in the current one. Moreover, participants from both ostensive labeling methods exhibited significantly higher scores of accuracy than those who underwent the “referent selection” study (GLMM, $Z_{\text{ost lab static-ref selec}}=-1.95$, $p=0.05^*$; $Z_{\text{ost lab videos-ref selec}}=-3$, $p=0.003^{**}$).). Note that the composition of the 2-year-old samples is highly similar across studies regarding the following variables: age, gender, proportion of subjects living with older siblings at home, proportion of subjects exposed to another language and IFDC score [Table 1, Table 3, Table 4 & Table 5].

4-year-olds: there was no significant difference in accuracy between the 4-year-olds who underwent the previous ostensive labeling experiment and the 4-year-olds who underwent the current experiment (GLMM, $Z=1.17$, $p=0.26$). Furthermore, 4-year-olds are similarly accurate during testing whatever the learning method (GLMM, $Z_{\text{ost lab static-ref selec}}=-1.4$, $p=0.16$; $Z_{\text{ost lab videos-ref selec}}=-0.36$, $p=0.7$) [Figure 54].

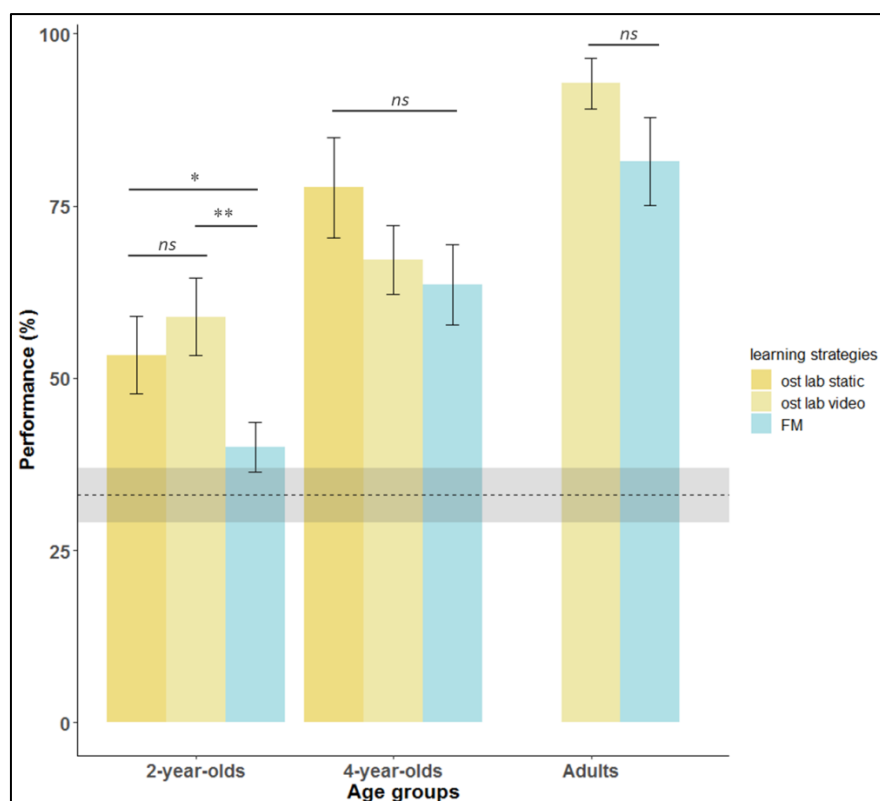


Figure 54 - Barplots summarizing the mean performance (% , \pm SEM) during testing of the three age groups according to different learning strategies (“**ost lab static**”, i.e. ostensive labeling with static images of the objects = current study ; “**ost lab video**”, i.e. ostensive labeling through short videos = study N°1 and “**FM**”, i.e. “fast-mapping” = study N°3). The dashed lines represent the 33% chance level (dark grey) and the 95% confidence interval around chance (light grey). Stars indicate the level of significance (*, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$).

Finally, we did not statistically analyze the mean performance of each age group according to the number of presentations during learning, and nor did we examine in detail the RT during testing, since the sample sizes are still too small, especially for the 4-year-old group. However, for the 2-year-old group, it can be noticed that the mean performance increased with the number of presentations during learning [Table 5] in a very similar way than the increase reported in study N°1 [Table 5]. This suggests that participants similarly benefitted from the repeated presentations of the pairs, and on the whole exhibited the same level of accuracy whatever the ostensive labeling method employed.

1.4 Discussion

This study aimed to know whether the difference in performance between our 1st (ostensive labeling) and 3rd study (“fast-mapping”) may not simply be due to high differences in the experimental settings. Specifically, the fact that 2-year-olds seemed to have outperformed in the former ostensive labeling experiment (1st study) in comparison to the “referent selection” experiment (3rd study) may simply be explained by an enhanced access to the material to learn in the 1st study (e.g. longer exposition to the novel objects, three repetitions in a row of the labels, causal information available, manipulation of the objects, etc.). Here, we wondered whether 2-year-olds would be equivalently accurate when removing all these putative learning enhancers. Furthermore, by standardizing the learning methods, we are now allowed to objectively compare the efficiency of the two learning strategies on the emergence of a memory trace.

To those aims, in this 4th study we applied the exact same parameters than those employed in the “referent selection” study (i.e. stimuli were static images, same size of the images, same amount of time to visualize the novel objects, same number of verbal repeats (1/trial), same neutral animation between trials, etc.). Nonetheless, what makes the main difference between the current study and the former “referent selection” one, is the learning method per se. Here, participants passively looked at the images of the novel objects being unambiguously displayed on the screen among two meaningless images and listened to the labels. In the previous “referent selection” experiment, participants had to actively disambiguate ambiguous trials by excluding the images of two familiar objects when hearing an unknown name.

Did 2-year-olds similarly learn from the two ostensive labeling learning methods?

Although this experiment is not fully completed, our preliminary data strongly suggest that 2-year-olds did not take advantage of the experimental procedure settled in our previous ostensive labeling study. For example, it does not seem that a video clip is more efficient for memorizing

the name of an object than a still image of that object. This is in accordance with a previous study showing equivalent results between children who were taught labels for moving images of objects and those who were taught labels for still images (Houston-Price et al., 2005). It also suggests that the three verbal repeats provided in the former ostensive labeling were not the source of success.

Consequently, it can be reasonably claimed that **ostensive labeling, whatever the labeling manner, is a very powerful learning method that can allow the establishment of a durable memory trace between an object and its associated name in 2-year-olds**. This finding is however surprising in regard to other studies from the literature that emphasized the importance of providing functional or causal information about the novel object, as well as non-verbal gestural cues (e.g. gazing at or pointing to the object, manipulation of the object, etc.), for enhancing the memorization of the name of that object (Booth, 2009; Booth, McGregor, et al., 2008; Markson & Bloom, 1997). These authors reported that learning improved with greater redundancy among cues, suggesting a strong role for socio-pragmatic factors on word learning in children from 28 to 31 months of age (Booth, McGregor, et al., 2008). Our current preliminary findings appear to contradict these results and instead suggest that a brief visual exposure to an object unambiguously accompanied with its label is sufficient to enable retention of the pair. Now, it would be interesting to examine whether the two pairs that were presented several times (3 and 5 repetitions) were important for the results or whether children of 2-years of age would also be able to remember pairs displayed and labeled only once without any external support (i.e. a replication of our study N°2 with this ostensive labeling design).

2-year-olds: why do they fail to retain fast-mapped associations?

Taken together, results from our first and current experiment show that 2-year-olds can succeed in learning the names of novel objects that were unambiguously taught, even when the learning conditions were more rigorous. Conversely, results from our third study demonstrated that 2-year-olds failed to remember the names of “fast-mapped” associations that followed the same rigorous conditions. Consequently, it can be argued that **it is precisely the disambiguation act that prevented 2-year-olds from remembering the associations after a delay**. Likewise, the difference in performance may be explained by the **presence of distractors during learning**. A study revealed that the number of competitors (2, 3 and 4) did not affect 30-month-olds’ ability to form the initial word-object mappings, but only children who encountered the lowest number of competitors (i.e. 2) during learning, demonstrated significant levels of retention after a 5-min break (Horst et al., 2010). In line with this idea, in another study, authors showed that highlighting the target object and/or reducing the salience of the distractor objects promoted retention (Axelsson et al., 2012). Accordingly, here, our preliminary results suggest that 2-year-olds

remember new words even better if there are zero (or only meaningless) competitors during learning. Altogether, these findings indicate that at that age, **word learning is maximized when ambiguity and distractions are minimized or absent**. In other words, a **passive unambiguous word learning method could be more efficient than being actively engaged in learning**.

What about the 4 years old?

Four-year-old participants performed well whatever the learning method. At that age, having a direct and unambiguous access to the meaning of a novel word or being actively engaged in mapping the words to their corresponding items by using inferential strategies triggered the same outcome on memory.

This finding goes against the unique study in the literature that compared the memory of 3.5-year-olds for words learned via inference or instruction (i.e. ostensive labeling) (Zosh et al., 2013). In that study, participants were taught six novel word-object pairs either in an inferential context (e.g. “point at the [label]”; note that here the target was pitted against a unique distractor) or in an instructional context (e.g. “this is a [label]”; no other object was present). In both cases, objects were images displayed on a screen for about 7sec and appeared only once during learning. Testing occurred immediately after learning and involved real 3-D versions of four out of the six novel pairings (i.e. 4 AFC task). Only one object among the four was queried during testing (i.e. testing = 1 trial). Children’s comprehension of the associations was measured by pointing to the target object after label onset. The authors found that the children who underwent the inferential learning context chose the target object significantly above chance, whereas those who underwent the instructional learning context performed at chance levels. This observed difference in performance was not attributed to a lack of attention during the learning trials since children looked longer to the novel object when it was presented in an instruction context (i.e. same analyzes as we did in the previous study). In a within-subjects replication of the design, 3.5 years old were taught 3 pairings in each context and exhibited similar results (Zosh et al., 2013, experiment 2). In contrast to our results, the authors found that children failed to retain novel object names learned via unambiguous naming and that only a disambiguation context favored word retention. Consequently, the authors concluded that the presence of a single distracter object may be more beneficial for 3- to 4-year-old word learners than with no competitors or with too many of them (Horst et al., 2010). They also pointed out a possible change of strategy across development.

The discrepancy between our finding (though preliminary) and theirs could arise from four major differences in the experimental design:

- (i) The delay between learning and testing: we tested our participants after a 30-min break, whereas Zosh et al tested their participants immediately after learning.
- (ii) The number of test trials: in our study, participants underwent 9 test trials and were tested on 3 word-object associations (3 test trials/association) against a unique test trial in Zosh et al's study (only one association was tested). Obviously our testing procedure was more demanding. In their paper, the authors indicated that pilot work suggested that performance declined with repeated testing (i.e. when children were tested on more than one association). But is it really fair to simplify as much a testing protocol to get a desired result? Unlike to their pilot observations, we did not find either negative or positive effects of a repeated testing across our different protocols.
- (iii) The number of presentations during learning and the number of associations to learn: our subjects had 3 associations to learn (against 6 in Zosh et al's first between-subjects design and 3 in their second within-subjects design). Moreover, our participants were exposed respectively 1, 3 and 5 times to the associations. In the study of Zosh et al, children participants were exposed only once to each pair.
- (iv) The number of distractors in the inferential condition: we pitted two competitors against the target object against one in Zosh et al.'s study. Note that in real life, children who are facing an ambiguous context often have to deal with many more distractors than only one or two.

Sadly, Zosh et al did not replicate their experiment on 2-year-olds which would have allowed them to investigate a possible developmental change in strategy if 2-year-olds would have exhibited the opposite pattern of result.

Obviously, the discrepancy between their finding and ours indicates that the effect of the different word learning strategies on memory formation for object names across development is not yet resolved and requires further investigation. Currently, we aim to complete our study by including at least 6 more 2-year-old subjects, 12 more 4-year-olds and a whole group of adults (~20 naïve participants) before drawing strong conclusions. We also need to measure the time children spent looking at the target during learning in order to assess whether it is truly equivalent to the mean time allocated to the target in the former inferential condition.

2 FIFTH EXPERIMENT. *Immediate versus delayed memory in children*

2.1 Introduction

So far, we were interested in highlighting the effect of different learning conditions and learning strategies on cross-modal memory formation during development (i.e. passive vs active word learning). We also investigated the effect of repeating the information during the learning session. We showed that overall, the more the participants encountered the sensory inputs to learn, the better there were at retaining this knowledge, irrespective of the learning method. We also showed that a single exposure to an object unambiguously labeled three times in a row was sufficient to induce a memory trace after a 30-min delay.

In this new ongoing study, we intend to foster our understanding of the factors that contribute to enhance the mnesic capabilities of children across development in the context of word learning. What could make them memorize better? Are the principles that drive the formation of long-lasting memories in human adults already operative in early-developing brain systems? In the first place, we question the importance of *time* between learning and recalling the newly learned information. For instance, it is now well established that sleep (overnight sleep or even a short nap) plays a critical role in the consolidation process of newly learned information in adults (e.g. Diekelmann, Wilhelm, & Born, 2009; Rasch & Born, 2013 for reviews). **Sleep as a memory enhancer has now also been established in children** (Backhaus, Hoeckesfeld, Born, Hohagen, & Junghanns, 2008). A recent study demonstrated that children who napped after learning exhibited enhanced retention of the words recently encountered compared to children who remained awake after learning (Axelsson, Swinton, Winiger, & Horst, 2018). In this thesis, we would have liked to implement an overnight or even a 45min nap interval between learning and testing, but for practical reasons (unavailability of the parents) this was unfortunately not feasible. We were however able to re-test the memory of subgroups of participants after longer-term delay (see study N°6 below). Although we could not implement a period of sleep prior to the testing, in the current experiment, we wonder whether the 30-min resting period during which children were engaged in other activities (thus potentially in novel learnings) was beneficial or conversely detrimental for the consolidation process of the recently learned material.

Specifically, we wanted to know whether reducing the delay between learning and testing would benefit the formation of memories. In other words, would our results have been improved if the delay between learning and testing would have been shortened? Could the discrepancy between

our finding and that of Zosh et al (discussed above) be predominantly attributed to differences in latency between learning and testing?

Intriguingly, in the literature, very few studies related to word learning in children tested memory after a delay superior to few minutes. As reviewed in Chapter 3, most studies tested children's comprehension of the meaning of the newly learned words either immediately after learning (e.g. Bion et al., 2013; Gurteen et al., 2011; Schafer & Plunkett, 1998; Zosh et al., 2013) or after a maximum of 5min delay (Axelsson et al., 2012; Axelsson & Horst, 2013; Horst & Samuelson, 2008; Horst et al., 2010; Kucker & Samuelson, 2012). Surprisingly, a study revealed that 3-year-old children, who were taught new words for unfamiliar objects or animals using causal information, performed at chance when tested immediately after training, but succeeded when that testing occurred several days after training (Booth, 2009). The authors attributed this finding to the enhancing process of sleep on consolidation. But even more surprising, in pilot work, the same authors reported successful retention after only a 10-min delay (and still performance at chance if the test occurred immediately after learning). The authors proposed that when children undergo intensive training (e.g. being taught too many novel words during a single learning session), they may be overwhelmed by fatigue and require time to refocus their attention on the task.

Given our persuasive finding that a single exposure to an object ostensibly labeled three times in a row triggered retention after a 30-min delay in children as young as 2-years of age (Study N°2), in the current study, we addressed the question of the effect of time by testing participants immediately after learning and by comparing this result to that obtained in study N°2 (i.e. after a 30-min break)

This study is in progress, and only preliminary results will be shown below.

2.2 Methods

2.2.1 *Participants*

At the moment, participants consist of 8 twenty-four-month-olds, 6 four-year-olds and 7 adults. Details about the current samples are summarized in [Table 6](#).

2.2.2 *Procedure*

This study is a replication of study N°2 with the unique distinction that testing occurred immediately after learning rather than after a 30-min break.

2.3 Results

Table 6 - Details about the composition and results of each age group.

Age groups	24-Mo	4-Y	Adults
N subjects	8	6	7
N females	2	4	5
Mean age (\pm SD)	25.3 Mo (\pm 0.72)	3.96 Y (\pm 0.30)	24.73 Y (\pm 2.95)
N subjects with older siblings at home	2	3	-
N subjects exposed to another language at home	0	3	-
IFDC voc in production (%) (\pm SD)	68 (\pm 42)	-	-
Mean overall accuracy (% correct trials) [95%CI]	38.8 [27-51]	29 [18-43]	85.7 [74-93]
<i>P</i> (exact binomial tests)	0.31	0.66	$<2.2e^{-16}$
Mean RT (sec) (\pm SD)	6.3 (\pm 4.53)	3.8 (\pm 2.03)	2.52 (\pm 1.01)

In this study, participants were ostensibly taught 3 novel word-object pairs introduced only once to them through short video clips and testing occurred immediately after learning. Preliminary results indicate that children participants of either 2 or 4 years of age poorly retained the names of the objects during the 3 AFC test trials [Table 6]. In contrast, adult subjects performed highly above chance in this condition ($M=85.7$, exact binomial $p<0.001^{***}$). Although more individuals need to be included in each age group, preliminary results suggest that only adult participants recognized the associations when testing immediately follows training.

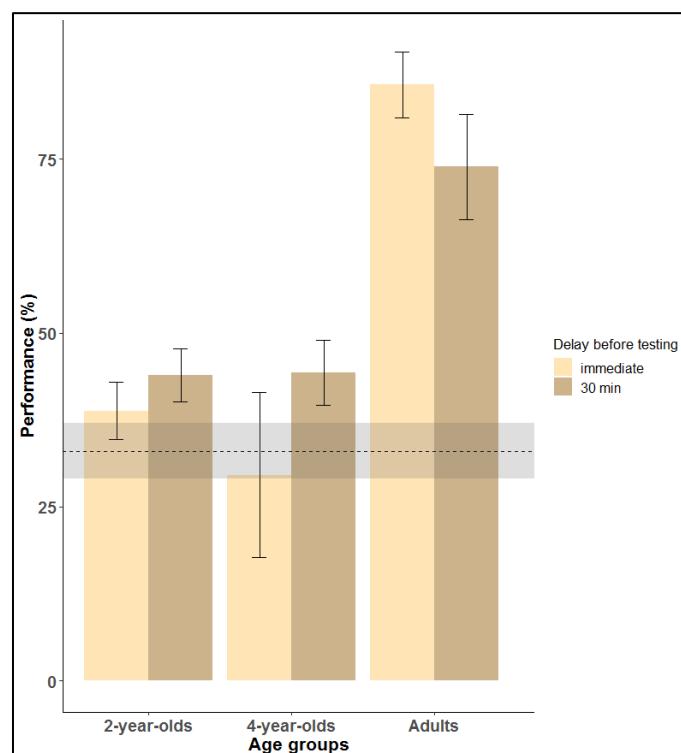


Figure 55 - Barplots summarizing the mean performance (% \pm SEM) of the three age groups to remember the names of objects seen once through a short video clip, according to the time that elapsed between learning and testing (0min = “immediate” (yellow) or 30min (brown)). The dashed line represents the 33% chance level (dark grey) and the grey shadow represents the 95% confidence interval around chance.

Next, we compared the preliminary results of this experiment to those obtained by participants of the same age who underwent a 30-min break before testing (i.e. study N°2, brown bar on [Figure 55](#)). Although there seem to be a dramatic drop in accuracy for the 4-year-olds who were tested immediately after learning compared to those who were tested after a 30-min break, no significant difference in accuracy could yet be established (GLMM, *4-year-olds*: $Z=1.8$, $p=0.07$). No significant difference could either be evidence in 2-year-olds, although they seem to exhibit the same decrease in accuracy than the 4-year-olds when no time elapsed between learning and testing (GLMM, *2-year-olds*: $Z=0.65$, $p=0.51$). Interestingly, the opposite pattern seems to govern the performance of adults (GLMM, *adults*: $Z=-0.6$, $p=0.54$) [[Figure 55](#)]. Naturally, more subjects should be included in each age group before interpreting in depth these data. Note that for now, the mean RTs in this condition are relatively similar to those measured in the 30-min break condition [[Table 3](#) & [Table 6](#)].

2.4 Discussion

This ongoing study intend to determine whether implementing a delay before testing young children on what they recently learnt is beneficial or not. To that purpose, we replicated the protocol we used in our study N°2 - namely exposing 2 and 4-year-olds and naïve adults to short video sequences displaying novel objects that are manipulated by an experimenter's arm and labeled three times in a row – but tested the subjects immediately after learning. The performance of each age group was compared to the performance of participants of the same ages who were engaged in other activities before being tested (results of our study N°2).

Interestingly, our preliminary results suggest that children fail to recognize the word-object associations if they are tested right after learning while they succeed if the test phase is delayed for about half an hour. Conversely, adults seem to perform equally well or even better if they are tested immediately after learning.

These interesting findings, though preliminary, may help understand why numerous studies failed to demonstrate word comprehension in young children who underwent test trials directly after learning (e.g. Bion et al., 2013; Horst & Samuelson, 2008, Zosh et al. 2013). It can be speculated that adding a delay before testing, would have improved the retention capabilities of those participants.

Why would immediate testing be detrimental compared to delayed testing?

First, as suggested by Booth, depending on the task complexity, children may suffer from a diminished arousal and may require time to refocus their attention on the task (Booth, 2009). Second, immediate testing does not target the same memory type than delayed testing, and the difference in performance may simply be explained by the fact that two distinct brain areas and mechanisms are recruited. **Immediate testing refers to short-term memory.** Short-term memory is defined as the ability to mentally retain information over short periods of time (from few seconds to few minutes), which is different from the ability to mentally manipulate the information (working-memory), though the distinction between short-term and working memory is still a matter of debate. The neural basis of short-term memory is still also uncertain but the governing assumption is that short-term memory (as well as working-memory) is supported by regions of the **prefrontal cortex**. As reviewed above, the prefrontal cortex shows a prolonged course of development compared to most other cortical regions. Consequently, the steady development of the prefrontal cortex coincides with the steady increase of short-term memory abilities measured at a behavioral level from the preschool years through to adolescence (Gathercole & Hitch, 1993; Gathercole, Pickering, Ambridge, & Wearing, 2004). This helps explain why our young children showed impairments on our current short-term test whereas adults, with a mature prefrontal cortex, succeeded when tested immediately after learning. It may also help explain why Zosh et al. found negative results for ostensibly labeled associations but significant ones for fast-mapped associations (Zosh et al., 2013). As discussed above, the “referent selection” process probably requires the involvement of prefrontal structures to logically disambiguate the task. It might be that because those structures were already recruited during learning, it benefitted short-term recall.

Then, if a sensory stimulus is arousing and relevant enough, the information may be consolidated and gradually transferred from short-term into long-term memory. This phenomenon occurs at many organizational levels in the brain (see Chapter 1) and involves cellular and molecular changes which typically take place within the first minutes or hours of learning. Ultimately, it results in structural and functional changes in various sets of neurons and neuronal circuits, such as the restructuring of existing synaptic connections or the growth of novel ones. Hence, **during the 30-min distractive period, the cascade of different metabolic events that constitute the consolidation process was already triggered.**

To summarize, **immediate versus delayed testing simply tap into different types of memories.** In the first case, the memorization process per say did not yet take place and children relied on their fragile maintenance abilities of recently encountered information while in the second case,

memory was well and truly tested. Remember that before 7 years of age, spontaneous rehearsal does not reliably occur, so the performance reported after a 30-min break cannot be attributed to an ability to maintain the information active for a longer period of time (Gathercole & Hitch, 1993).

Despite this, if our preliminary findings are confirmed with the inclusion of more participants, this would have a direct and influential implication for educational purposes. It would suggest that immediate tests disrupt learning and that applying a delay benefits the consolidation process of newly learned information. It would also mean that assessing word learning immediately after learning is probably not an adequate method in early childhood.

3 SIXTH EXPERIMENT. *Long-term memory of word-object pairs in children*

3.1 Introduction

In this ultimate study, we aimed to know if memory can survive longer delays during early childhood. The M4 project attempts to examine the principles and factors that promote the establishment of long-lasting sensory memories in human beings. A central goal of this project is to determine whether few exposures to a novel sensory input are enough to form a memory that can last a lifetime. Previous studies from colleagues working on that project revealed that decades after the initial exposure, adults were able to recall audio-visual stimuli (TV programs) that could not have been re-experienced in the intervening period (Christelle Larzabal et al., 2017). In another astonishing study, the same first author showed that young adults were capable of recognizing above chance level drawings that had been presented three times, two seconds each time on average, approximately 12 years earlier (Larzabal, Tramoni, Muratot, Barbeau, & Thorpe, 2018).

But do durable memories also exist in young children? Considering the phenomenon of juvenile amnesia (though mainly related to episodic memories), the possibility that children of 4-years of age could retain sensory information after a delay of several months than after a delay of 30min is uncertain. In the literature, only a handful of studies examined the question of the strength of the memory trace for complex bimodal information in 4-year-olds, especially in the context of “fast-mapping”. They all demonstrated that 4-year-olds successfully recognized the name of an object they fast-mapped 1 month ago (Carey & Bartlett, 1978; Markson & Bloom, 1997b; Vlach & Sandhofer, 2012; Waxman & Booth, 2000). However, as already discussed earlier, since they had

only one object to fast-map, it is unclear whether those children truly retained the name of the object or simply recognized which object was given a special treatment during learning (during testing, the target object was pitted against few completely novel objects).

To disambiguate this issue, in the current study, children participants were tested on three word-object associations. Specifically, we tested children from any of our previous studies whose parents have agreed to bring them back into the lab were re-tested on their knowledge for novel word-object pairs after a delay of either one month or six months. Since we used pseudo-words and pseudo-objects, the possibility that children participants were re-exposed to those stimuli during the time interval was unlikely. Parents were instructed not to talk about the experiment with their child during the entire intervening period. Moreover, prior to the re-test, we asked the parents if their child spontaneously recalled the labels or talked about what they have experienced in our lab since the last visit. If so, we did not include their data in the following results. Thus, the absence of potential re-exposure to the material was strictly controlled since we wanted to explore if a memory trace could remain dormant for few months in young children.

Long-term performance of 4-year-olds was compared to that of adult subjects who were also unexpectedly re-tested after either one month or six months. As only few participants from each learning condition came back to the lab, we pooled the data of participants, whatever the learning condition they underwent. Overall, irrespective of the learning condition, we expect adults to outperform 4-year-olds, especially after a 6-month interval. We also expect 4-year-olds to perform at chance on those long-term test trials.

3.2 Methods

3.2.1 *Participants*

During initial recruitment, caregivers could decide to bring their child either once (learning and testing on the same day) or twice (second test after either one or six months) at the lab. Those criteria were not imposed on the caregivers because we were afraid that most of them would decline to enroll their child in our study. Parents who accepted to come back for another visit were instructed not to talk about the experiment at home during the entire time interval and were explained how this criterion was important for our scientific purposes. At the beginning of the second visit, parents were asked if their child spontaneously recalled the labels at home. None of the children in the following sample did so. Adults who were tested a second time were mostly acquaintances since it would have been difficult to justify a return to the lab to strangers who already “fell into our trap” during the first memory test (i.e. they were not aware about the

memory test). If we would have told them that the request would be about a long-term memory test, they would probably have tried to develop strategies to remind themselves about the associations. Since we wanted children and adult participants to be in the closest learning and testing conditions to enable trustworthy comparisons, the long-term memory test had to be unexpected for both the children and the adults.

The current sample consists of **28 four-year-olds** and **26 adults**. Eighteen 4-year-olds were re-tested after a 1-month delay and ten were re-tested after a break of 6 months. Twenty-one adults were tested after a break of 1 month and 5 adults were re-tested after a 6-month delay. Among the 21 adult participants who were tested after a 1-month break, 9 did not perform a first memory test after a 30-min break (i.e. they performed the long-term memory test without having being tested at all). Too few 2-year-olds came back for another memory test, and so we did not include their data in the current study.

Not all 4-year-old and adult participants of the current sample underwent the same type of learning. Instead, the sample consists of a few participants from each learning condition. Specifically, the sample of 4-year-olds consists of: 10 participants who underwent our 1st study (ostensive labeling with video clips, multiple exposures/associations), 4 participants who underwent our 2nd study (ostensive labeling with video clips, 1 exposure/associations), 9 participants who underwent our 3rd study (“Fast-mapping”, i.e. FM, multiple exposures/associations) and 5 participants who underwent our 4th study (ostensive labeling with static images, multiple exposures/associations) [Table 7].

The adult sample consists of: 11 participants who underwent our 1st study, 4 participants who underwent our 2nd study and 11 participants who underwent our 3rd study [Table 7].

3.2.2 Procedure

Participants were tested on the associations taught either 1 month or 6 months earlier. They had not been re-exposed to the objects or to the pseudo-words prior to the test. No warm-up trials preceded the test. The memory test implemented the same procedure than the test performed after a 30min delay, i.e. 3 test trials/associations (9 trials).

3.2.3 Analysis

After each latency (1 month and 6 months), the performance of each age group was compared to levels expected by chance (i.e. 33%). Generalized linear mixed-models (GLMM) were used to assess whether differences in performance could be established between the different latencies and between the two age groups.

Given the low number of participants arising from each type of learning, it was not possible to compare the effect of the different types of learning neither the effect of the number of presentations during learning on longer-term retention. For both 4-year-olds and adults, data were pooled but a color coding helps distinguishing the type of learning accomplished by each participant [Table 7].

3.3 Results

Children whose parents have agreed to come back to the lab, underwent a long-term memory test after a delay of either one month or six months. Some adult participants who underwent a first test after a 30min break were unexpectedly re-tested after the same amounts of time. In order to explore whether the first memory test benefitted long lasting retention, another group of adult subjects did not undergo any test between the learning phase and the long-term testing phase. Results are represented in details in Table 7 and in Figure 56.

Overall, the mean performance of the 4-year-old participants who make up this sample was highly above chance level after a delay of 30min ($M=69.43$, $SD=28.12$, exact binomial $p=2.2e^{-16***}$). Those of them who were re-tested after a break of 1 month still performed significantly above chance ($M=58.64$, $SD=27.7$, exact binomial $p=2.5e^{-11***}$). A significant difference in performance could be established between the two latencies, indicating that those 4-year-olds were significantly better at recognizing the associations after a 30min delay than after a 1month delay (GLMM, $Z=-2.3$, $p=0.018^*$).

In contrast, the 4-year-olds who underwent a 6-month time interval with no re-exposure to the material completed the second memory test at levels expected by chance ($M=33.33$, $SD=23.2$, exact binomial $p=1$) [Figure 56]. Additionally, those 4-year-olds performed significantly worse during this second memory test than during the first test completed 30-min after learning (GLMM, $Z=-3.42$, $p=0.0006***$). They also performed significantly worse than the 4-year-olds who underwent their second memory test after a break of 1 month (GLMM, $Z=-2.4$, $p=0.016^*$).

Table 7 – Details about the composition of the two samples (4-year-olds and adults) who underwent a long-term memory test.

	ID	Sexe	Age (Y)	Learning condition	Perf. at 30 min (%)	Perf. at 1 month (%)	Perf.at 6 months (%)
4-YEAR-OLDS	1	F	4.04	FM	100	77.77	
	2	M	3.8	FM	100	100	
	3	M	4.48	FM	100	100	
	4	M	4.22	Ost lab static	100	100	
	5	M	3.85	Ost lab static	88.88	44.44	
	6	F	3.65	Ost lab static	88.88	44.44	
	7	M	5.05	FM	88.88	33.33	
	8	F	3.63	Ost lab static	88.88	55.55	
	9	F	4.35	Ost lab static	77.77	88.88	
	10	M	4.6	Ost lab video	77.77		11.11
	11	F	4.7	Ost lab video	77.77		33.33
	12	M	4.3	Ost lab video	77.77		33.33
	13	F	3.85	FM	66.66	33.33	
	14	F	5.01	1 repetition	66.66	33.33	
	15	F	4.6	1 repetition	66.66	88.88	
	16	M	5.05	1 repetition	66.6	77.77	
	17	M	3.9	FM	55.55	44.44	
	18	F	4.03	Ost lab video	55.55		0
	19	F	4.32	Ost lab video	55.55		33.33
	20	M	4.19	Ost lab video	55.55		0
	21	F	4.07	Ost lab video	55.55		66.66
	22	F	4.93	Ost lab video	66.66		55.55
	23	F	4.04	Ost lab video	66.66		44.44
	24	M	4.09	FM	33.33	33.33	
	25	M	4.48	1 repetition	33.33	22.22	
	26	F	4.07	Ost lab video	100		55.55
	27	M	3.93	FM	22.22	44.44	
	28	F	4.27	FM	11.11	33.33	
ADULTS	29	M	23	1 repetition	100	100	
	30	M	22	1 repetition	100	77.77	
	31	M	22	1 repetition	100	11.11	
	32	M	22	Ost lab video	100	88.88	
	33	F	21	FM	100	77.77	
	34	F	21	FM	100	44.44	
	35	F	22	FM	100	88.88	
	36	F	47	FM	100	66.66	
	37	F	20	FM	100	100	
	38	F	28	FM	88.88	100	
	39	M	37	1 repetition	88.88	66.66	
	40	M	22	FM	88.88	22.22	
	41	F	30	Ost lab video	88.88		88.88
	42	M	23	Ost lab video	100		100
	43	F	27	Ost lab video	77.77		77.77
	44	F	21	Ost lab video	77.77		33.33
	45	M	24	Ost lab video	33.33		33.33
	46	F	21	FM		100	
	47	F	22	Ost lab video		100	
	48	F	23	Ost lab video		100	
	49	F	23	Ost lab video		88.88	
	50	F	23	FM		33.33	
	51	F	25	FM		33.33	
	52	M	22	Ost lab video		33.33	
	53	M	21	Ost lab video		33.33	
	54	M	20	Ost lab video		33.33	

Adults that make up this sample performed highly above chance level after a delay of 30-min ($M=90.04$, $SD=17.7$, exact binomial $p=2.2e^{-16***}$). Adult participants who were re-tested after a break of 1 month still completed the test highly above chance level ($M=70.37$, $SD=30.09$, exact binomial $p=2.56e^{-15***}$) but they performed significantly worse than after a 30min delay (GLMM, $Z=-4.55$, $p=5.27e^{-6***}$) [Figure 56]. Adult participants who did not undergo a first memory test after a 30min delay similarly recognized the associations 1 month after learning ($M=61.72$, $SD=33.8$, exact binomial $p=1.57e^{-7***}$). No significant differences could be established between the performance of adults who underwent a first memory test after a delay of 30min and those who did not (GLMM, $Z=0.31$, $p=0.75$) [Figure 56].

Finally, the adult participants who underwent a long-term memory test after a delay of 6 months also recognized the associations greater than what would be expected by chance ($M=66.2$, $SD=33.4$, exact binomial $p=4.31e^{-6***}$) but succeeded comparatively worse than they did after a 30min delay (GLMM, $Z=-2.07$, $p=0.038^*$). No significant differences could be established between the performance of adults who underwent a second memory test 1 month after learning and those who underwent their second memory test 6 months after learning (GLMM, $Z=-0.4$, $p=0.68$). Note that more adults need to be re-tested, especially after a 6-month delay, to enable us to draw solid conclusions.

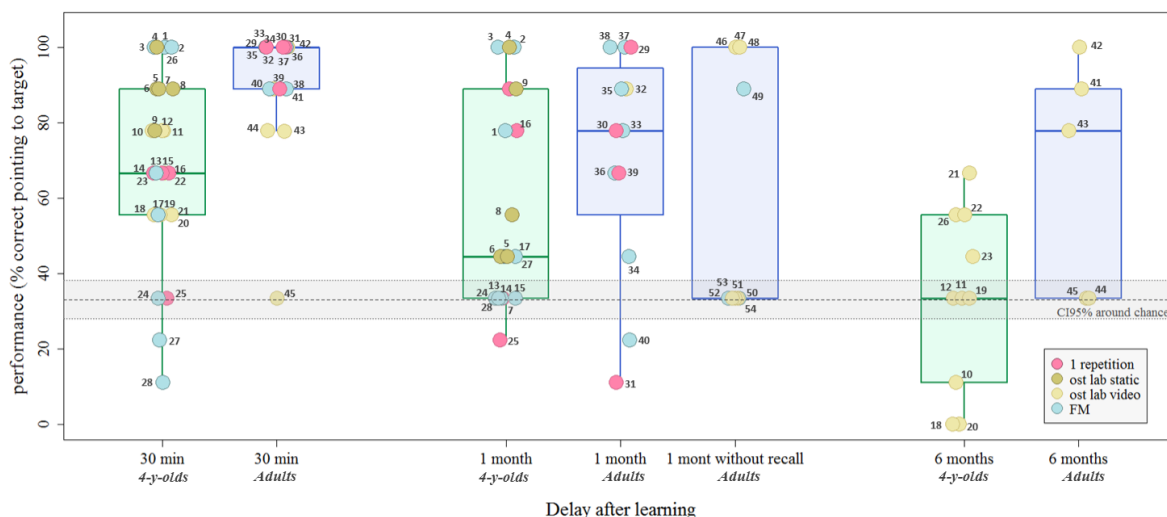


Figure 56 – Boxplots representing the performance of 4-year-old and adult participants to recognize word-object associations after a 30-min delay (left), a 1-month delay (middle) and a 6-month-delay (right). One group of adults did not perform a previous test after a 30min break (i.e. “1 month without recall”). The ID numbers correspond to those utilized in the former table (i.e. refer to table for more details about each subject). The green boxes correspond to the 4-year-olds and the blue boxes to the adults. The mean performance of each participant is represented by a filled circle with a color coding that corresponds to the type of learning (light yellow=ostensive labeling via video clips (1,3 and 5 presentations), 1st study ; dark yellow=ostensive labeling via static images (1, 3 and 5 presentations), 4th study ; pink=ostensive labeling via video clips (1 presentations), 2nd study ; turquoise=“fast-mapping”, 3rd study). The grey dashed line represents the level expected by chance and the 95% confident interval.

When comparing the long-term retention skills of 4-year-olds with those of adults, no significant differences could be established after a delay of 1 month (GLMM, $Z=1.12$, $p=0.26$). However, after a delay of 6 months, adults significantly outperformed 4-year-olds (GLMM, $Z=2.33$, $p=0.02^*$).

3.4 Discussion

This study addressed the question of the robustness of the memory of 4-year-olds to recognize cross-modal stimuli a long time period after initial learning without reactivation of the information in-between. To that aim, a sample of 4-year-old participants were re-tested on the word-object associations they learnt either 1 month or 6 months earlier and their accuracies were compared to those they accomplished 30-min after learning. Moreover, retention skills of 4-year-olds were compared to those of adults who were similarly unexpectedly re-tested after the same lengths of time.

Although this study is ongoing, current results suggest that 4-year-olds are indeed capable of recognizing word-object associations after a 1 month intervening period but not after a 6-month delay. In contrast, adults were as accurate to recognize the associations after a 1-month break as after a 6-month break. Four-year-olds were as accurate as adults to recognize the associations 1 month after learning but were significantly worse than adults after a break of 6 months with no reactivation of the information in-between.

First of all, these preliminary results support the assumption that children of 4 years of age are capable of remembering the names of objects presented a very few times - or even only once - 1 month ago. Previous studies demonstrated that 4-year-olds could recognize an association that was fast-mapped one month earlier (Carey & Bartlett, 1978; Markson & Bloom, 1997b; Vlach & Sandhofer, 2012; Waxman & Booth, 2000). If our preliminary results are confirmed with the inclusion of more participants, this would definitely attest that 4-year-olds are capable of maintaining at long-term (at least 1 month) the memory of not only one but several word-object associations.

However, we cannot exclude the possibility that this result is mainly attributable to the first memory test our participants underwent 30-min after encoding. Indeed, the critical role of retrieval practice in long-term retention is now well-established (Roediger & Butler, 2011 for a review). Retrieval practice plays the role of a powerful mnemonic enhancer, producing even often larger gains in long-term retention relative to repeated studying (Roediger & Butler, 2011). We have started to address this issue with a first sample of adult participants who did not undergo any

test trial between the encoding phase and the long-term testing performed 1 month later. Our preliminary results suggest that adult participants did not benefit from the first memory test since no significant differences were established between the subjects who underwent a first test after a 30min break and those who did not. Whether this finding would also hold for 4-year-olds is still unclear. Clearly, this condition should now be tested with young children as well.

Second, this ongoing work suggests that although retention is reported after a delay of 1 month in young children, the memory trace seems to vanish during the following months if no reactivation occurred. Indeed, we found that after an intervening period of 6 months after last practice, the performance of the 4-year-olds as a group to recognize the associations was not significantly different from the level expected by chance, although few of them (three) seemed to have performed a slightly better than the rest of the sample (i.e. outliers). In contrast, adults as a group still recognized the associations after a 6-month delay, suggesting that unlike children, the memory trace remains stable at long-term in adults which is in accordance with the findings of previous studies (Larzabal et al., 2018; Larzabal et al., 2017). This result is particularly interesting regarding the phenomenon of juvenile amnesia. Here, we suggest that a memory for complex and abstract bimodal associations (sound patterns arbitrary paired to visual items) can survive at least one month but less than 6 months in preschoolers without re-exposure. This would mean that not only episodic memory but also semantic knowledge is affected by this phenomenon. Many studies reported the cases of adults who emigrated or were evacuated from their native country (because of a war for example) as children and who never returned in their inborn country since (e.g. Mattsson, Maliniema-Piispanen, & Aaltonen, 2015; Portes & Hao, 1998). Intriguingly, all of them completely lost their mother tongue (Mattsson et al., 2015) or were no longer fluent in their parental languages (though there were wide variations among immigrants in the extent of their parental linguistic retention) (Portes & Hao, 1998). In Mattsson et al's work, the majority of adults were aged between 2 and 4 years by the time of the evacuation (another one was aged 5 years and a last one 7 years) (Mattsson et al., 2015). Although the emotional trauma may have played a critical role in the loss of the mother tongue of those persons, it may also be possible that the absence of re-exposure to the native language for such a long time period made the information inaccessible.

A perspective to that work would be to test children after latencies of 2, 3, 4 and 5 months to establish whether the memory strength declines linearly with time or brutally after a certain time delay. It would also be very interesting to test 2-year-old participants after such longer delays in order to highlight potential differences between 2 and 4-year-olds in their ability to keep the memory of bimodal sensory information "dormant" in their brain. A recent study demonstrated

that 30-month-olds were able to retain the names of objects presented four times each over a week (participants also underwent a first memory test soon after learning) (Wojcik, 2017). It would now be interesting to explore whether similar results could also be found after a delay of at least 1 month in 2-year-olds.

At the moment, we do not have enough participants in each learning condition to allow fair comparisons of the results according to the type of learning. Nevertheless, at first glance, it would appear that the results are fairly homogeneously distributed regarding the type of learning (e.g. participants arising from the “ostensive labeling with videos” learning condition were not systematically worse or better than those arising from the “fast-mapping” learning condition). In adults, one observation caught our attention. It seems that the adult samples can be fairly divided into two subgroups whatever the delay of testing (1 month or 6 months) and whatever the learning method: those who well-recognized the associations (hits>80%) and those who did not (hits≈33%, chance level). There seems not to be in-between results in adults. This would suggest that some adults perfectly maintained the information in memory and were accurate from the first trial, while others completely “lost” the information or at least had not access to the information anymore. In contrast, one explanation for the results of children who performed “in-between” is that they failed on the first trials, but then were able to use a completion strategy to increase their performance across trials (for example, they may have randomly attributed objects to the first two labels they heard during testing, then they truthfully recognized the object corresponding to the third label and finally adjusted their response choices during the following test trials). This mechanism of *pattern completion* involves the CA3 field of the hippocampal system which operates as an auto-association network providing completion of the whole memory during recall (e.g. Rolls, 2010, 2013 for reviews). Given the prolonged maturation of the hippocampal system, it might be that 4-year-olds can rely on this mechanism to improve their performance and to recall a significant part of their memories after a one-month delay, but not anymore after a 6-month delay. If too little information is accessible after a prolonged time interval (weak synaptic weights), pattern completion may not help 4-year-olds to recover the memory.

Finally, when looking a little longer at individual performances, it can be noticed that most 4-year-olds who perfectly succeeded (88-100% accuracy) during the first memory test similarly well-recognized the associations after longer delays which would suggest that once an information is reliably and robustly encoded, that knowledge does not decay much with time. Those who relatively well-performed (55-77% accuracy) during the first test, either performed in a similar range at long-term or completely failed to recognize the pairings. Finally, the 4-year-olds who

mostly failed during the first memory (22-44% accuracy) test non-surprisingly similarly failed at long term.

To conclude, this experiment - though in progress – clearly suggests that **a small number exposures to novel cross-modal information are enough to induce the consolidation of that knowledge into more permanent representations in children as young as age 4 years**. The memory trace can survive longer delays without rehearsal but seems to have disappeared less than half a year after encoding. This is in line with the assumption that long-lasting memories are unlikely in early-developing brain systems that undergo heavy brain maturational changes. This would mean that **for semantic information acquired during early childhood to survive until adulthood, the information may need to be updated** (in the sense of reactivated) perhaps many times spaced in time. To test that hypothesis, it would be interesting to compare long-term retention skills of two groups, one provided with a rehearsal of the material to learn during the time interval while the other not. In contrast to children, here we showed that adults can form and maintain at very long-term, memories for arbitrary stimuli encountered very few times during a single learning session.

4 GENERAL DISCUSSION

In this first part of my thesis, we investigated the ability of young children and adults to remember the names of newly learned objects, as well as the effect of different factors and learning strategies on the formation of those memories. To reach these aims, we developed a tightly controlled but playful and ecologically valid design. For the first time, participants from different age ranges (18-month-olds, 24-month-olds, 4-year-olds and adults) underwent learning procedures (either (i) ostensive labeling with video clips, (ii) ostensive labeling with static images or (iii) “fast-mapping”) under the exact same conditions (stimuli, apparatus, experimenter, experimental room, delay, number of repetitions, etc.) allowing direct comparisons between age groups and between learning conditions which had never been done before at such a large scale.

To conclude on the main results we obtained through our six studies (though 3 are still in progress) and to link our findings with the issues of the M4 project, it can be said that from 2 years of age human beings can recognize complex cross-modal stimuli they encountered only once half an hour ago. Moreover, we highlighted the importance of providing repetitions during the encoding phase and showed that even after only a 30min delay, the strength of the memory increased roughly linearly with the number of presentations, whatever the learning method. Thus, it can be suspected that repetition is even more important for a memory trace to survive longer

delays during childhood. Indeed, recognition after a delay of 1 month was possible in 4-year-olds for whom the memory was not refreshed during the intervening period, but recognition was absent if the delay was increased up to 6 months. This finding has presumably to be linked with the maturational processes that occur early in life (e.g. synaptic pruning within the brain regions involved in memory formation), since the memory of adults in our experiments was much less affected by an absence of reactivation during a 6-month delay. According to my opinion, during early childhood, a neuron does not necessarily lose its selectivity and perhaps remains “dormant” for a long time period, but in the absence of regular (direct or indirect) re-exposures to that particular material, it is the neural pathway that is damaged or pruned. Only the neural networks that are frequently activated would be preserved, and thus at a behavioral level, only the information that are often renewed would be maintained at long-term. This hypothesis makes sense with the fact that children will not be able to recall personal events when adults, but in contrast will show a preserved semantic memory for information that they encountered many times (remember that in the absence of re-exposures to general knowledge, even a semantic information will not survive until adulthood; e.g. people exposed to another language early in life). However, our finding that even a single visual exposure was sufficient for children as young as 2 years of age to remember object names after a 30min delay seemed at first glance surprising regarding the literature and Hebbian theories. Nevertheless, we discussed the possibility that multiple very brief visual repetitions occurred during the trial sequence and also suggested that repetitions are still required for the information to survive longer delays. We also assumed that recognition of these pairings following a single exposure was specific to the learning strategy implemented.

Indeed, we also demonstrated in this thesis that not all learning methods were efficient to promote the formation of a memory trace during childhood, highlighting the extreme importance of using adequate and age-fitted methods when teaching knowledge to young children. In the context of word learning, we showed that there is a developmental change in the efficacy of an active learning procedure and that this shift occurs between 2 and 4 years of age. While younger children mostly disambiguate a situation for their immediate communicative requirements, being actively involved in the learning process is a very powerful method in older children (4-year-olds). Interestingly, the efficacy of such a learning method is also sustained in young adults but then declines with age during adulthood.

In addition, we showed that the amount of extra-cues (object manipulation, object functionality, length of exposition, number of verbal repeats, etc.) does not seem to be vital in post-vocabulary

spurt children who are instead able to extremely rapidly learn and maintain in memory the names of novel objects provided that they are unambiguously presented and labeled.

Finally, we also tried to understand some of the potential mechanisms that could have accounted for the results we obtained. We assumed that a minimal threshold of attentiveness and language skills (i.e. word comprehension and production) are required to rapidly form a memory between both an object and its associated label. The results showed that 18-month-olds who were on average the significantly less attentive participants compared to the other age groups and who were predominantly pre-vocabulary spurt children, failed to recognize the pairings after a 30-min delay unless the associations were presented at least 5 times. Moreover, we also suspect that our findings were largely driven by the maturational stage of the underlying brain structures. As reviewed in detail in the chapters 1 and 3, and as already discussed in earlier discussions, there is still a debate about the brain structures involved in the consolidation process of semantic information during early childhood. There is also an active debate about the implication of the hippocampus in the “fast-mapping” learning process and the question of whether ostensive labeling and “fast-mapping” depend on similar mechanisms and brain structures. To briefly summarize, some researchers postulate that word learning essentially relies on extrahippocampal neocortical regions, especially antero-temporal and temporo-parietal regions (Atir-sharon et al., 2015; Schmolck et al., 2002; Sharon et al., 2011; Vargha-Khadem, 1997). According to them, semantic memory is only acquired through a slow consolidation process that requires statistical regularities (McClelland, 2013). However, under certain conditions (e.g. “fast-mapping”), the neocortex may be capable of rapid learning independently of the hippocampus (Atir-sharon et al., 2015; McClelland, 2013).

Other researchers consider that the first stages of word learning are mediated by slow processes independent of MTL structures, and reliant on repetitions (~procedural learning), but later fast word learning would truly and primarily depend on the declarative memory system (i.e. especially the MTL structures) (Friedrich & Friederici, 2011a).

Finally, other authors assume that the declarative memory system supports word learning from its beginning and that the behavioral shift from a slow, repetition-based learning to a fast, trial-unique word learning is mainly attributable to the maturation of specific MTL structures (especially the dentate gyrus) (Jabès & Nelson, 2015; Nelson, 1995). According to this last viewpoint, in the first stages of language acquisition, word learning would be sustained by a hippocampus-dependent “pre-explicit” memory system. Only neuroimaging studies may help reconcile these different points of views, although fMRI studies involving young participants are known to be very challenging to carry out.

PART 2



LEARNING STRATEGIES AND LONG-TERM MEMORY OF WORD-OBJECT PAIRS IN DOMESTIC DOGS



Moka

Abstract

In the past, animal studies have rarely given interest to memory, which has always been perceived as a “higher order mental process”. Aristotle’s assertion that animals are incapable of higher mental processes, in addition to the solid traditional view arising from the *behaviorism*, were until recently, predominant in general view. In the late 1970s, advances in human memory research attracted many animal researchers and psychologists, who initiated experiments and transferred memory models developed in humans to the study of animal cognition. Specifically, researchers attempted to discover if cognitive principles or mechanisms, still believed to be unique to humans (e.g. language, memory, etc.), may be ultimately shared with other living creatures. Despite this, there is still currently a huge gap in our knowledge about animals’ long-term memory skills and the underlying mechanisms. Therefore, this second experimental project was dedicated to testing the generality of some memory principles that form the heart of the M4 project to a distantly related animal model: the domestic dog (*Canis lupus familiaris*). Since a handful of studies showed that individual dogs were capable of learning the referential relationship between words and objects, we decided to focus our investigations onto long-term memory in dogs on that specific complex cross-modal paradigm.

The first chapter will be devoted to the state-of-the-art about canine cognition, and will show why dogs represent nowadays a model of choice in animal cognition research. In a second chapter, we will present our findings about long-term memory for word-object pairs in a heterogeneous sample of dogs. More precisely, 40 dogs initially naïve about the task were recruited and underwent a six months training period to acquire the word-object associational concept. Comprehension tests were carried out at the end of the training phase and memory was assessed after a break of three months without rehearsal during the intervening period. We also explored if dogs’ cognitive processes follow similar rules to those of humans, such as an age-sensitivity to learning and remembering a novel abstract task at long-term. In the last chapter, we will present the series of exploratory experiments we conducted with our best word-trained dog, which should enhance our comprehension about the conditions required for memorizing novel information, such as novel object names. It should also help better characterize the nature of the word-object paradigm in dogs, allowing suppositions about the underlying type of memory and thus about the recruited brain systems. To reach these aims, we reproduced the learning strategies employed with children to examine their hypothetical extensibility to this unrelated animal and their efficiency on learning and memorizing new words. Lastly, we will discuss the meaning that could be given to “word learning” by dogs and relate recent findings that encourage us to believe that dogs do create internal representations of sensory information.

CHAPTER VI.

CANINE COGNITION RESEARCH: STATE-OF-THE-ART

For thousands of years, dogs have shared humans' daily lives. As the oldest domesticated species, dogs, more than any other species, have acquired the ability to understand and communicate with humans and there is a general consensus that most people like dogs. However, despite their high prevalence in modern human societies, very few have found them interesting from a scientific perspective. Nevertheless, dogs' impressive social and non-social skills, shaped by millennia of cohabitation with humans and of intense selective breeding, have only caught the attention of researchers in the last two decades or so. This chapter will provide an overview of canine research from the origin of the process of dog domestication to the neural correlates underlying canine cognition.

1 Origins and evolutionary social skills of the “man’s best friend”

The dog is a young species on the current phylogenetic tree as evidenced by mitochondrial DNA sequence analyses. Indeed, dogs originated a proper clade about 100,000 years ago, as they evolved away from the common ancestor they shared with wolves (Vilà et al., 1997). The dog is the **first domesticated species**. Dogs have lived alongside humans since the Upper paleolithic period (~ 35,000 years BP) and evolved in an extraordinarily large range of phenotypes according to natural and artificial selection. Firstly, the change around 10,000 to 15,000 years ago from nomadic hunter-gatherer societies to more sedentary agricultural populations may have imposed new selective regimes on dogs resulting in adaptive phenotypic changes (Vilà et al., 1997). Secondly, evidence from genomic DNA studies shows that dogs and wolves continued to exchange genes after the origin of dogs from a lupine ancestor. These backcrossing events may have enriched domestic dogs through periodic interbreeding and consequently played a critical role in the diverse evolution of dog breeds (Vilà et al., 1997). By the time of the Ancient Egyptian civilizations (~ 3,000 years BP) distinctive breeds of dogs had been characterized and were selectively bred for precise purposes. Specifically, dog breeds served as hunters, guards or life companions (Galibert, Quignon, Hitte, & Andre, 2011) [**Figure 57**].



Figure 57- Illustrations of dogs throughout Human history. Upper left corner: Renegade Canyon, California, about 4000 BP. Upper right corner: Dog of Fo, Chinese polychrome statuette, about 960–1000. Burchard gallery. Lower left corner: Artemis goddess, greek oenochoe. From Athens, about 450 BC, Louvres Museum. Lower right corner: “Très riches heures du duc de Berry, janvier”, details. Limbourg brothers, 1416. From Galibert, Quignon, Hitte, & Andre, 2011.

The fragmentation of several hundred various breeds took place during the Middle-Age and Renaissance periods as well as during the more recent centuries as a consequence of an **intense selective pressure** on dogs to produce phenotypes that were better **adapted to human purposes** (Galibert et al., 2011). Many breeds were indeed developed to perform multiple additional tasks [Figure 58] such as herding, helping the physically impaired, identifying diseases, rescuing people, helping to maintain public safety by finding drugs and explosives, etc. (see Bensky, Gosling, & Sinn, 2013 for a review) [Figure 58].

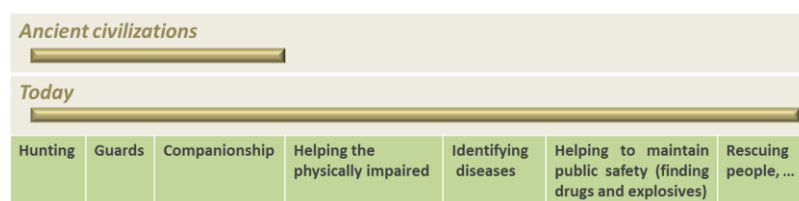


Figure 58- Illustration of the numerous functions dedicated to dogs during the ancient civilizations and nowadays. Such specific functions follow an intensive selection on dogs to generate purpose-adapted breeds.

Noticeably, this selective pressure favored the emergence of **social-cognitive capabilities of dogs** (Hare, Brown, Williamson, & Tomasello, 2002). Indeed, the roles dedicated to dogs - including companionship - are cognitively challenging. To achieve such tasks, dogs need to learn numerous environmental and social contingencies and extend this knowledge to a wide range of contexts, as well as problem solving and decision making. The way dogs process information has also been profoundly impacted, enabling them to understand human social and communicative behaviors in

unique ways as shown by multiple studies (e.g Frank, 2011; Hare et al., 2002; Hare & Hare, 2007).

In an **object-choice paradigm**, domestic dogs, semi-socialized wolves and chimpanzees had to find under which of two opaque containers a food reward was located by using conspicuous communicative cues provided by the experimenter (for example looking at, pointing at, tapping on, or placing a marker on the correct container, **Figure 59**) (Hare et al., 2002). The results demonstrated that dogs were more skillful at using human social cues than one of humans' closest extant primate relatives - chimpanzees. In this study, dogs also outperformed their closest evolutionary relatives - wolves. Finally, dog puppies were as skillful as dog adults to solve this human-guided paradigm, irrespective of their age and rearing history with humans, ruling out the training hypothesis (Hare et al., 2002). Altogether, this research supports the evidence that dogs, unlike chimpanzees, developed human-like skills at **understanding human communicative signals** and that these skills likely evolved during the **domestication process**, as they are neither learned nor inherited from wolves (see also Hare & Tomasello, 2005 for a review).



Figure 59- Dogs are more skilled than chimpanzees at using human behavioral cues (e.g. pointing). The experimenter points in the direction of the correct cup and lets the dog choose a cup. *From Hare & Tomasello, 2005.*

A replication of this study with higher controlled socialization levels to humans refined the conclusions on this topic; It showed that socialized wolves raised in similar rearing conditions to dogs were indeed able to learn about human cuing, but significantly less well than dogs (Miklósi, Kubinyi, Topál, Gácsi, Virányi, & Csányi, 2003). Similarly, opponents of the domestication hypothesis to explain understanding of human communicative signals by domesticated dogs argued that some cases of non- or less-domesticated species (or at least individuals of those species) showed high levels of comprehension of human cues (e.g. the well-known case of a horse “Clever Hans” as being able to “count and read”; Candland DK, 1993) (see Miklosi & Soproni, 2006 for a review). As their key message, these authors advised to be highly cautious when

comparing cognitive abilities of broad species: one should consider both the individuals' heterogenic history and previous experience with humans as well as the experimental design before reaching firm conclusions.

Despite the lack of scientific consensus about which hypothesis is more likely to explain social cognitive evolution in dogs, such capacities remain unique in the cross-species relationship domain and could have converged with those of humans through a **phylogenetic process of evolutionary and ontogenetically constraints** (i.e. selective pressures on specific traits and artificial selection during domestication) and/or through **convergent cognitive evolution** with humans as a result of **thousands of years of cohabitation** (i.e. similar derived traits in distantly related species).

2 The ontogeny of lab research on canine cognition

2.1 Canine research in few figures

Dogs have become a fascinating model of research in understanding the evolutionary aspect of complex forms of dog-human communication. Not only this, they have also spurred an increasing interest in overall **canine cognition** over the **last 20 years**. From 1995 to 1999, only 14 articles about canine cognition were published (Bensky et al., 2013). This number has grown significantly between 2000 and 2004 and this increasing progression continues nowadays [**Figure 60**].

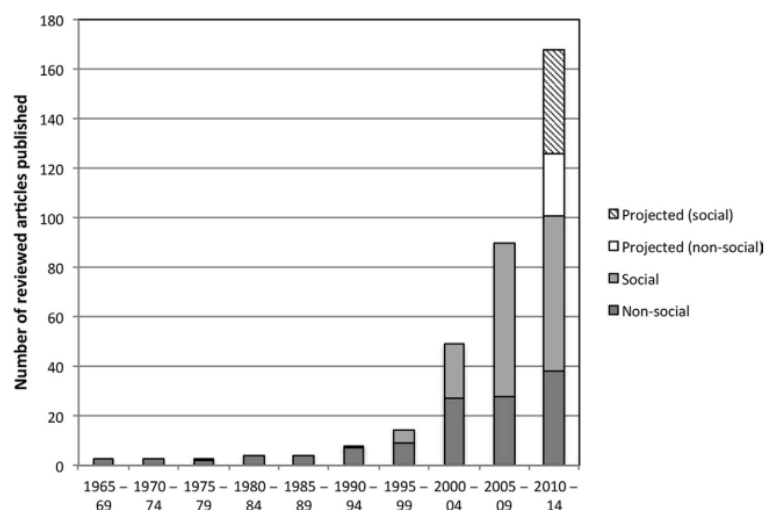


Figure 60 - Articles on dog cognition published every five years since 1965. Publications are divided into nonsocial and social publications. Projections through 2014 are based on the average number of articles published each year between January 2010 and December 2012. From Bensky, Gosling, & Sinn, 2013.

In brief, most of the studies about canine cognition focused on the visual sensory modality (~73%), while fewer used auditory related protocols (~19%) or olfactory manipulations (~8%) (Bensky et al., 2013).

Intriguingly, relatively few canine cognition studies recruited working-dogs (~9%). Instead, it appears that most of the subjects were pets volunteered by their owners (~71.9%) and relatively few were raised specifically for laboratory research purposes (~19%) (Bensky et al., 2013).

Additionally, Bensky et al. reviewed that a higher proportion of dog subjects were above one year of age (i.e. sexually mature adults; 87.7%) and fewer were below one year old (i.e. puppies; 13.4%). Moreover, only 12.6% of the studies focused on developmental research using cross-sectional or longitudinal techniques (Bensky et al., 2013).

Since the year 2000, wide scientific disciplines have illuminated research on canine cognition including ethology, developmental psychology, evolutionary anthropology, behavioral analytics and neuroscience. They have covered very broad topics including both social and nonsocial cognition.

To briefly summarize, *canine social cognition* comprises three major subcategories of research: (1) responses to human cues and **dog-to-human communication** (i.e. factors of effectiveness explaining how humans and dogs communicate with each other); (2) perspective taking (i.e. dogs' ability to evaluate and behave differentially depending on the perspective of others), and (3) **social learning** (i.e. how demonstrators and social interactions impact learning and expression of behaviors). *Nonsocial cognition* aims to investigate how dogs develop **mental representations** of physical stimuli from their surrounding environment and how they use them to solve a wide variety of tasks. To date, nonsocial cognition includes research on **discrimination learning**, **object permanence**, **object learning**, **categorization** and **inferential reasoning**, object manipulation in **problem solving**, **quantitative understanding**, **spatial cognition** and **memory** (for a review see Bensky et al., 2013). The research studies are driven by a major interest in canine social cognition as shown in [Figure 60](#), but there is a notable trend toward nonsocial cognition in the last fifteen years (Bensky et al., 2013).

2.2 Brain correlates of canine cognition

Since 2012, research on canine cognition is no longer restricted to behavioral studies, owing to the fact that notable improvements in the methods and equipment have recently permitted investigations on the underlying brain mechanisms. Indeed, **electroencephalography** (EEG) and

fMRI approaches in awake dogs have provided a unique opportunity to probe the neural underpinnings of the social and nonsocial capabilities of dogs. Although these methods are in their early stages, many discoveries have underpinned current knowledge, especially about the unique dog-human social bonding described above.

For example, it has recently been demonstrated that dogs are capable of **discriminating emotional expressions in human faces** (Muller, Schmitt, Barber, & Huber, 2015). The authors controlled that this ability did not merely depend on simple visual cues but instead did truly depend on extensive interaction with humans and/or domestication. Similar results were established in a cross-modal paradigm involving sensory emotional information (faces and vocalizations) (Albuquerque et al., 2016). Together, this suggests that dogs have developed high level **cognitive representations of human faces**. Such a remarkable performance is likely due to a **face-selective area** within the **ventral-posterior region of both temporal lobes** of dogs, shown to specifically respond to images of dog and human faces (Cuaya, Hernández-pérez, & Concha, 2016; Dilks et al., 2015). These findings provided evidence that a **neural mechanism dedicated to face processing** is not unique to primates and offers one possible explanation about the sensitivity of dogs toward human social cues, such as emotional states. Note that face selective neurons have also been reported in the temporal cortex of sheep, and that the responses of the highly selective cells were similarly probably influenced by factors related to social interaction (Kendrick & Baldwin, 1987). It can therefore be assumed that a face-selective area is not restricted to primates.

Another study demonstrated that dogs were able to discriminate a specific human scent from others. In this research, the **ventral caudate** was shown to be **maximally activated to the scent of the human** with whom the dog had a **positive experience** during the experiment (not necessarily the dog's handler) compared to self, strange human, familiar dog or strange dog scents (Berns, Brooks, & Spivak, 2014). Moreover, the **nucleus accumbens** was significantly more active to **reward-predicting stimuli** irrespective of the reward type, food or praise, relative to control unrewarded stimuli, supporting the evidence that this nucleus serves as a **neural predictor of reward in dogs** (Berns, Brooks, & Spivak, 2013; Cook, Prichard, Spivak & Berns, 2016). In humans, the corresponding primary reward regions (consisting of dopamine neurons of the ventral tegmental area that projects to the nucleus accumbens, a part of the ventral striatum) also consistently respond more strongly to rewarded than unrewarded stimuli (e.g. Haber & Knutson, 2009), suggesting analogies between the brain mechanisms in humans and dogs.

Other studies suggest that dogs present analogous brain functions with humans and/or developed apparent brain functions tuned to human properties. This has been particularly investigated in the

field of **cross-species vocal communication**, since the evolutionary processes not only enhanced dogs' ability to engage in acoustic communication with humans (Pongrácz, Molnár, & Miklósi, 2010; Pongrácz, Molnar, Miklosi, & Csanyi, 2005) but also enabled them to **develop sensitivity to human vocalizations** (e.g. Gibson, Scavelli, Udell, & Udell, 2014; Scheider, Grassmann, Kaminski, & Tomasello, 2011). In 2014, the first comparative neuroimaging study of dogs and humans demonstrated that for both species, **auditory regions coding for voices** were more strongly **activated for conspecific vocalizations** than for heterospecific vocalizations or nonvocal sounds (Attila Andics, Gacsi, Farago, Kis, & Miklosi, 2014) [Figure 61]. Interestingly, dogs also presented **auditory subregions** responding preferentially to **human vocalizations** (13% of all auditory voxels) or nonvocal sounds (48%). In humans, the subcortical medial geniculate body (MGB) responded stronger to dog sounds than to other sound types (10%).

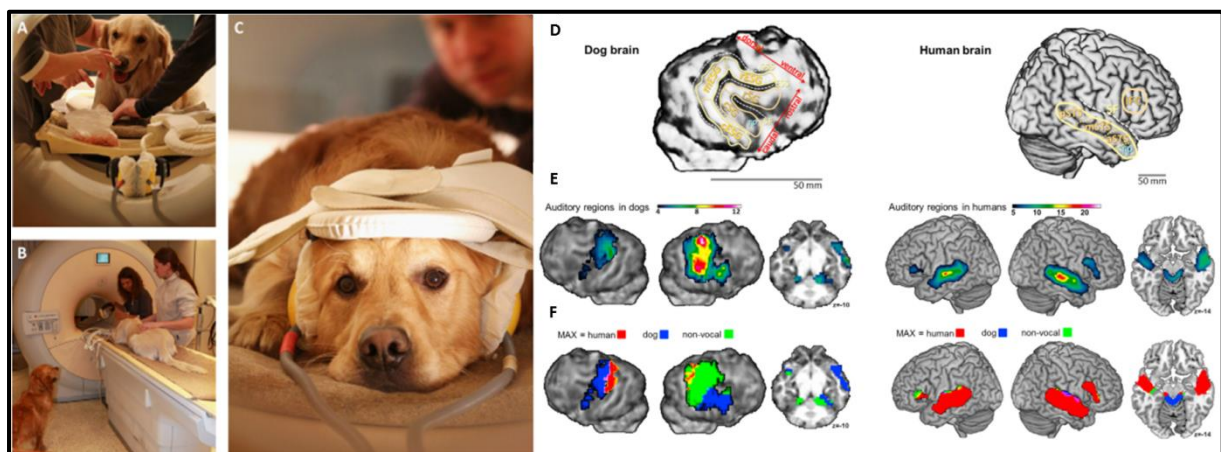


Figure 61 - A. Dog lying on a scanner bed, being rewarded with food and socially by the owner ; B. The tested dog is praised while receiving earphones from an experimenter ; C. The upper element of the coil is fixed with strips on the top of the dog's head ; D. Schematic representations of sound-sensitive perisylvian regions in dogs and humans, superimposed on rendered brains. Dog abbreviations: c, caudal; m, middle; r, rostral; ESG, ectosylvian gyrus; ESS, ectosylvian sulcus; SF, Sylvian fissure; SG, Sylvian gyrus; SSS, suprasylvian gyrus; TP, temporal pole. Human abbreviations: a, anterior; m, mid; p, posterior; IFC, inferior frontal cortex; SF, Sylvian fissure; STS, superior temporal sulcus; and TP, temporal pole ; E. Auditory regions as determined by the all sounds versus silence contrast in dogs and humans, thresholded at $p < 0.001$, FWE corrected at the cluster level, using the uncorrected voxel threshold $p < 0.001$ for dogs (in a whole-volume search space of 90 cm³) and $p < 0.00001$ for humans (in a whole-volume search space of 1,277 cm³). Color heatmaps indicate t values, superimposed on rendered brains and selected axial slices ; F. The same auditory maps as in (E). The color code refers to the sound type that elicited the maximal response in each voxel. Adapted from Andics et al., 2014.

The authors also identified regions sensitive to **emotional valence** in both dogs and humans. Conspecific and heterospecific **emotional vocalizations** perceived as more **positive** elicited greater neural responses near to the **right primary auditory cortex**, but no regions responded stronger to negative vocalizations. Their findings suggest that a similar mechanism may be used to extract and process affective information from heterospecific vocalizations in both species. In another **fMRI study** the same authors disentangled **lexical information** (praise words, neutral words) and **intonational information** (praised intonation, neutral intonation). They showed that

awake unrestrained dogs presented **left-lateralized cortical responses for lexical information** but no specific intonational effects in this hemisphere, supporting the evidence of a left hemispheric bias for processing intonation-independent meaningful words (A. Andics et al., 2016). The acoustic cues of **affective speech intonation** were processed in the **right middle ectosylvian gyrus** independently of word meaning. As with dogs, speech perception and emotional intonation processing in humans involves distinct brain regions (Belin et al., 2014; Farago et al., 2014).

Altogether, this study revealed **functional analogies** between dog and human brain mechanisms for **lexical and intonational information processing**. This raises the question of whether this ostensibly shared characteristic arises from an **evolutionary history** dating back to the common ancestor of dogs and humans (~100 million years ago) or from **convergent evolution**. Furthermore, although **neural evidence for lexical processing** in nonhumans is scarce, the neural capacity to process humans' spoken words does **not appear to be a uniquely human specificity**. As broader evidence, **humans' word comprehension by nonhuman species** has also been established at a behavioral level.

3 Word comprehension by dogs

3.1 Word learning within the animal kingdom

Words could be defined as **arbitrary sounds sequences mapped to a precise meaning**. Words are the basic building blocks of human languages. What makes humans unique is their verbal ability to produce and combine words into complex sentences and syntax, which represents a powerful **communicative tool**. Although this striking capacity of **meaning abstraction** and **flexible associations** is **restricted to humans**, vocal learning capacities are also patchily found in nonhuman vocal communicative systems. Indeed, only few taxa are able to produce sounds learned by imitation from a model: three distantly related groups of mammals (bats, elephants and some cetaceans) and three clades of birds (parrots, hummingbirds and songbirds) (Doupe & Kuhl, 1999; Janik & Slater, 1997; Knornschild, Nagy, Metz, Mayer, & Von Helversen, 2009). Nevertheless, whether or not nonhuman vocal learners produce **sound-specific meanings** that could be flexibly elicited according to various contexts is less clear. There are only few cases of nonhuman animals being able to associate arbitrary vocalizations with specific meanings. For example, meaningful lexical compositions are found in Campbell monkeys: they combine basic loud calls each associated to a highly specific meaning (e.g. designation of specific predators) into

different sequences and/or contexts, which influences the behaviors of the other members of the group (see Collier, Bickel, van Schaik, Manser, & Townsend, 2014 for a review).

While vocal learning is only sporadically found within the animal kingdom, learning and **discriminating heterospecific arbitrary sound sequences** seems to be a little more widespread. Here we will specifically focus on the ability of nonhuman animals to learn some aspects of human language. **Word (or sign) learning** has been clearly established in nonhuman animals that lived in **close contact and interaction with humans** and that were subjected to training on **human linguistic skills** (words, sign language, symbols, etc.). A few *chimpanzees* (e.g. Gardner, Gardner, & Van Cantfort, 1989), *bonobos* (e.g. Savage-Rumbaugh & Lewin, 1994), *gorillas*, *sea lions* (Schusterman & Krieger, 1984), *bottlenose dolphins* (Herman & Wolz, 1984), *African grey parrots* (Pepperberg, 2002) and more recently *domestic dogs* (Kaminski, Call, & Fischer, 2004; Pilley & Reid, 2011) have successfully decoded **unique or multiple-item “sentences”** composed of words (or signs) for locations, actions, objects, objects features, recipient, etc., spoken or elicited in variable combinations.

In 1984, two **sea lions** (*Zalophus californianus*) (a 6-year-old female, Rocky and a 3-year-old male, Bucky) were reported as being able to **discriminate distinctive gestural signs** (produced by movements of a trainer’s arm and hand) referring to types, attributes, locations of objects and actions to be taken in symbolic communication experiments (Schusterman & Krieger, 1984).



Figure 62 - Rocky orienting to a gray Clorox bottle. Other objects floating on the surface of the water include black and white Clorox bottles and black, white and gray water wings. The blindfolded trainer is holding the gestural sign referring to Clorox bottle until the sea lion returns it on the toe of the experimenter’s boot. The trainer had already transmitted the gesture referring to gray in a three-sign construction *GRAY CLOROX BOTTLE MOUTH*. From (Schusterman & Krieger, 1984).

Rocky and Bucky had a comprehension vocabulary of 20 and 16 signs respectively. They accurately decoded 3-sign combinations, such as *black-ball-mouth*, for “go over to the black ball (and not the gray or white balls) and place your open mouth on it” [Figure 62]. In each case, at least six objects varying in color and shape (e.g. balls colored in black, white or gray and baseball bats colored in black, white or gray) were present in the scene. These abilities required nearly 24 months and 20 months of training for Rocky and Bucky, respectively.

In another study, two **bottlenose dolphins** (*Tursiops truncatus*), Akeakamai and Phoenix, were taught **specific sounds** (generated by a computer-controller system projected underwater) and gestures for agents, objects, actions or modifiers of place or direction. They accurately responded with up to 5-sounds randomly recombined, which represented hundreds of uniquely meaningful “sentences” (Herman & Wolz, 1984). But, probably one of the most impressive language comprehension skills is attributed to **great apes**, who learned to decode **sign language**.

Beginning in the 60s, the first **sign language-trained chimpanzee**, **Washoe**, a female common chimpanzee (*Pan troglodytes*), inspired decades of research on the significance of human-like language abilities in apes. The Gardners reported that Washoe acquired 132 distinct and meaningful signs after 51 months of training (Gardner & Gardner, 1969; Gardner et al., 1989; Miles, 1991). After many years, she learned approximately 350 words of sign language. Over her entire life, she signed about objects present or absent, she made contrasts as in “same/different”, used negation, signed combinations, used signs to communicate new information, was able to generalize and, more impressively, invented her own signs. Her trainers, the Gardners, stated that such abilities share great similarities with those of very young children. Another famous ape, a **male bonobo** (*Pan paniscus*) named **Kanzi**, was known as having acquired spontaneous language, without being reinforced nor rewarded (e.g. Savage-Rumbaugh & Lewin, 1994). He was reported as being able to understand spoken English even though he had never been trained to, and initiated responses to English requests by using a keyboard composed of plastic visuographic symbols (lexigrams). Savage-Rumbaugh and colleagues claimed that such performances are at similar levels to that of a two and a half year old child. Later on, consistent results for sign-language were also demonstrated with gorillas and orangutans.

More recently, researchers revealed that not only are large-brained marine mammals or the closest human phylogenetically species, such as the great apes, capable of word learning, but that Grey parrots (*Psittacus erithacus*), a **human distantly related species** with a very small brain size, also compete in this field. Indeed, a **Grey parrot** named **Alex** developed remarkable **word comprehension skills** after many years of training. He knew the labels of more than 50 objects, 7 colors, 5 shapes, 3 categories (color, shape and material), etc. (Pepperberg, 2002).

Obviously, these human language-trained animals also inspired a vast and still topical debate about whether it reflects real “**language-like communication skills**”, results from reductionist “**operant conditioning**” learning rules, or from other non-defined processes, a topic that will be further discussed in a later section.

3.2 Word learning by dogs

3.2.1 *The word-referent concept: dogs retrieve objects by name*

Considering the growing evidence about canine cognitive competences, the presumably domestication-based high sensitivity of dogs to human communicative signals described previously, and the functional analogies between dog and human brain mechanisms for lexical processing, one would easily expect dogs to show word comprehension skills, like the other linguistically trained animals presented above. Dog owners often boast that their dogs obey multiple words for specific actions or requests, or that they understand few words referring to objects. However, such apparent skills have to be disentangled from gestural cues that owners explicitly or unconsciously address toward their pet dogs (e.g. pointing or looking at the object, wearing a coat while saying “let’s go for a walk”, etc.). Therefore, until recently, word comprehension by dogs suffered from a lack of controlled experiments that rigorously demonstrate this presumed ability.

In their pioneering study, Kaminski et al. provided the first evidence that a dog is indeed capable of language-learning abilities. A 9-year-old Border collie, **Rico**, was reported by his owners to know the **labels of over 200 items** (mostly children’s toys and balls) that he progressively acquired since he was 10 months of age, and which he apparently correctly retrieved upon request. A tightly controlled experiment was designed to discard any potential “**Clever Hans**” **effects**¹⁹ that might accounted for his performance. The experiment consisted of assigning 20 sets of 10 different objects randomly chosen among the 200 items reported as being familiar to the dog. The objects of each set were dispersed in an adjacent room, thus neither the dog nor his owner had visual access to the objects. Rico was requested to fetch randomly chosen objects by name from the separate room (one object after the other). The highly significant results of this experiment indeed confirmed that Rico knew the name of these items and that this performance could not be attributed to subtle or visual cues that guided his behavior (Kaminski et al., 2004). Thus, this experiment demonstrated for the first time that a dog is able to acquire a receptive “vocabulary size” comparable to that of the language-trained animals described above.

¹⁹**Clever-Hans effect:** *in the beginning of the 1900s, an Orlov Trotter horse was claimed as being able to solve arithmetic and other intellectual tasks. During a formal investigation, a psychologist discovered that the horse was actually not performing any mental task but simply “decoding” subtle cues that were unconsciously provided by the body of his owner or of the audience. In the absence of these involuntary cues, the “Clever horse” was no longer clever anymore. The discovery of this artifact led to a better control of the methods used in animal research. It is of major importance to rigorously control this effect in dog research, because dogs are gifted at attending to and interpreting subtle human social cues (e.g. Lakatos et al., 2016).*

In a second pioneer study, another Border collie named **Chaser** learned the **names of over 1000 proper objects** over a 3-year period of intensive training since her early puppyhood (Pilley, 2014; Pilley & Reid, 2011). Her owner, the retired psychologist John Pilley, aimed to investigate whether Rico’s acquisition of over 200 words represented an upper limit of word learning by dogs, or whether an intensive training could extend this genuine skill. Since Chaser was 8 weeks old, she was provided 4-5 hours of daily training. She was first taught general obedience, classical and operant conditioning on discriminative and associative tasks, herding, agility and tracking behaviors. In her fifth month, she was first introduced to fetching items by name, and was rewarded only with play and praise each time she retrieved a correct object upon request. She was taught one or two proper-nouns per day and had daily rehearsal testing for the objects for which she already knew the name. Therefore, she gradually increased her “receptive repertoire”. Over a period of 3 years, she acquired proper names for over 800 cloth animals, 116 balls, 26 Frisbees and over 100 plastic items. All objects differed in size, weight, texture, color, shape, etc. In rigorous experiments, analogous to those used by Kaminski with Rico, her ability to retrieve objects exclusively on the basis of verbalized names was positively confirmed.

A few other dog cases brought light to this field. Interestingly, a Yorkshire terrier case, a lap dog species that has not been bred neither for work nor for command obedience, was capable of learning a large vocabulary of spoken names (>120) (Griebel & Oller, 2012). This case study contributed to refute the speculation that only working dogs (like Border collies), specifically selected for their ability to understand human signals were able to associate labels to referents. The case of a mongrel dog able to respond to verbal requests composed of two independent words (object-action) also argues against this speculation (Ramos & Ades, 2012).

3.2.2 Multiple-items “sentence” comprehension by dogs

Kaminski’s pioneer study about Rico triggered massive **debates** about the “**linguistic-like**” **abilities of dogs** that she had emphasized. Soon after her publication in *Science*, Bloom conveyed his skeptical perspective about the nature of referential understanding by dogs (Paul Bloom, 2004). To him there is no proof that dogs reliably understood the label-referent concept. Alternatively, their abilities might be **limited to specific contexts** (e.g. the owner is the only one who pronounced the requests) and **specific routines** (fetching items). For example, in the utterance “*fetch the sock*”, Bloom wondered about Rico’s ability to separately treat the sound “*sock*” as referring to the specific sock item and the sound “*fetch*” as a behavior he should produce toward this specific item. Instead, he might have treated this sound as a **one-word proposition command** “*fetch-the-sock*”, which would have **little to do with language learning** in a human sense [Figure 63]. Therefore, Bloom claimed that further experiments were needed to help resolve these issues. More precisely, he addressed the following questions: “*would a dog learn a word for something other than a small fetchable object?*”; “*can a dog display knowledge of a word in some way other than fetching?*”; *would a dog be able to recognize an object with instructors other than those involved in the initial training process?*; *do dogs appreciate, like children do, that words can also refer to categories?* Another underlying question was related to dogs’ ability to disentangle words from actions in a two- or multiple-word utterance (Paul Bloom, 2004).

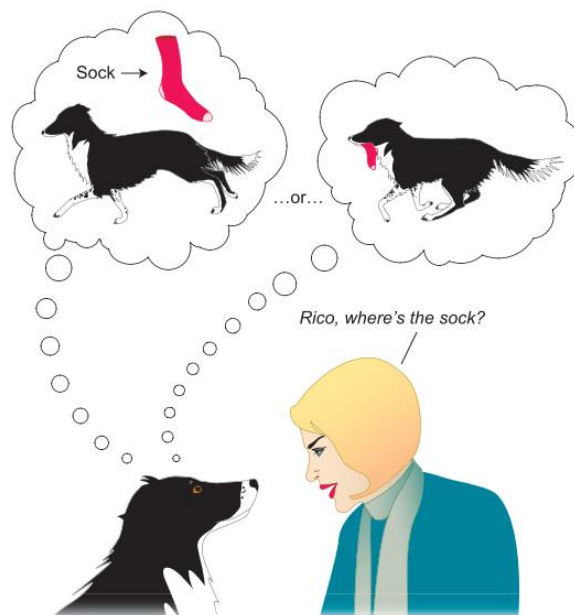


Figure 63 - Illustration of Bloom’s concerns about Rico’s ability to treat the utterance “fetch the sock” as two separated identities: one referring to the object and the other to the action to produce toward this specific object. If so, Rico might effectively understand that the word “sock” refers to a category of objects in the world and “fetch” (or the rest of the verbal command) refers to the fetching action. Alternatively, Rico’s understanding of this utterance might be limited to associating the word spoken by his owner with a specific behavior such as fetching a sock. Bloom, 2004.

To address Bloom's concerns, Pilley & Reid tested whether Chaser treated the name of an object independently from the given command (Pilley & Reid, 2011). The authors randomly paired different familiar commands (take, paw and nose) and familiar proper nouns (e.g. Lips, ABC, Lamb) that Chaser had never heard in combination prior to the tests (e.g. “*take Lamb*”, “*nose ABC*”, etc.). Testing was carried out in a double-blind procedure across 14 independent trials. In that experiment, Chaser performed without any errors, demonstrating that she could reliably produce the required responses oriented to each target object independently. This first evidence ruled out Bloom's concern and rather supported the evidence that Chaser brilliantly **dissociated independent meanings for references from verbal commands** (Pilley & Reid, 2011). In a succeeding study, the first author demonstrated that Chaser was even capable of appropriately treating each word of **sentences embodying three elements of grammar**, namely a prepositional object, a verb and a direct object (e.g. “to ball take Frisbee”) (Pilley, 2013). She was similarly successful when confronted with objects that had never been used in the training of three elements of grammar sentences. The author claimed that, considering these findings, she demonstrated **syntax and semantic understanding**.

In another study, the case of the mongrel dog already introduced earlier, provided further empirical evidence about the ability of dogs to appropriately respond to **combinatorial “sentences”** composed of a word designating an action and a word referring to an object (Ramos & Ades, 2012). Moreover, this dog also performed above chance when the experimenter inverted the order of the sentence items. According to the authors, this demonstrates that the performances were due to the learning of the sound sequences as single discriminative stimuli (Ramos & Ades, 2012).

3.2.3 *Word generalization abilities by dogs*

- *Word generalization to categories of objects*

As described in the section above, another concern addressed by Bloom pointed out dogs' restricted ability to learn proper nouns, i.e. sound sequences referring to a particular object. He wondered whether dogs, like children, would be able to treat words as **common nouns**, i.e. **sound sequences referring to a category of objects**.

In the last few decades, it has been widely demonstrated that categorization is a widespread cognitive ability within the animal kingdom, including dogs. As an example, Heffner's pioneer work demonstrated that dogs were able to **categorize auditory stimuli based on their source** (i.e. “dog” versus “nondog” sounds) and showed successful transfer to novel instances of these

categories (Heffner, 1975). Another study using a computer-automated touch-screen technology revealed that dogs successfully distinguished natural pictures of dogs from pictures of landscapes in a forced two-choice task procedure (Range, Aust, Steurer, & Huber, 2008). Dogs also successfully extended these categories to novel stimuli. The features used by dogs to solve this sort of task are poorly known, but data from other taxa may provide insight into how dogs might treat salient visual cues (e.g. Aust & Huber, 2001; Fabre-Thorpe, 2003). Furthermore little is known about the ability of dogs to form categories identified by the name of objects.

Pilley & Reid were the first to demonstrate that dogs could **learn labels for categories** (Pilley & Reid, 2011). Chaser was indeed successful in discriminating by name objects belonging to two subcategories of her toys sharing relatively similar visual properties: “balls” (round shape objects) and “Frisbees” (disk-shaped objects). More interestingly, she was also accurate in recognizing items from a “toy” category, thus objects that differed widely in their physical features, but that shared an abstract function: Chaser had been allowed to play with those “toys” in contrast to many other available objects with physically similar features with which she was forbidden to play, considered as the “non-toy” category [Figure 64]. “Balls” and “Frisbees” were two subsets of “toys”. Therefore, Chaser learned the label “toy” and “ball” for 116 different bouncy balls as well as each of their proper names. Similarly, she knew that the label “Frisbee” designated the 26 disk-like objects, for which she also learned each proper name as well the “toy” category label.



Figure 64 - Photographs showing the three categories of objects for which Chaser knew the common nouns. In the left photograph, the 16 objects used to test Chaser’s comprehension of the common noun “ball” (8 balls and 8 non-balls); in the middle photograph, the 16 objects used to test her comprehension of the common noun “Frisbee” (8 Frisbees and 8 non-Frisbees); in the right panel, the 16 objects used to test her comprehension of the common noun “toy” (8 toys (the tagged objects on the photograph) and 8 non-toys). *Adapted from Pilley & Reid, 2011.*

The testing procedure consisted of arranging sets of eight random exemplars of a category with eight random non-exemplars of the category (e.g. eight balls and eight non-balls) in an adjoining room out of the vision of both Chaser and the experimenter. Chaser was asked to retrieve each of the exemplars representing the category upon request, one after the other (e.g. “fetch a ball”, “fetch another ball”, etc.). For the three categories Chaser had never made any error. These results

indicate that Chaser **mapped one common noun** (“ball”, “Frisbee” or “toy”) **onto many objects**. She also demonstrated that she could **map up to three meaningful labels onto the same object** (e.g. the name of the functionally defined category “toy”, the name of the subcategory “ball” and the proper names of each corresponding balls). In humans, although there is a statistical tendency for words to be mutually exclusive, young children rapidly learn to break this word learning principle (e.g. a “cat”, “kitty” and “pet” designate the same referent), which might contribute to the rapid growth of their receptive vocabularies (Markman & Abelev, 2004). **Mutual exclusivity** is a default principle that is widespread in general cognitive tasks, i.e. not restricted to language or to humans. Therefore, Chaser’s demonstration of one-to-many or many-to-one label-referent mappings is a relevant support for the **referential understanding by dogs**. In addition, it suggests that mutual exclusivity was not the assumption by default that drove her performances.

- ***Word generalization to novel items***

Unfortunately, Pilley & Reid did not test whether Chaser was able to extend her word knowledge to novel instances of the known categories or, in the case of proper nouns, to unfamiliar objects sharing similar features. In another case study, a Border collie reported to know 54 object names was tested on the physical features that drove his word generalization ability to novel items (Van der Zee, Zulch, & Mills, 2012). The testing procedure consisted of presenting the dog with objects that varied in size (smaller *vs* same size *vs* bigger), texture (smooth *vs* rough) and shape (2 different shapes) relative to previously introduced objects. The testing took place either after a brief familiarization with a novel word-object pair (~10 min) or after an extended familiarization with the novel pair (~39 days). The results showed that when **briefly familiarized with novel word-object mappings**, the dog tended to **generalize the word to objects with the same size**. However, after a **prolonged familiarization phase**, the **texture** seemed to be the most important physical property that accounted for the dog’s retrieval choices. These results appear to contrast with the shape bias for word generalization observed in humans (L. B. Smith & Jones, 1988) but have to be cautiously considered as the authors only replicated the experiment twice, i.e. only tested generalization to novel items for two introduced objects (Van Der Zee et al., 2012).

- ***Word generalization to novel voices and accents***

The experiments described previously provided evidence that dogs are able to generalize known objects to categories of items shaped by physical or abstract properties. Moreover, during the generalization process, dogs seem to rely on information about texture for well-known objects and conceivably about size for newly introduced objects. Here the authors investigated whether dogs

could generalize their recognition of words to novel voices, thus to different voices than those of their trainers (Griebel & Oller, 2012).

In this experiment, similarly controlled forced-choice task trials were designed to assess the dog's performance. In a first case, a female with a German accent requested the dog to fetch random objects by name. In a second case, a male Native American English speaker ordered the retrievals. Both were novel speakers with novel accents from the dog's perspective. The results were highly significant for both speakers and thus clearly suggest that this dog (another Border collie) was able to **recognize words from his “receptive repertoire”** even when emitted by **unfamiliar voices** (Griebel & Oller, 2012). This demonstration put to rest Bloom's concern about word recognition with new speakers (Paul Bloom, 2004).

- *Decontextualized generalization*

There are a few more findings in the literature claiming that dogs' performances were not restricted to one specific training context. At least one dog was shown able to **generalize his “knowledge” to novel conditions**. Indeed, the mongrel dog introduced earlier, was capable of obeying verbal commands in a large range of contexts: 1) when he had **no visual access to the eyes or mouth of the experimenter** (which is in contrast with previous findings (Fukuzawa, Mills, & Cooper, 2005)); 2) when the requests were **pronounced by an unknown person**; 3) when the **spatial location of the objects was changed**; and 4) when **testing happened outside the laboratory** (Ramos & Ades, 2012).

To summarize, although it is necessary to remain cautious as all these reported cases are, to date, isolated cases of trained dogs showing intriguing “language-learning” skills, they have shed light on their **putative capacity to extract and process relevant verbal and visual features in a relative independence from contextual parameters**.

3.2.4 *“Fast-mapping” abilities by dogs*

We already reported numerous advanced skills of dogs in social and nonsocial cognitive tasks and we especially focused on their ability to learn the relation between both a word and its referent. Such skills, almost unique within the animal kingdom, may be inherited from the domestication process and/or from thousands of years of narrow cohabitation with humans. Researchers on canine cognition even claim that these abilities resemble “language learning” in humans and allow comparisons between the word learning aptitude of dogs and that of a 2 year old child (Kaminski et al., 2004; Pilley & Reid, 2011). Obviously this inspired animated debates and commentaries in the scientific community- as already discussed above (Bloom, 2004; Markman & Abelev, 2004).

Another effervescent point of debate comes from the **nature of the word-referent concept per se**. Unlike children, dogs' referential understanding seems limited to mapping words to referents after intensive training, which could be purely attributed to **mere associative learning mechanisms**, according to these later authors. If dogs indeed appreciate this complex concept, then they should be able to go beyond fetching routines with familiar objects, by at least showing abstract reasoning abilities about this concept.

In her pioneer study, Kaminski was the first to address this question. She tested whether Rico would be able to infer the referent of a new word on the principle of **exclusion reasoning** (i.e. the ability to base a decision on the exclusion of potential alternatives), that is to “**fast map**” and to retain this new mapping over time (Kaminski et al., 2004). The authors placed a new object along with seven familiar ones in an adjacent room. In the first one or two trials, Rico was asked to retrieve a familiar item. In the second or third trial, Rico was asked to fetch an object for which he had never heard the label. By excluding objects that did already have a label, Rico successfully retrieved the novel item in most of the cases (7 out of 10 sessions involving new sets of objects each time). **One month after this single exposure**, Rico's retention of the relations between the words he accurately mapped to the novel items was tested. During retention testing, each target item was placed together with four completely novel and four familiar items in a separate room. Again, in the first or first two trials, Rico was asked to retrieve a familiar item and in the second or third trial to retrieve the target one. The dog correctly retrieved 3 out of the 6 target items he mapped by exclusion one month previously (one target item was not tested for retention one month later). When retention occurred only 10 min after selecting the objects by exclusion (for other sets of objects), Rico performed slightly better as he correctly retrieved 4 out of 6 target items. Pilley & Reid obtained similar results with Chaser. Using the exact same paradigm, Chaser also successfully mapped novel words to novel referents by exclusion. She was equally successful in retrieving the newly named objects when tested immediately or 10 min after learning (she found 5 out of 8 objects after this time delay). However, she showed no sign of retention when tested with a 24 hours delay (1 out of 8) (Pilley & Reid, 2011).

According to these results, the authors assumed that their dogs acquired the abstract concept that things can have names for two main reasons. First, dogs' accuracy to map new words to their referents by exclusion **could not be explained by associative learning mechanisms**, because the **name and the referent were not presented together in temporal contiguity** (Pilley & Reid, 2011). Second, Rico stored the name of at least some of these referents in long-term memory, which suggests that **extensive training is not the only parameter that drives word learning** in dogs (Kaminski et al., 2004).

Nevertheless, the validity of Rico's pioneer demonstration of exclusion learning was questioned by Markman and Abelev (Markman & Abelev, 2004). According to them, there are two potential issues that could compromise the results. First, they pointed out a **lack of control for baseline preference**. It is well known that human and animal subjects show **neophilia** (i.e. "novelty bias") when facing novel and familiar stimuli (Kaulfuß & Mills, 2008). Thus, Rico could have retrieved each novel object just because they were **more salient, rather than making cognitively complex decisions** based on a "**Novel Name-Nameless Principle**" for example (cf. Part 2. for more details) (Kaulfuß & Mills, 2008; Markman & Abelev, 2004). Fischer et al. argued that Rico did not make any error when retrieving the familiar items prior to the new one. Thus if neophilia was influencing Rico's choices, he had at least exhibited an ability to inhibit any novelty preferences when accomplishing the familiar retrievals (Fischer et al., 2004). As a further reply to this concern, Pilley & Reid proposed a paradigm where they arranged two novel and eight familiar objects in an adjacent room and successively asked Chaser to retrieve each of the familiar items, one at a time. This procedure was replicated eight times with eight different sets of objects. Chaser never brought a novel item back, showing that her baseline rate for choosing objects based on novelty alone was null (Pilley & Reid, 2011).

Second, Markman & Abelev draw attention to the **reward** given to Rico subsequently to the correct referent selection trials as it could have **mediated the retention trials**. They questioned whether long-term retention following reasoning via exclusion would be possible with no positive feedback given to the dog (Markman & Abelev, 2004).

Finally, the authors stressed the fact that the procedure used to assess retention by both Kaminski et al. and Pilley & Reid may be controversial. To measure retention, they displayed the target item together with four familiar and four novel objects in a separate room. However, Rico's success could be merely attributed to a sort of "**extended exclusion**" (Griebel & Oller, 2012). Indeed, if he retained that the newly mapped object had been rewarded previously, he may have excluded the unrewarded objects (i.e. the novel ones), and the known items (i.e. those for which he already learned a label) when he heard the target label (Griebel & Oller, 2012). According to these last authors, a similar forced choice task where the items that had been successfully identified by exclusion were pitted against each other would be a more appropriate testing design. This procedure is the one typically used for measuring retention of fast-mapped words in children (e.g. Bion et al, 2013; Horst & Samuelson, 2008; Kucker & Samuelson, 2012; Axelsson et al. 2012).

To conclude this section, it is important to keep in mind that, although there are several studies showing reasoning by exclusion in nonhuman animals (e.g. Aust, Range, Steurer, & Huber, 2008; Herman & Wolz, 1984; Marsh, Vining, Levendoski, & Judge, 2015; Pepperberg, Koepke,

Livingston, Girard, & Hartsfield, 2013; Schusterman & Krieger, 1984), the specific **cross-modal “fast-mapping” ability displayed by dogs appears to exceed that seen in nonhuman primates**. However, undoubtedly, further experiments on “fast-mapping” are required and should provide insight on whether this remarkable skill translates a human-like understanding of the referential concept by dogs or if it only results from mere exclusion and working-memory skills.

4 Memory in dogs

The theoretical models of memory developed for humans struggle to have equivalences in models deprived of verbal production. In the **absence of language**, employing the term “**declarative memory**” can rationally be **questioned**. Given the lack of alternative animal-adapted models of memory, most researchers are still attached to studying animals’ memory abilities from a human perspective. Researchers also often directly compare memory skills of animals to those of humans, in order to investigate whether functions of memory systems are shared between humans and nonhuman animals. This comparative cognitive research has led to long-running debates. The most notable **debates stand around episodic memory in animals**. Indeed, for animals to fulfill the behavioral criteria for episodic memory, they should demonstrate an **explicit recollection** of knowledge about “*what*”, “*where*” and “*when*” an event occurred, as well as a “**mental time travel**” ability to internally replay the memory (Tulving, 1972). Therefore scientists cautiously talk about an “**episodic-like**” memory when examining this type of memory in animals. However bringing to light memory systems that handle incidental memory such as episodic memory could be very important for elucidating the degree to which memory functions depend on language.

4.1 Behavioral research on memory in dogs

4.1.1 *Working memory*

Memory is a major component of animal cognition. However, investigations on memory in dogs are seriously lacking. Currently, only working-memory and episodic-like memory have received actual attention. **Working memory is a limited capacity responsible for temporally holding information available for processing**. Working memory allows a **mental maintenance and manipulation of the information as long as the subject requires this information**, usually for a **limited time delay** (from seconds to minutes or hours according on the circumstances, paradigms and species). Being endowed with such cognitive aptitude could well be critical for animals to survive in the wild; For example, hunting often requires the predator to track and locate hiding prey. Fiset et al. replicated this phenomenon in an object permanence task (Fiset, Beaulieu,

& Landry, 2003). They applied various time delays between the disappearance of a moving object behind a box and the subsequent search for this box by the animal (amongst a choice of 4). Up to a 240s intervening period, dogs' accuracy to find the correct box was above chance levels but declined as a function of the length of the retention intervals (respectively 0, 30, 60, 120 and 240s) [Figure 65].

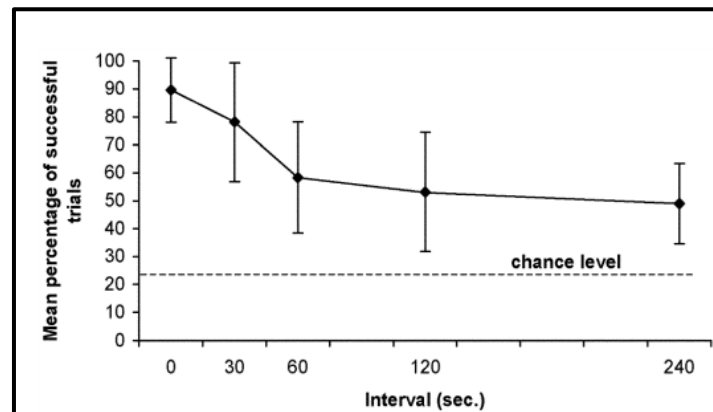


Figure 65 - Mean percentage of successful trials as a function of the retention interval in the object permanence task developed by Fiset et al. The experiment was designed to evaluate the limits of working memory for disappearing objects in dogs. From Fiset et al., 2003.

Delayed-NonMatching-to-Sample tasks (DNMS, i.e. avoid the stimulus presented prior to the delay and choose the novel one) have also been used to determine the duration of dogs' working memory. Intriguingly, dogs had difficulty acquiring a visual version of this task, especially aged dogs compared to younger ones (Milgram, Head, Weiner, & Thomas, 1994). However those who reached the 10s criterion within the 400 trials were subsequently tested for retention using variable delay intervals ranging from 10 to 50s. Most of the dogs performed significantly above chance irrespective of the delay intervals. The authors indicated that only a small correlation between accuracy and time delay was observed (Milgram, Head, Weiner, & Thomas, 1994). A spatial version of the DNMS task, the **Delayed-NonMatching-to-Position** (DNMP) task, demonstrated that dogs were more gifted to reach the 10s criterion in this case which suggests the **pre-eminence of spatial information on visual cues in dogs** (e.g. Chan et al., 2002; Head, Mehta, Hartley, & Kameka, 1995; Milgram et al., 1999). Those dogs who reached criterion still remained above chance levels for delay intervals up to 110s but their performances gradually deteriorated as the delay interval increased. Moreover, aged dogs showed analogous impairments to acquire the task and to perform higher than chance with increasing time intervals, as in the DNMS task presented above (e.g. Chan et al., 2002; Head, Mehta, Hartley, & Kameka, 1995; Milgram et al., 1999). These results suggest that cognitive functions like **visual or spatial working-memory seem to be similarly age-sensitive in dogs** as it is commonly described in humans and nonhuman primates or rodents.

The decrease in the dogs' performances as a function of the retention intervals both in object permanence tasks and in DNMS or DNMP tasks supports the hypothesis that dogs relied on **mental representations** to encode and memorize spatial or visual information. In 2014, a study demonstrated **deferred imitation** of novel and familiar actions in dogs for the first time (Fugazza & Miklósi, 2014). Deferred imitation is the ability to learn from the observation of the behavior of another in order to later reproduce this behavior; thus to encode, retain and reproduce a demonstrated action after a delay. In their study, dogs accurately reproduced the observed actions with retention intervals ranging from 0.4 to 10 min in different conditions and also if they were engaged in several distractive activities during the time interval (Fugazza & Miklósi, 2014). From a cognitive perspective, imitative behaviors that occur after a minimum delay of 1 min **cannot be considered as contagion, reflexive behaviors or response facilitations** but would rather be the fruit of **representational abilities** (e.g. Zentall, 2006). However, if dogs effectively rely on internal representations to solve these types of tasks, the **underlying mechanism is still unclear**. One could speculate that dogs used a **mental rehearsal** of the information but no one has yet investigated this explanation. Others are reluctant to draw such conclusions based on procedures where animals are repeatedly and actively trained to retrieve their memory traces. According to them, dogs could have developed such specialized and outdoing skills as a mere result of an **extensive training** (Fujita, Morisaki, Takaoka, Maeda, & Hori, 2012). In contrast, the question of whether dogs are able to form an incidental memory of a single experience, also defined as episodic memory, appeared as a more central framework of research for these authors.

4.1.2 *Episodic-like memory*

As briefly introduced above, establishing episodic memory in animals is one of the most **challenging issues** of current research. It requires the animal to **recollect knowledge** about the content of an event ("**what**"), the location where the event took place ("**where**") and details about when it occurred ("**when**"). The subject should also be able to **mentally travel through the event** by recalling details about the source, and about previous or later episodes; thus to precisely **situate the event in a timeframe**. Pioneer foraging tasks have been used to tackle the question of the existence of episodic memory in animals, specifically in **scrub jays** (*Aphelocoma coerulescens*). Using two different types of food rewards, including one that was preferred by scrub jays but time-sensitive (wax worms versus peanuts), these birds selectively foraged for the two rewards based on the length of the time interval since their last visit to the site (Clayton & Dickinson, 1998). Thus, scrub jays appeared able to remember *what* they were searching for, *where* it was located and *when* it was initially cached.

Recently, researchers questioned whether dogs could also demonstrate episodic-like memory. To date, such studies mainly focused on the *what* and *where* components. Kaminski tested two dogs, including Rico, by asking them to retrieve several objects by name from two separate rooms after being given a chance to initially view the location of each object (Kaminski, Fischer, & Call, 2008). Both reliably retrieved the correct objects by their names, thus displaying an ability to remember *what*, but only Rico's search pattern indicated potential integration of information about the object's location, thus the *where*. Although the second dog manifested a more stereotyped search strategy, at least one dog subject provided evidence for remembering **knowledge about both *what* and *where*.**

Other studies focused on the retrieval of memory incidentally formed during a single past experience related to a foraging context (Fujita et al., 2012). Owners were instructed to walk their dogs - all experimentally naïve to the setup - to four baited food containers. Dogs were allowed to eat from two of them chosen at random. After a 10-min walk, dogs "unexpectedly" returned to the experimental room and could freely explore the containers. Interestingly, dogs firstly visited one of the containers where they had not been allowed to eat the treat. To exclude an odor-based choice in the test phase, the containers were all replaced with identical but empty counterparts, but the arrangement of the objects was kept the same. Physical or inadvertent cues from the owner were also discounted to explain the results as the dogs were unleashed as soon as they entered the testing room and were turning their back to their owners.

If the dogs relied on operant learning strategies to solve the task, they would presumably have first visited the containers from which they obtained a reward during the exposure. Conversely, it appeared that such association learning did not take place in this case as they explored significantly more the containers from which they were not allowed eating previously. This search pattern clearly suggests that dogs **retrieved and adaptively utilized knowledge from their past experience** as they correctly remembered from *where* they could collect more rewards.

In a further experiment, 39 new dog subjects had access to four novel containers: two containing food in but only one of them is allowed to be eaten from, one containing an inedible item and one being empty (Fujita et al., 2012) [Figure 66]. After a similar 10-min walk, the dogs returned to the room and searched freely. Thirty out of the 39 dogs first visited the containers that originally contained food (e.g. containers 1 and 3 on the example below). Nineteen of them went to the container that they had initially not been allowed to eat from (e.g. container 1 on the example below) and eleven chose the container where they had already eaten the treat (e.g. container 3 on the example below). According to their results, the authors claimed that these dogs demonstrated a **memory about *what* and *where* previous food treats were located**. In the wild, being endowed

with such episodic memory could increase the animals' ability to successfully collect uneaten food.

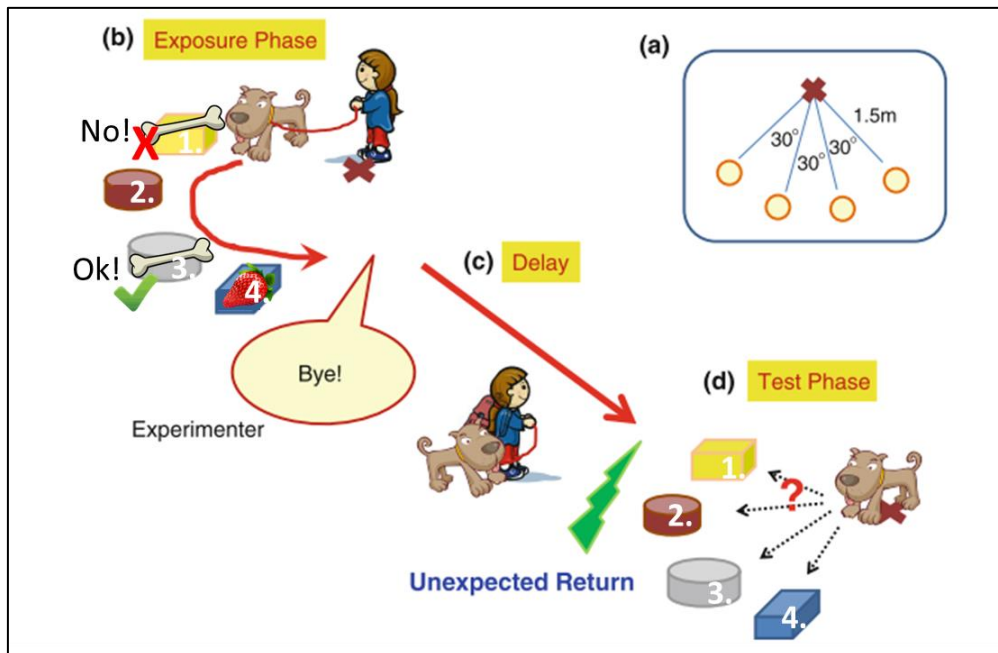


Figure 66 – (a) A schematic top view of the arrangements of the apparatus; (b)-(d) a schematic drawing of the testing procedure. During the exposure phase, two containers contained edible items (e.g. 1 and 3) and the dog was allowed to eat from one (e.g. 3), one container contained a non-eatable food (e.g. 4), the last one was empty. *Adapted from Fujita et al. 2012.*

Despite this evidence for dogs conceptualizing both *what* and *where* components, the defining feature of episodic memory involving **time representation** and “**mental time travel**” in dogs is **sparse**. A study revealed that dogs were affected by the amount of time left home alone, as they demonstrated significantly more greeting behavior (tail wagging, lip licking, body shaking) toward their owners upon reunion after longer separation durations (2 and 4 hours) relative to shorter ones (1/2 hour) (Rehn & Keeling, 2011). Another study indicated that different delay periods affected dogs' willingness to engage in a cooperative exchange task (dogs could choose not to eat a small valued food item in order to exchange it later for a much larger item) (Leonardi, Vick, & Dufour, 2012). The results showed that **dogs anticipated delay duration** as they were less willing to cooperate (to wait in order to make the exchange) if the required time was too long (>40s).

Clearly, research on episodic memory and more particularly “mental time travel” and the conceptualization of *when* by dogs will need more robust investigations but this fruitful area of research is truthfully only at its beginning.

Nevertheless, the **delay between learning and testing** encountered in all those studies ranged from **seconds to minutes** (or to a few hours in one study). Whether dogs are able to remember

knowledge or information about events days, months or years later is not yet well documented. Although this species is believed to remember familiar people after long time periods of absence, or traumatic events months after they happened (e.g. an injection at the veterinary clinic), empirical evidence for long-term memory capacities in dogs remain missing.

4.1.3 *Long-term memory in dogs*

In the past, animal studies have rarely been directly aimed at the problem of **long-term memory**, which has always been considered as a “**higher order mental process**”. To date, only a handful of studies contributed to the actual knowledge about long-term memory in animals. The majority of them tested memory for various abstract concepts in nonhuman primates. In 1973, Johnson and Davis reported that eight rhesus monkeys (*Macaca mulatta*) retained abstract oddity tasks (i.e. discrimination among novel sets of stimuli based on the perceptual “difference”; symbolically: ABB, ABA, BAA, etc.) nearly perfectly when retested **7 years later** on the same task (C. K. Johnson & Davis, 1973). Burdyn et al. documented that three squirrel monkeys (*Samiri sciureus*) trained on an oddity concept also showed evidence of retention on this relational concept over **two years later** (Burdyn, Noble, Shreves, & Thomas, 1984). Finally, gorillas (*Gorilla g. Gorilla*) that had demonstrated a win-stay, lose-shift strategy during training on a series of discrimination reversal problems were posed this same problem again after 2.5 years (T. L. Patterson & Tzeng, 1979). Their performances were comparable to their best prior performances. For these three studies, it was clear that the **problem-solving frameworks established during training were remembered by individuals over very long periods of time**.

The first report of **long-term conceptual memory in a non-primate species** dates back to 2002 (Reichmuth Kastak & Schusterman, 2002). This study reported the case of a sea lion tested for memory for an associative concept (non similarity-based classes of stimuli in a simple discrimination repeated-reversal procedure) **one year after the last practice** and for a relational concept (generalized identity matching) after about ten years. The memory tests revealed no decrement in performance in either of the two problem-solving strategies. Later, another study demonstrated that horses accurately remembered stimuli they were trained on during discrimination and categorization learning that originally occurred between **six years to a decade earlier** (Hanggi & Ingersoll, 2009). These horses also recalled and applied a relative size concept rule that they had learned seven years prior to the test.

Finally, the first evidence that dogs could remember a task after a long time period opened the door of this new and certainly fruitful area of research. In one study that used a touch-screen apparatus, dogs were tested on a **clip art picture discrimination task 6 months after** they

acquired the task (Wallis et al., 2016). A discrimination task consists of a forced choice procedure where two stimuli (in this case colored clip art pictures) are simultaneously displayed on the screen, one being the positively reinforced stimulus (S+) and the other the negative stimulus (S-) [Figure 67]. Over the trials, dog subjects learned to selectively choose S+ and avoid S-. In this experiment, the learning criterion was set at 87.5% correct choices (≥ 28 correct first choices in 32 trials) in five out of seven consecutive sessions.

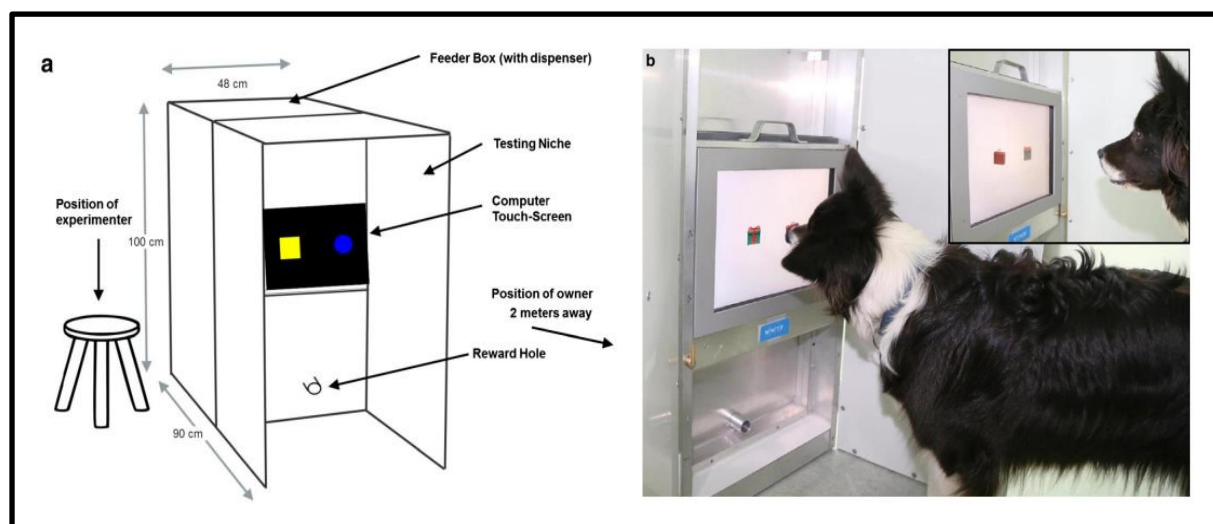


Figure 67- Schematic drawing of the apparatus (a) photograph of a dog working on a visual discrimination task in the testing niche with one side open. From Wallis et al., 2016.

Among the 82 dogs who reached the learning criterion, 46 participated in the memory test after a 6 month interval. Forty-two of them scored significantly higher than chance level (i.e. 0.5) in the first testing session (≥ 22 correct first choice out of the 32 trials). The results revealed dogs' aptitude to discriminate between visual stimuli such as clip arts displayed on a screen, and demonstrated memory skills for **recognizing reinforced stimuli at long term based on their visual properties exclusively** (Wallis et al., 2016).

In another study, laboratory dogs (beagles) were trained on a traditional obedience task (i.e. to go to the basket and stay) by means of operant conditioning and shaping (Demant, Ladewig, Balsby, & Dabelsteen, 2011a). Dogs' retention of the task was tested four weeks after acquisition and revealed a high level of recall irrespective of the frequency of training they were exposed to. The authors suggested that once a task is learned, dogs are likely to remember it for a period of at least one month after the last practice.

Finally, as described in the previous section, at least one dog, Rico, showed retention of **object labels mapped via inferential strategies one month earlier** (Kaminski et al., 2004).

To the best of our knowledge, no other evidence for long-term retention abilities by dogs are currently reported in the scientific literature.

4.2 Canine brain anatomy of the mnesic system

The canine brain recapitulates many important structural aspects of the human brain, including an **over-developed neocortex**, **gyral folding patterns** and a **lobar organization** [Figure 68]. Although brain anatomy in dogs is relatively well described, the **functions underlying each brain region are less clear**. However, as the basic relations between principal structures of the brain remain similar from mouse to human, it can be speculated that dogs' **brain areas that anatomically equate to the human homologous areas share similar functions**. Nevertheless, the brain areas responsible for encoding, consolidating and retrieving information in humans reveal, for a part, structural differences in a dog's brain.

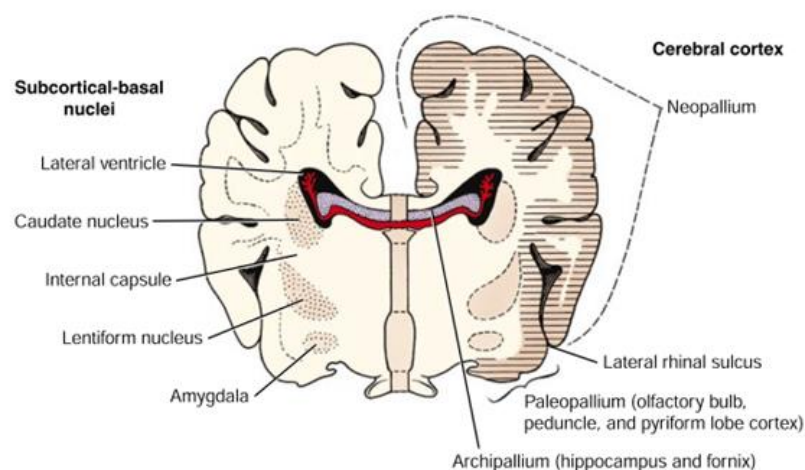


Figure 68 - Drawing of a dog frontal section showing the development of the telencephalon. From *the Veterinary neuroanatomy and clinical neurology*, De Lahunta.

As already described earlier in this manuscript, in humans, the **hippocampus** is one the major components responsible for these cognitive functions. It is part of the **medial temporal lobe in primates**, whereas for **other mammals it is located under the cerebral cortex** (De Lahunta, Glass, & Kent, 2014). Dogs exhibit “**two hippocampi**” per hemisphere, an upper “rodent-like” and lower “human-like” hippocampus [Figure 69].

In dogs, the hippocampus is an **internal gyrus**, an **area of cerebral cortex that has been rolled into the lateral ventricle during development**, and that is not visible on the external surface (archipallium) [Figure 68 & Figure 69]. As with humans, it is located **close to the amygdaloidal body** and belongs to the **limbic system**. In dogs, the hippocampus extends in a curve with a C like shape, starting from the amygdaloid body ventrally in each piriform lobe and progressing caudally

and dorsally and then rostrally over the diencephalon [Figure 69]. It forms part of the medial and dorsal wall of the lateral ventricle ventrally, and part of the medial and ventral wall of the lateral ventricle dorsally (De Lahunta et al., 2014).

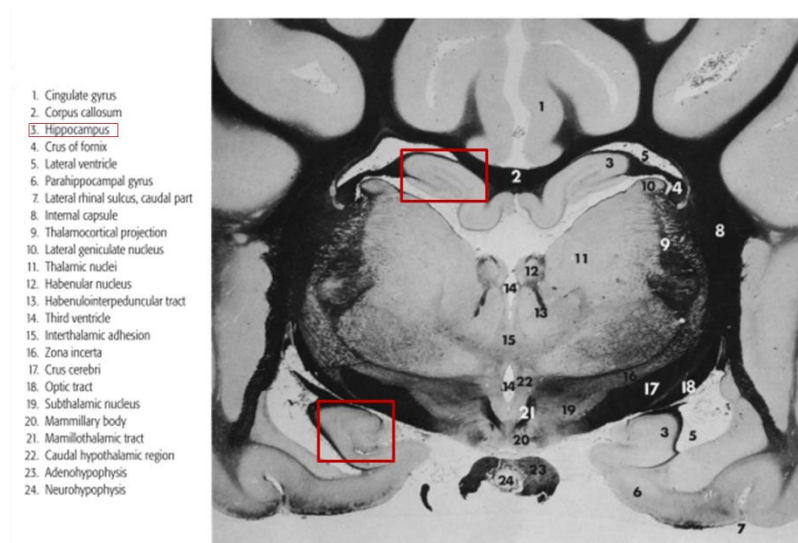


Figure 69 - Transverse brain section from the rostral part of a dog's brain. White matter has been stained and appears black; grey matter is relatively unstained. Adapted from the *Veterinary neuroanatomy and clinical neurology*, De Lahunta.

5 Conclusion of the chapter and presentation of the experimental project

There is considerable current interest to shed light onto complex cognitive functions in nonhuman species. The dog has become an inspiring model of nonhuman cognition research which now spreads beyond the narrow scope of studies conducted on rodents and primates. In contrast to primates, a dog is a much more cooperative model to work with. The thousands of years of cohabitation with humans in addition to a stringent selection of specific physical and behavioral traits, that definite most of the current breeds, conferred to dogs unusual social and unsocial skills in which the scientific community finally showed an interest.

Nowadays, the domestic dog is no longer just the daily life companion of humans. Dogs have become a promising model of research that have already revealed numerous of their remarkable and sometimes unexpected aptitudes over the last years. As detailed in this chapter, in many tasks, dogs have been shown able to outperform our closest primate relatives. Moreover, the unrelated genetic pattern between dogs and humans - and thus the evident disparity between the brain anatomy of dogs and humans - also inspired research on the **evolutionary emergence of some cognitive abilities** shared by both species. Comparative studies are not only specifically designed

to merely compare cognitive aptitudes between species; they sometimes also help us to **understand where the roots from complex cognitive functions might originate.**

In this thesis, we will focus on an apparent communicative ability that some dogs have already been shown able to acquire, namely the **word-referent concept**, in order to **examine long-term memory abilities by dogs.** As developed above, this cognitive function has been poorly documented in dogs. However, it is of major importance to investigate whether learning principles considered as the building blocks for memory formation in humans are similarly shared by unrelated species like the domestic dog. **At a behavioral level, if the conditions that drive memory formation in humans follow similar rules in distantly related animals,** like dogs, this might offer an interesting approach to better understand the origins of the underlying brain mechanisms.

So, the central goal of this thesis is to test the **generality of the repetition-based memory mechanism**, which constitutes the heart of the M4 project as introduced in the first part of the manuscript.

More precisely, in the context of a complex **cross-modal paradigm** (the word-referent concept), we first aimed to know if **dogs would be able to exhibit long-term memories** for such complex information as humans are able to. The impressive cases of Rico and Chaser who are able to learn the names of hundreds of objects suggest that they could remember such complex associations for long periods of time. However, there is no clear evidence of their retention ability for object names after a prolonged intervening period of months or years, without being exposed to the stimuli during the time interval. We will firstly focus on this crucial question in this thesis. Furthermore, we will draw a special attention on the heterogeneity of dogs' cognitive performances on this task, and we will try to find reliable explanations of this phenomenon.

The second goal of this experimental project was intended to **examine the effect of a repetition-based learning process on memory formation.** Would dogs' memory performances be correlated with the number of presentations of the pairings during learning? Moreover, in a **case study** we tested the **minimum number of repetitions** the dog required for storing such bimodal sensory information in memory.

Third, we started to investigate whether the **efficiency of the learning strategies** encountered with young children would be equivalent in dogs. Therefore, we applied learning conditions using either ostensive labeling, fast-mapping or reinforcement learning, similarly to those that children encountered, to examine their effects on dogs' retention of novel object names.

CHAPTER VII.

EMPIRICAL EVIDENCE OF LONG-TERM MEMORY OF WORD-OBJECT PAIRS IN DOGS

General introduction

This project was designed to test whether dogs are able to form **long-term memories** for complex information after a long period of time **without rehearsal**. This project also intended to investigate whether the conditions for long-term memories to be formed and survive after a delay follow similar rules in nonhuman animals to those reported in humans. This would help us to understand to what extent the memorization processes that is already well-defined in humans might be generalizable to other animals. The **domestic dog** appeared as an interesting model of research for this aim firstly because dogs are **genetically distant from humans**, in contrast to nonhuman primates. Secondly, because they are usually **keen to cooperate with humans**. This means they easily agree to perform tasks when asked by humans, in contrast to any other wild or domesticated animals. And finally, because the impressive communicative abilities they developed to understand human signals enable us to **teach them cognitively complex tasks**. Therefore, our long-term memory experiments on dogs involved **complex cross-modal associations**, specifically the names of objects, as previous studies showed that dogs seem able to acquire this word-referent concept. This concept is particularly demanding from a cognitive perspective, because it requires encoding the visual features of the objects, the auditory sound sequences that constitute the labels and the associative link between each specific item and label. The recruited dogs were all naïve about this task. Therefore, the first challenging part of this study was to teach them the word-referent concept, which constituted the **6-month “training phase”** of the experiment. We considered that dogs had reliably acquired the task as soon as they displayed abilities to retrieve an object by its name, in various contexts and among various sets of objects. For rigorous scientific purposes, accuracy was measured through a stringent setup, similar to the one originally developed by Kaminski et al. (2004). Then, we conducted a series of experiments in order to explore the central questions addressed in this project.

In a first experiment, we directly addressed the question of whether dogs are able to form long-term memories for complex information such as cross-modal associative stimuli. Dog participants

underwent a **3-month intervening period** with no visual nor auditory access to the pairings before memory test. Moreover, we focused on the **effect of age on the acquisition and retention** of this cognitively complex task. We attempted to figure out whether age would influence learning and memorizing new information in dogs as it does in humans.

In a second phase, we carried out a series of **short exploratory experiments** on our best experimented dog, which constituted a **case study**. We tried to understand whether a **repetition-based learning strategy** would influence learning and long-term memory formation. If so, what would be the **minimum number of exposures** to an object associated with its label for the dog to reliably remember it after a delay? We also aimed to know if **memory remains stable over time** or if it **declines as a curve of forgetting** after a certain delay. Finally, in this case study we also attempted to highlight whether learning techniques used in humans can also help the acquisition and retention of new knowledge in dogs. For instance, we compared the effect of **ostensive labeling, fast-mapping** and **reinforcement learning** on word learning.

1 FIRST STUDY. *Age effect on the acquisition and retention of a high-order cognitive task in dogs: the word-object pair paradigm*

1.1 Introduction

Previous case studies demonstrated the remarkable aptitude of some dogs to learn hundreds of words referring to particular objects (e.g. Kaminski et al., 2004; Pilley & Reid, 2011). This ability appeared not to be exclusive to Border collies - considered for a long-time as being the cleverest species of dog- since two recent studies have reported that a Yorkshire Terrier and a mongrel dog were also capable of associating words to referents (Griebel & Oller, 2012; Ramos & Ades, 2012). This inspired the idea that perhaps any dog might be able to acquire this complex task. A recent study explored the brain regions for novel word detection in twelve dogs originating from various breeds, who were trained on the word-referent concept (Prichard, Cook, Spivak, Chhibber, & Berns, 2018). By implementing Pilley's protocol (Pilley, 2014; Pilley & Reid, 2011), dogs were trained over a period of 2 to 6 months to map two words to their corresponding referents and were then tested in an fMRI imaging experiment. Authors' learning criterion was set at 80% correct retrievals for at least one object with the other at or above 50%. The number of objects taught to the dogs as well as the learning criteria required by the authors was much lower than the impressive performances of dogs reported in the other studies. It might be argued that higher levels of performance were not necessary for the purpose of this study. But perhaps the dogs had problems performing the task, and were unable to learn more object names, which

would suggest that cross-modal referential understanding is perhaps difficult for dogs to apprehend.

To test this hypothesis, the following experiment examined whether any dog belonging to **any breed** and with any life history would be able to acquire the complex word-referent task. To that aim, we replicated Pilley's protocol on a **heterogeneous group of dogs** and tested their abilities to retrieve objects by name after a **6 month training period**. We expected to observe a disparity in the ability of dogs to learn the task and hypothesized that the variability could partly be attributed to similar parameters than those found in humans. Specifically, we presumed that **age would play a crucial role** in the ability to learn a novel task. In humans, it is assumed that learning is a cognitive process that increases rapidly from infancy to early adulthood and steadily declines during adulthood (e.g. Baltes, 1987). **Age-related learning impairments** are reflected by an increased number of trials to reach a learning criterion as well as increased perseverative responding (e.g. Craik & Salthouse, 2008; Salthouse, 1996). If learning is similarly age-sensitive in dogs, then young dogs should acquire a novel task better and faster than adult dogs, an idea suggested by some previous studies. For instance, relatively old laboratory beagles (>5 years of age) showed impairments to reach the learning criterion of DNMS and DNMP tasks compared to younger beagles (<2 years of age) who were typically able to perform above chance (criterion not reached within 400 trials at a 10-s delay within the group of older dogs) (e.g. Chan et al., 2002; Head, Mehta, Hartley, & Kameka, 1995; Milgram et al., 1999). Similarly, in a discrimination learning task (i.e. a two-choice procedure displaying two stimuli and only one led to a reward), Wallis et al. found a significant effect of age on the number of trials dogs needed to reach criterion (as age increased, discrimination learning ability decreased) (Wallis et al., 2016). Moreover, all those studies reported a high **degree of perseveration in the older dog groups** (e.g. old dogs developed a side preference and/or showed persistent responding to previously rewarded stimuli) which contributed to the deficits to acquire the discrimination learning tasks (e.g. Milgram et al., 1994; Wallis et al., 2016). If these patterns are generalizable to other types of learning, and if age is truly a critical parameter in the acquisition of a novel task in dogs, we expect our younger dogs to demonstrate higher levels of performances in comparison to older dogs.

The second purpose of this study was to benefit from this cognitively demanding task to investigate whether dogs not only learned, but also **stored this knowledge in memory**. Currently, very little is known about long-term memory formation in dogs. Dogs have been shown to be able to **recognize visual stimuli** that were **positively reinforced 6 months prior to the test** in a **discrimination learning task** (Wallis et al., 2016). In the current study, we also addressed this

question by measuring dogs' ability to recognize the objects they were trained on (thus positively reinforced objects) amongst distractors (not positively reinforced).

Furthermore, we aimed to test if dogs also **stored the entire mapping information** between objects and words in memory. Did they only remember the perceptual properties that characterized the target objects (e.g. color, texture, shape, etc.) in order to recognize them among distractors or did they also memorize the sound sequences that referred to the objects during training? We examined this question by conducting **retention tests after a 3 month intervening period** with no rehearsal of the material during this time delay in order to prevent reactivation of the memory. For these tests, we also expected to record inter-individual differences. If age is similarly critical for the consolidation and retrieval processes in dogs as it is in humans, **the older dogs would be less likely to remember the names of the objects** compared to younger dogs. In Wallis et al.'s visual discrimination task, there was no difference in performance in any of their age groups after the 6 month break (Wallis et al., 2016). One explanation is that object discrimination tasks are not difficult enough to highlight age effects. Therefore, we hypothesize that our complex associative design would emphasize the appearance of **age effects**, if age is a responsive parameter to memory formation and maintenance in dogs.

1.2 Materials and methods

1.2.1 *Collaborations*

Since 2016, we have established **two major partnerships** in order to achieve this project. The first involved the **Veterinary School of Toulouse**, namely "l'Ecole Nationale Vétérinaire de Toulouse" (ENVT). Nathalie Priymenko, veterinarian and lecturer at the ENVT, and Elisabeth Jeunesse, research engineer at the ENVT, helped with the recruitment, offered access to the facilities where the experiments took place (neutral rooms or classrooms of the ENVT), and gave appropriate advice about the procedures according to their expertise.

Second, Lucie Negro, a **professional dog trainer** since 2006 and manager of the company "Chien Complice" devoted a lot of time to supervising the training sessions at the Vet School. Together, we developed the training procedure of the protocol, according to the methodologies of previous studies that we had to adjust to each dog. She also gave precious advice to the owners to help them with teaching the task to their dogs at home.

In 2017, we also started a collaboration with the "**Ecole Chiens Guides d'Aveugles**" (ECGA) of Toulouse in order to replicate our setup with dogs of the same breed (i.e. Labradors) and from the same working environment. The aim was first to enlarge the number of individuals in the cohort,

and second to examine whether this relatively homogeneous group of dogs would generate less variability in performances. Unfortunately, 3 months after the beginning of the protocol, we had to put an end to this collaboration, as most of the dogs had to quit the experiment for various reasons as detailed in a later section.

1.2.2 *Ethics Statements*

The study was approved by the **Ethical Committee in Animal Experimentations** “Science et Santé Animales N°115” (SSA N°115) of Toulouse, under the **Number SSA_2016_012**. Owners freely enrolled in this study and gave written consent for their dog’s participation in the experiment.

1.2.3 *Recruitment*

Subjects were pet dogs recruited from owners who volunteered to participate in the project. Owners were predominantly vet students in their first years of studies at the ENVT. Recruitment took place at the ENVT and at the ECGA. Due to the time requirement to supervise each dog-owner couple efficiently, two cohorts of dogs were recruited over two consecutive years: October 2016 – September 2017 and October 2017 – September 2018. The experimental setup as well as the training conditions (i.e. facilities, time period of the year, access to the dog trainer’s advices, etc.) were identical for both cohorts. For both cohorts, recruitment took place at the end of October 2016 and 2017 at the ENVT (and at the end of October 2017 at the ECGA).

Announcements distributed per email firstly invited potentially interested owners to attend a meeting at the ENVT (and at the ECGA). The meeting detailed the objectives, procedure, inclusion criteria and required investment from owners to achieve the task during the following year. At the end of the meeting, owners who enrolled in the study were later met individually with their pet dogs. During this subsequent session, each dog’s behavior was scrupulously evaluated according to Mrs. Negro’s behavioral evaluation grid. It consisted of measuring dogs’ general obedience, natural interest in playing and fetching objects, ease of being manipulated by unknown individuals (e.g. experimenters), absence of signs of aggressiveness, etc. Only dogs who fulfilled at least 15 out of the 21 behavioral points were allowed to participate in the protocol.

1.2.4 *Subjects*

Subjects were pet dogs recruited at the ENVT and ECGA of Toulouse over two consecutive years. Dogs had to fulfill several inclusion criteria to integrate the study:

- (1) All types of breeds were admitted except dogs from the 1st and 2nd category as defined by the French Rural Code, art. 211-1 of the Law 99-5 of 06/01/199 concerning dangerous and stray animals (i.e. “attack dogs” such as for example Staffordshire terrier, American Staffordshire terrier, Mastiff, Tosa types, etc.; and guard dogs and molossoids such as Rottweiler, for example).
- (2) Both genders were included except gestate females. Non-sterilized females could not participate in the weekly working sessions when they were in heat.
- (3) There was no age limit to participate in the study. However, dogs above 7 years old were only included if they had undergone a recent medical examination including a blood test with no suspect issue.
- (4) All dogs had to be identified (chip or tattoo) and vaccinated against the Kennel Cough, Canine distemper (also called Carre’s disease or hard pad disease), Infectious Canine Hepatitis (ICH), Leptospirosis, Canine Parvovirus and rabies.
- (5) Dogs had to fulfill at least 15 out of the 21 behavioral points measured by Mrs. Negro.

Forty companion dogs of owners who volunteered to participate in the experiment respected these criteria (ENVF, N=35; ECGA, N=5). The **experiment lasted about one year** and required owners to **train their dog at home three times a day and half an hour a week under the supervision of the experimenters**. Due to a lack of time to satisfy these requirements, several owners abandoned the study before the end of the training phase (N=9). One owner moved from Toulouse with his dog during the time period of the study (N=1), another interrupted the study because his dog became suddenly critically ill (N=1) and a few abandoned for unknown reasons (N=3). The five dogs recruited at the ECGA had also to leave the study as they were removed from the ECGA because of major behavioral or health troubles. Altogether, from the 40 dogs originally recruited, only **21 performed the final comprehension and retention tests**. Therefore, the following data only involved these dogs.

The 21 remaining dogs came from various breeds, as detailed in [Table 8](#). Except for one dog (Yuki), all dogs were “**medium sized breeds**” that had relatively similar life expectancies and developmental periods.

When the training phase began, dogs ranged in age from **3 months to 8 years** (see [Table 8](#)). The dogs were split into 2 age groups, namely *puppies* and *adults*, taking into consideration the main developmental periods of “medium sized dogs”. The cutoff was established at one year old and was partly based on sexual maturity which stands at around one year of age for most of the recruited breeds.

Finally, all of the dogs were experimentally naïve to the task prior to the experiment. None of them had previously learned words to refer to specific items. However, most of them knew words

for specific actions (e.g. “sit”, “lie down”, “wait”). Furthermore, some of the dogs evolved in various environmental conditions and were exposed to various enrichments (e.g. agility, cani-cross, dog training, etc.) that could not be controlled in this experiment.

Table 8 - Table indicating participants’ ID number, name, age, age group, breed, sex and number of objects pitted against each other on floor during testing.

N° ID	Name	Age (months)	Age group	Breed	Sex	Number of objects on floor
1	Fenrir	14	Adult	Husky	M	4
2	Aleombre	12	Puppy	German x Tervueren shepherd	M	5
3	Miko	12	Puppy	Korthal	M	5
4	Alaska	12	Puppy	Australian shepherd	F	6
5	Oreo	3	Puppy	Border collie	M	5
6	Nova	5	Puppy	Border collie x malamute	F	4
7	Neurone	7	Puppy	Beagle	M	6
8	Nami	9	Puppy	German shepherd	F	4
9	Mango	12	Puppy	Border collie x white shepherd	M	5
10	Nayla	8	Puppy	Golden retriever	F	5
11	Nausicäa	18	Adult	Border collie x Husky	F	3 + 4
12	Jeika	24	Adult	Border collie x Bernese mountain	F	3 + 4
13	Harley	96	Adult	Labrador x hunting dog	F	4 + 4
14	Happy	54	Adult	Labrador	F	4 + 4
15	Tanga	72	Adult	Mix Shepherd	F	3 + 4
16	Hyuri	54	Adult	Springer spaniel x Border collie	M	2 + 3
17	Yuki	30	Adult	Boston terrier	M	4
18	Alba	11	Puppy	Mix Shepherd	F	4
19	Moka	6	Puppy	Hunting dog	M	8
20	Glee	60	Adult	Australian shepherd	F	5
21	Elis	84	Adult	Border collie	M	3 + 4

1.2.5 Stimuli

Stimuli consisted of toys for children (soft toys) or for dogs (e.g. plastic toys, stuffed animals, knotted ropes, rubber toys, squeaked toys, etc.) as illustrated in **Figure 70**. The objects differed in size, weight, shape, texture, color, design and material. Despite some similarities, each object was characterized by unique features enabling discrimination. For each dog, there were no duplicates of the objects.

A distinct proper name was attributed to each object. Labels were **pseudo-words** composed of two identical syllables (e.g. “lili”, “dudu”, etc.), two phonologically distinct syllables (e.g.

“rivou”, “peno”, “tuda”, etc.), or consonant-vowel-consonant constructions (e.g. “BUX”, “ZAV”, “PIM”, etc.) [Figure 70]. As with young children, the use of pseudo-words ensured that dogs had no auditory access to these stimuli during the time interval that separated learning from memory tests.



Figure 70 - Photographs of some objects with their associated labels, belonging to the dog Moka.

1.2.6 Procedure

The experimental setup first consisted of a **6-month training period** since the recruited dogs were naïve about the concept that words can refer to items. First of all, owners attended a meeting that explained the theoretical framework of the training procedure they had to apply at home. Then, **three times a day** (3 x 5min) they taught the word-referent concept to their dogs by scrupulously following the protocol. Short training sessions spaced in time were shown to be more efficient for dogs to acquire a task than working sessions performed in a row (Demant et al., 2011a). Owners were asked to record in detail each of their working sessions in an experimental book provided by the experimenter. Owners had to transcribe the schedule and duration of their sessions, the objects they worked with, any issue they encountered and the fetching scores obtained by their dogs. Moreover, owners met the experimenter and the dog trainer **once a week at the ENVT**. It consisted of **individual and personalized training sessions**. A session lasted about 30 minutes and was essential to track the dogs’ progression [Figure 71]. During this working-session, owners were asked to reproduce the exercises they performed at home in front of the experimenter and the dog trainer. Both could appraise each dog’s progression and often had to adjust the original protocol to each dog according to the difficulties encountered. Sometimes, the experimenter or dog trainer took advantage of this session to train the dogs themselves which also constituted an opportunity for the owner to clarify the teaching procedure. Owners were then given **personal advice** and received an **individualized training program** for the following week. The entire training protocol and its main steps are detailed in the later section.

Exactly 6 months after the beginning of the training phase, the dog’s ability to retrieve objects by name was evaluated using a **tightly controlled procedure**. The **testing** occurred at the ENVT and

included **two to four consecutive sessions**, in order to reflect in the most appropriate way the **dog's comprehension of the task**. Indeed, dogs (and more largely animals) are very sensitive to environmental changes that can easily disrupt their motivation to engage in a task or their level of concentration (e.g. physical environment such as noise in the corridor, unusual odors in the experimental room, etc.). Internal state might also account in the performances (e.g. physiological status or mood (being tired, bored about the task, too excited), etc.). These external and internal parameters have to be considered when evaluating an animal's performances. Therefore, comprehension of the task was measured as the mean score obtained over two to four test sessions performed at the end of the 6 month training period. Five dogs could only be evaluated once because their owners were unavailable.

Following this comprehension test, dogs underwent a **3-month period** during which they had **neither visual access to the objects nor auditory exposure to the pseudo-words** [Figure 71]. Retrieving objects by name was not rehearsed with other items during this time interval either. Furthermore, dogs had **no access to the experimental room**, which could have reminded them of the task.

After the 3 month intervening period, dogs were first tested on a **discrimination task**. It consisted of **recognizing the objects** that they were trained with amongst a set of distractor objects (i.e. dogs' unlabeled toys) that dogs had not seen for the same amount of time. During this discrimination test, target objects were not labeled. Next, **retention of the word-object associations** was measured using the same **tightly controlled procedure** as for the comprehension tests [Figure 71]. The memory test included a single testing session.

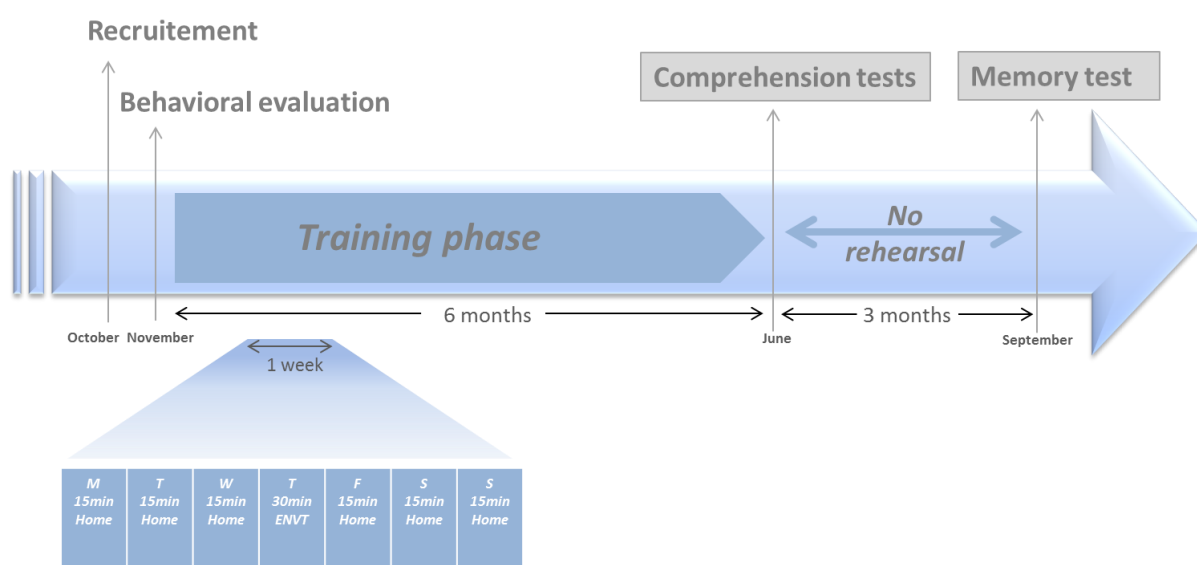


Figure 71 - Schematic of the experimental setup for evaluating dogs' memory performances to remember object names after a 3 month intervening period. The design was conducted during two consecutive years, starting from October 2016 and October 2017.

1.2.6.1 *Word-referent training procedure*

During a 6 month period, dogs were taught object names and trained to retrieve items by their corresponding labels. The training took place in familiar environments (e.g. at home) and with well-known individuals (e.g. the owner or owner's partner) to make dogs feel confident about the task. Most of the training procedure implemented Pilley's protocol that he established for Chaser (Pilley, 2014; Pilley & Reid, 2011). As Pilley did with Chaser, owners dedicated a great part of the training sessions playing tug and fetch with their dog and the objects while verbally reinforcing the names of those objects. Next, the word-referent concept per se was mostly taught by means of **classical** and **operant conditioning** associated with **positive reinforcement**. **Rewards were play, praise and/or food**, depending on each dog. For example, excited dogs were typically provided a neutral food pellet instead of play which would have been too stimulating for them. In contrast, dogs that mainly focussed on the food reward rather than on the task were reinforced with praise and play (see Prichard, Cook, Spivak, Chhibber, & Berns, 2018 for equivalent methodology). Moreover, dogs were taught the **clicker training method** since the click provides a more **neutral signal of accuracy** than a voice signal. This classical conditioning method consists in clicking each desired or accurate behavior that will be subsequently reinforced by a reward. Thus, from a dog's perspective, each click predicts a reward that will be delivered in a short time window. The clicker is an efficient method as it enables the capture of desired behaviors precisely when they occur which helps subjects to rapidly comprehend what the targeted behavior consisted of.

The first object was introduced to the dog, who was allowed to freely play with it, while its owner verbally repeated the pseudo-word allocated to this object. The owner also held and pointed to the object while labelling it. These playing while labelling sessions lasted three days. Then, the object was placed at a distance of 4-5m and the owner instructed the dog to "go get [*object*]" or "where is [*object*]" or simply "[*object*]". No other objects were available on the floor for retrieval, so errors were unlikely. As soon as the dog grasped the object in his mouth, the owner reinforced that behavior by a click and the associated reward as described above. Then, the dog was progressively rewarded for fetching and bringing the object to his owner's feet rather than simply grasping it. After retrieval, owners provided several rehearsals with the associated label.

Following this initial training in the absence of other objects, the newly introduced object was placed on the floor among dogs' familiar toys that had never been labeled by the owner before. The dog was instructed to retrieve the newly learned object exactly in the same manner as when displayed alone. The dog was equally first rewarded for grasping the target object and as soon as it was confident with this novel configuration of the task, it was required to bring the object to the

owner's feet in order to get the reward. When the dog was 100% accurate on this task, the second object was introduced following the same procedure. At this step, the second object was always presented isolated from the first labeled item. As soon as the dog was equally accurate in retrieving this second object among familiar unlabeled toys, the owner alternated between both introduced objects every succeeding training session.

Next, both newly learned objects were pitted against each other among familiar toys serving as distractors. The owner randomly requested one of the objects by uttering the same instruction as previously (e.g. "go get [*object1*]"). The dog was rewarded only for correctly selecting the target object if it was his first selection. Otherwise, if the dog selected the wrong object, the owner remained neutral, made neither remark nor gave punishment, recalled his dog and gave him another opportunity to fetch the requested item. If the dog similarly failed to retrieve the object upon command on the second or third trial, the owner removed the other objects and gave his dog additional training on an easier configuration to reinforce the word-object mapping. Owners had to always remain attentive and ready to rehearse easier exercises in order not to lose their dog's interest and motivation to engage in the task. When the dog successfully retrieved the target object placed beside the other newly learned object, it was removed and the dog was requested to fetch the second item upon command. This **procedure without replacement of the objects** facilitated the dog's comprehension of the concept. Indeed, since an object was correctly retrieved, the probability for dogs to retrieve this reinforced object in the consecutive trial is high. To minimize errors, the second trial of the sequence only involved the remaining learned object among distractor toys. Objects were then rearranged to limit learning by position and another sequence could start. The order of the targets was also randomly counterbalanced to avoid learning a sequence. A training session always ended with a correct performance. Dogs were also always given the opportunity to play with the objects at the end of the session as objects were not available to dogs aside from in the training sessions.

Novel objects were only introduced when few errors were made on the fetching tests of learned objects. Thus, each dog could progress at their own pace. The same procedure of teaching the names of novel objects was applied for each newly introduced object. Familiar toys serving as distractors were removed as soon as the dog knew 4 or 5 object names.

Complementary exercises were established in accordance with the dog trainer and were especially designed for dogs who had difficulties in acquiring the task. For many dogs, the sole act of interacting with an object constituted a source of reward whether or not the object was the target. Thus, some dogs preferred interacting with their favorite object rather than fetching the requested one, even though it would not have brought a reward afterwards. To avoid this behavior

happening, two complementary exercises were designed. The first consisted of holding two newly learned objects in the hands - one in each hand - facing the dog who sat in front of his owner about 1m apart [Figure 72]. The owner looked straight away and randomly instructed his dog to fetch one object (e.g. “where is [*object2*]”). If the dog selected the target object at first, he could interact with the object and get the reward. Instead, if he went to the wrong object first, the owner rapidly closed both hands thereby not allowing his dog to interact with the object. This method helped the dogs make fewer errors, in addition to reinforcing correct retrievals.

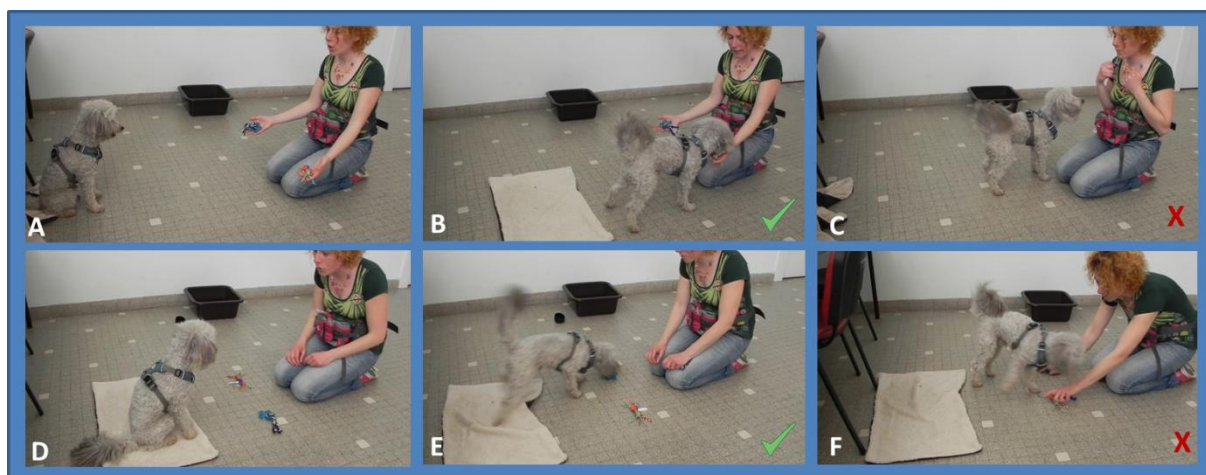


Figure 72 - Photographs illustrating the complementary exercises designed to help dogs in comprehending the word-referent concept. The dog stands about 1m apart from his owner waiting for the fetching request (panel A & D). The dog correctly selects the target object at first and is allowed to interact with (B & E). The dog selects the wrong object at first and is not allowed to interact with (C & F). One exercise consists of holding the objects in the hands (A, B & C), while for the second exercise the owner displays the objects in front their knees (D, E & F). Pictures are photographs from the dog trainer and her dog (not included in the sample).

As unconscious cues might have been utilized by dogs to solve the task, a second exercise consisted of displaying the objects in front of the owner’s knees. Similarly, the owner could easily prevent his dog from making an error and from getting self-rewarded [Figure 72]. Nevertheless, inadvertent cues might also have accounted for the dog’s accuracy in this configuration. To ensure dogs accurately comprehended the task and to control for any “Clever Hans effects”, a black panel was introduced about 3.5 months after training began and that separated the dog and its owner from the objects. Thus, dogs also became accustomed to the stringent testing procedure that occurred at the end of the 6-month training period as well as after the 3-month intervening period.

1.2.6.2 Word-referent testing procedure

At the end of the 6 months of training, dogs’ were tested on their ability to discriminate between the trained objects upon command. Performance was measured as a **score of correct retrievals**. The testing procedure also followed Pilley’s methodology (Pilley, 2014; Pilley & Reid, 2011).

The experimenter randomly displayed the objects behind an **opaque panel** out of view of both the owner and the dog. The owner was thus **naïve about the distribution of the objects**. Only newly learned objects were pitted against each other. There were no familiar unlabeled objects among them. For a few dogs, the set of trained objects was divided into two groups: a group that contained only the favorite objects and a group composed of the neglected objects. For these dogs, when objects of both groups were placed beside each other, they always fetched the preferred object and ignored the others whatever the label was. This behavior merely reflects fussiness and an absence of application of the task. In order not to exclude these dogs from the sample and for them to still get a chance to demonstrate their acquisition of the task, they were tested on their favorite and less preferred objects separately.

The dog was positioned beside his owner and both were facing the panel [Figure 73]. Both had no visual access to the randomly displayed objects. This rigorous experimental control ensured that dogs selected objects exclusively on the basis of their verbalized names as they **could not rely on inadvertent cues** provided by the owner. For the same reason, the experimenters turned their back away from the objects and did not look at the dog [Figure 73]. They stood about 1m apart from the dog plus owner, and launched the computer program. A **self-developed program** generated through Python Software randomly designated a target object for each trial. Each object was only tested once during the session. For each trial, the experimenter wrote the name of the elected object on a white board readable by the owner who verbally instructed his dog to fetch that object (e.g. “go get [*object4*]” or simply “[*object4*]”). Only the object that the dog brought from behind the panel counted as a correct or wrong answer. The objects the dog potentially pointed to or grasped in his mouth without bringing them back to the owner’s feet were not taken into consideration in the analyses. A camera stabilized on a tripod videotaped each testing session.

The dog got feedback for his response which was either a click associated with a reward when he accurately retrieved an object (i.e. praise, play and/or food) or a neutral remark when he made a mistake (e.g. “let’s try again”). When he failed on one trial, the dog was given another opportunity to select the correct item, although this second trial did not account in the analyses. The trial was considered as successful or incorrect according to the first retrieval. Since the testing procedure was designed **without replacement**, objects correctly retrieved were not replaced after selection (see Pilley & Reid, 2011 for equivalent methodology). The objects were not rearranged from one trial to another during the entire session. The dog was instructed to fetch all objects until the last one which was not taken into account in the analysis since the dog had 100% chance of retrieving it without error (see Kaminski et al., 2008 for equivalent methodology).

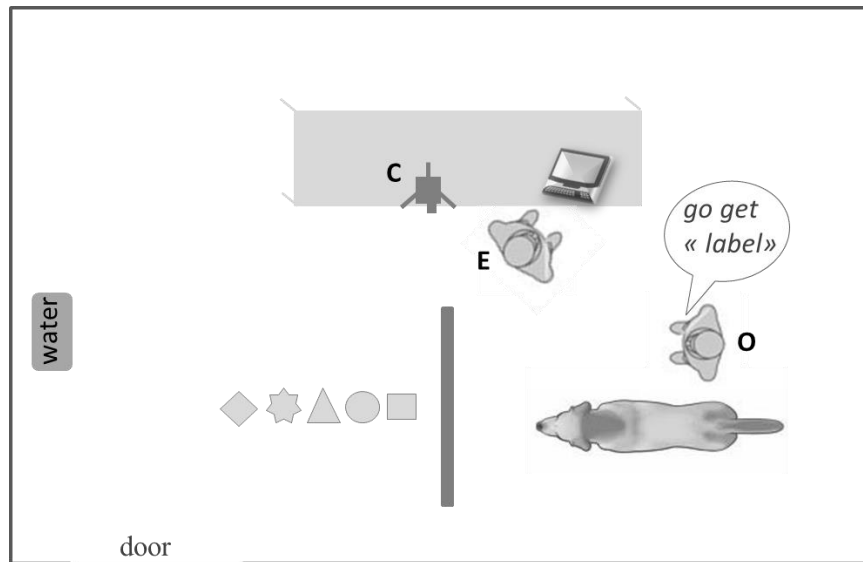


Figure 73 - Illustration of the experimental set up during testing. The dog and his owner (O) were placed about 1m behind an opaque panel that separated them from the objects. For each trial, the experimenter (E) launched the computer program that randomly generated the name of one object. E transcribed the elected name on a whiteboard readable for O. O instructed his dog to fetch the targeted object by verbal command. The dog had to bypass the panel to access the objects. The object brought by the dog was considered as the dog's selection, and was counted either as a success (correct retrieval) or as a mistake (incorrect retrieval). A camera (C) fixated on a tripod videotaped the entire testing session.

1.2.7 *Analyses, modelling and statistics*

Performance was calculated using the **mean proportion of correct retrievals** discarding the last trial (100% chance of being correct). For dogs who performed more than one test - for instance when comprehension tests were repeated or when the set of objects was divided into two distinct subgroups as explained above – performance corresponded to the mean score obtained across those multiple tests. The calculation of a dog's mean performance of an entire testing session followed this formula:

$$\text{Mean Performance} = \frac{\sum_{i=1}^{N-1} (x_i) = x_1 + x_2 + \dots + x_{N-1}}{N-1}$$

Legend. N : number of objects; i : trials; x : subject's responses (0=error or 1=success)

In order to determine whether or not dogs acquired the task, performances were compared to a **theoretical probability to complete the task by chance**. Since the tests were carried out **without replacement of the objects**, the **probability of success** (i.e. the probability of correctly selecting the target object by chance) varied for each trial, following a **hypergeometric distribution**. We simulated this distribution using a **Monte-Carlo algorithm** (see Grassmann, Kaminski, & Tomasello, 2012 for equivalent methodology). It consisted of modelling random object selection

for each trial by implementing different chance values resulting from not replacing the objects after each selection. Thus, in this model, we implemented the **deterministic probability of success by chance for each trial according to the number of objects available on the floor**. The model considered N number of objects available on the ground, $N-1$ number of trials (since the last was not taken into account), and x the response for each trial (i.e. 0=error or 1=success according to the binomial law).

The following formula summarizes the theoretical mean probability of success of the entire experiment according to the initial number of objects available for retrieval:

$$\text{Mean } P_{\text{succes}} = \frac{\sum_{i=0}^{N-1} \left(\frac{1}{N-i} \right)}{N-1} = \frac{1}{N} + \frac{1}{N-1} + \frac{1}{N-2} + \dots + \frac{1}{N-(N-1)}$$

Legend. N : number or objects; i : trials

The simulation of the Monte-Carlo algorithm was repeated a large number of times (2,000 simulations) to generate the **theoretical distribution of the performances due to chance**. To facilitate interpretation, the simulation could be represented by a **curve of success expected by chance according to the number of competing objects** (see results section).

Next, the level of performance of trained dogs was compared to the level of the simulated distributions (i.e. chance level) as calculated above, using adapted **binomial tests** that were **revised in accordance with a hypergeometric law** (see Griebel & Oller, 2012 for the use of equivalent adapted binomial tests):

$$\text{Binom.test} \left(\sum_{i=1}^{N-1} (x_i), N-1, \text{Mean } P_{\text{succes}} \right)$$

Legend. N : number or objects; i : trials; x : subject's responses (0=error or 1=success)

The first argument of such a binomial test corresponds to the **number of successes observed across the experiment** (i.e. number of correct retrievals); the second reflects the **number of possible successes of the entire experiment** (i.e. the number of trials) and the last argument indicates the **mean probability of success by chance**. The last argument is the one we adapted to our design since it was carried out without replacement. This adapted version of the binomial test was used to assess comprehension and retention of the task at a group level and also at an individual level.

When comparison between subgroups was required (e.g. females vs males, puppies vs adults), **generalized linear mixed-models** (GLMM) were conducted. These models consider **repeated measures on individuals as random effects**, **performances as binomial data** (which had to be specified in the model) and the **variable(s)** to investigate as the **fixed effect(s)**.

1.3 Results

1.3.1 *Comprehension of the word-referent concept by dogs*

In order to determine whether dogs were reliably able to associate words to referents, performance was tested at the end of the 6-months training period during **tightly controlled comprehension tests** as described above.

For each dog, one to four test sessions took place according to owners' availability (1 test-session: N=5; 2 test-sessions: N=9; 3 test-sessions: N=3; 4 test-sessions: N=4). For dogs who performed more than one test, results from each test session were considered. Moreover, dogs acquired a heterogeneous number of object names during the training period ($M= 5.66$; $SD= 1.42$; range: [4-8]). Learned objects were pitted against each other during testing, except for few dogs for whom two subgroups of objects were tested separately (cf. **Table 8** in the methods). In this case, results obtained for each subgroup of objects were conserved.

Testing consisted of retrieving objects from behind a black panel upon verbal commands. For each trial, the object brought by the dog was scored as a correct (i.e. 1) or mistaken (i.e. 0) response. Since objects were not replaced after retrieval, the probability to retrieve the target item by chance increased from one trial to another following a **hypergeometric distribution**. Thus, **adapted binomial tests** were used to **compare dogs' performances to levels expected by chance according to a fluctuating number of objects pitted against each other**. Results showed that **dogs, as a whole, performed significantly greater than chance** (binomial test, $p<0.001$) [**Table 9**]. Although binomial tests had to be adjusted to the design, it remained the most conservative statistical tool as it preserved the performance of each trial of each individual. **One-tailed univariate t-test against chance** - which gives a less refined statistical outline of the results as it is based on the mean performance – indicated the same level of significance of the results ($t=3.75$; $p=0.0012$).

Table 9 - Table indicating the number of subjects, total number of trials performed, number of overall correct retrievals, mean expected chance level (in %), p-values from adapted binomial tests, mean performance and CI95% allocated to each group.

Age groups	N subjects	N trials	N successes	Mean expected chance level%	P-values	Mean performances%	CI95%
All	21	171	86	33	$4.11e^{-6}$	50	43-58
Puppies	10	99	55	30	$9.85e^{-8}$	56	46-66
Adults	11	72	31	37	0.32	43	31-55

To go further, a **Monte-Carlo algorithm** simulated the experiment **2,000 times** in order to visualize the distribution of the performances that would have been expected by chance, according to the number of objects available for retrieval at the first trial. As shown in **Figure 74**, **the majority of trained dogs appeared to have performed above the chance levels generated by the algorithm**. Nevertheless, when the performance of each dog individually was compared to its corresponding mean expected chance level, only a small number of dogs displayed significant results (binomial tests; dog N°4, $p<0.05^*$; N°7, $p<0.05^*$; N°10, $p<0.05^*$, N°11, $p<0.05^*$ and N°19, $p<0.0001^{***}$). Four of these five well-performing dogs were puppies. Moreover, only one dog demonstrated knowledge of eight object names and was highly accurate to retrieve these objects by their names (dog N°19). Interestingly, this dog was the third youngest dog of the sample when training began.

Thus, we observed that the notable between-subjects heterogeneity to acquire the task is not only reflected by accuracy but also by the number of object names dogs were able to learn in 6 months. Further analyses tried to shed light on this variability. To test whether age affected the ability of dogs to learn object names, the sample was divided into **two age categories**: *puppies* who comprised dogs below one year old by the time they enrolled on the experiment and *adults* composed of dogs above one year old. The one year old cutoff corresponded to the median of the sample.

For each age group, performance was first compared to levels expected by chance using similar adapted binomial tests. Results showed that *puppies performed significantly above chance* (binomial test, $p<0.0001$) [**Table 9** & **Figure 74**]. In contrast, the performance of *adults as a group was not significantly different from chance level* (binomial test, $p=0.32$) [**Table 9** & **Figure 74**]. To ensure that the significant result demonstrated in puppies was not merely due to a higher number of data available – which might have increased the power of the statistics – since they performed more trials overall, an algorithm randomly selected the same number of trials amongst all as performed by the adults. Each time the algorithm was run it revealed that the mean performance was still highly significant.

Second, heterogeneity was also evidenced by the difficulty of the task, since dogs did not have the same number of competing objects at the first trial. Comparison between *puppies* and *adults* revealed that **puppies had on average significantly more competing objects at the first trial compared to adults** ($M_{puppies}= 5.3$ objects; $SD_{puppies}= 1.16$; $M_{adults}= 3.7$ objects; $SD_{adults}= 0.6$; Kruskal-wallis, $X^2 = 10.99$, $p<0.001$).

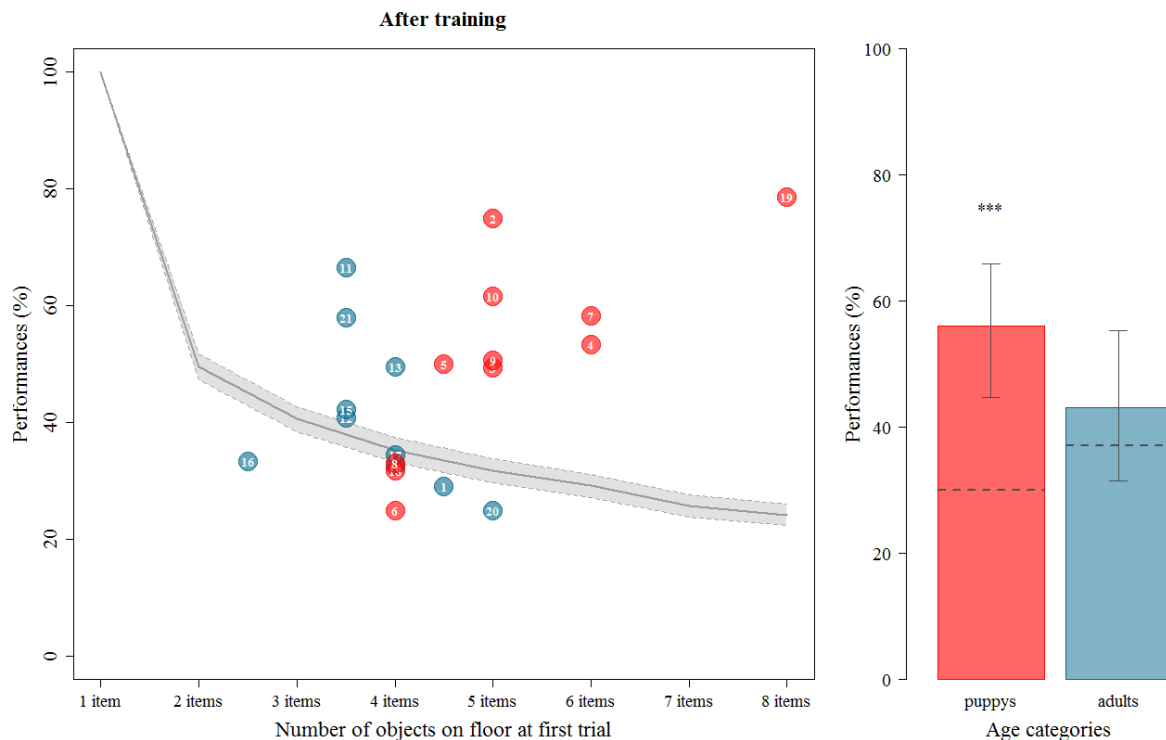


Figure 74 - The left graph represents the mean performance of dog participants to correctly retrieve objects by name during the comprehension tests performed 6 months after training began, according to the number of objects pitted against each other at the first trial. Circles depict the mean performance of each individual dog identified by its ID number. Puppies are represented in red (red circles), and adults in blue (blue circles). The mean probability of success expected by chance according to the initial number of objects available on floor is indicated by the gray line (mean of 2,000 repetitions of a Monte-Carlo algorithm). Dashed gray lines indicate the upper and lower limits of the standard error of the dispersion. The plot on the right recapitulates the mean performances of puppies (red) and adults (blue) as distinct age groups. Dashed lines represent the chance levels (calculated as the mean chance levels according to the numbers of objects on floor at each trial for each age group separately). Standard errors are indicated. Stars indicate significant differences against chance (***, $p < 0.0001$).

In order to assess whether performance was age-category dependent, a **generalized linear mixed-model** was conducted including age group (categorical variable: *puppies* vs. *adults*) and difficulty of the task (i.e. number of objects on the floor at first trial (continuous variable)) as fixed effects, repeated measures on individuals as random effects, and accuracy as binomial data. The model revealed a **significant effect of age on the performances; puppies showing significantly higher rates of accuracy** (GLMM, $Z_{age_categ} = 2.31$, $p = 0.02^*$). There was no significant interaction between the “age-category” and “number of object on floor” variables.

Finally, no sex effect was established (GLMM, $Z = 1.15$, $p = 0.2$).

1.3.2 Long-term retention of object names by dogs

Dogs came back to the experimental room 3 months after last practice and were first tested on their ability to **discriminate the target objects**. This involved displaying each of the objects they had been trained with during 6 months behind the opaque panel beside 2 or 3 distractors (i.e.

familiar objects (dog's toys) that had never been associated with a label). As for the target objects, dogs had no access to the distractors during the intervening period, to exclude choices being simply based on “novelty”. For each trial, dogs were instructed to “go get” and no label followed the command, nor was the label uttered after retrieval. They were as many trials as the number of objects dogs acquired during training. Each trial involved different distractors. Dogs were **highly accurate in discriminating the target objects** ($M=95.1\%$, $SD=10.7$, binomial test, $p<0.0001^{***}$).

Table 10 - Table indicating the number of subjects, total number of trials performed, number of overall correct retrievals, mean expected chance level (in %), p-values from adapted binomial tests, mean performances and CI95% allocated to each group.

Age groups	N subjects	N trials	N successes	Mean expected chance level%	P-values	Mean performances%	CI95%
All	21	98	48	34	0.002	49	39-59
Puppies	10	49	28	30	0.0001	57	42-71
Adults	11	49	20	37	0.65	40	27-56

Next, **retention of the associations** between objects and labels was examined. Performance was measured by the mean proportion of correct retrievals compared to levels expected by chance as for the comprehension tests. Results demonstrated that **dogs as a whole performed significantly above chance** (binomial test, $p=0.0018^{**}$) [Table 10 & Figure 75]. At an individual level, however, only one dog, **a puppy, was significantly more accurate than chance** (dog N°19, binomial test, $p<0.001^{**}$).

Accuracy was also examined separately for each age group. **Puppies showed significant retention** of the object names (binomial test, $p<0.001^{**}$) whereas **adults were at chance** ($p=0.65$). A GLMM that included age and number of competing objects at first trial as fixed effects revealed that **puppies had significantly better retention scores than adults** (GLMM, $Z_{age_categ}= 2.25$, $p=0.024^{*}$).

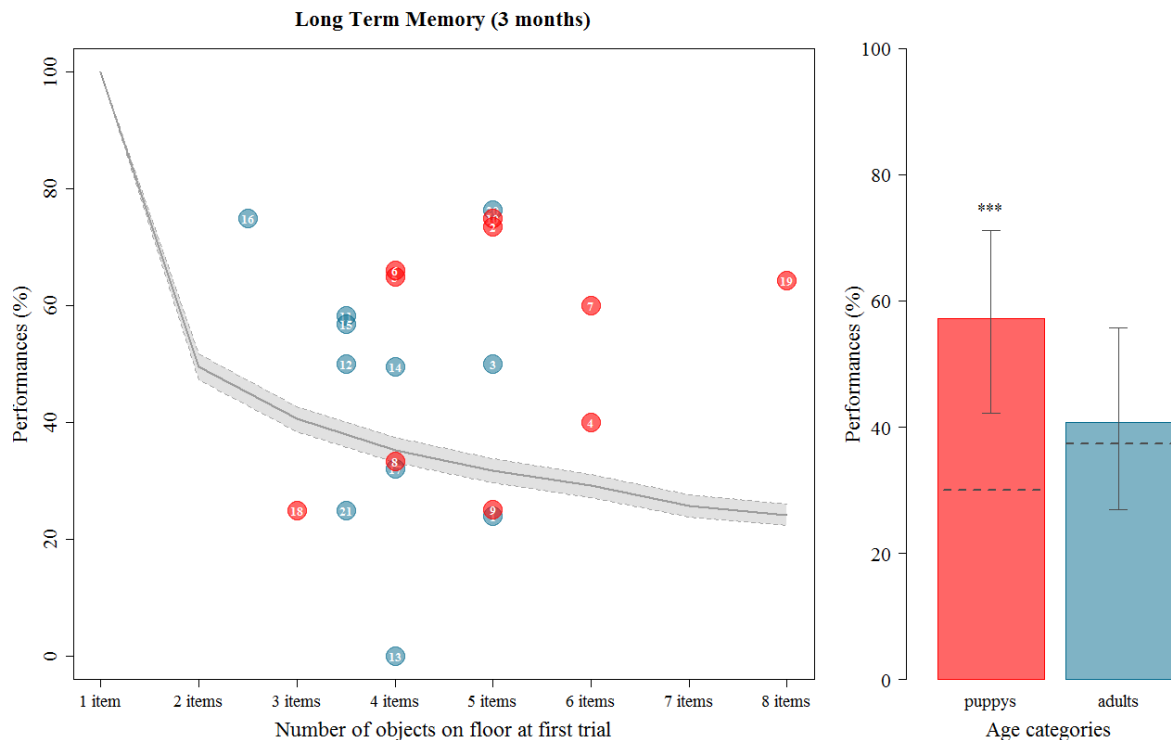


Figure 75 - The left graph represents the mean performance of dog participants to remember the associations between words and objects 3 months after the last comprehension test, according to the initial number of competing objects. Circles depict the mean performance of each individual dog identified by its ID number. Puppies are represented in red (red circles), and adults in blue (blue circles). The mean probability of success expected by chance according to the initial number of objects available on the floor is indicated by the gray line (mean of 2,000 repetitions of a Monte-Carlo algorithm). Dashed gray lines indicate the upper and lower limits of the standard error of the dispersion. The plot on the right recapitulates the mean performance of puppies (red) and adults (blue) as distinct age groups. Dashed lines represent the chance levels (calculated as the mean chance levels according to the numbers of objects on the floor at each trial for each age group separately). Standard errors are indicated. Stars indicate significant differences against chance (***, $p < 0.0001$).

1.4 Discussion

This study first intended to investigate whether dogs originating from various breeds, ages and life histories would be equally able to acquire the cognitive concept that words can refer to objects, as demonstrated by a few dogs previously (e.g. Griebel & Oller, 2012; Kaminski et al., 2004; Pilley & Reid, 2011). Since Kaminski's pioneer study with Rico in 2004, other authors have published on this topic and revealed that the ability to learn object names is not limited to just Border collies. Indeed, a Yorkshire terrier and a mongrel dog have also provided robust evidence of their abilities to learn words as object referents (e.g. Griebel & Oller, 2012; Ramos & Ades, 2012). But for all that, are they isolated cases of genius dogs or do these findings support the assumption that perhaps, with a minimum training, any dog might be skilled to learn word-object associations? To tackle this question, we recruited a heterogeneous group of dogs composed of numerous breeds and ages. Moreover, dogs were companion dogs recruited from owners who volunteered to

participate in the study meaning the dogs grew up in different environments and were all differently educated.

Acquisition of the word-referent task

Dogs were trained on the word-referent concept during a six month period by implementing Pilley's protocol (Pilley, 2014; Pilley & Reid, 2011). They were mainly trained by their owners at home during short daily sessions and once a week by an experimenter and a dog trainer in a neutral room. Dogs' acquisition of the task was tested at the end of the training phase by performing tightly controlled comprehension tests. Results showed that dogs as a whole performed significantly greater than chance levels. Nevertheless, when focusing on individual performances, only a small number of dogs seemed to have acquired the task. Several reasons might explain the discrepancy between our results and the impressive performances of dog cases reported in the literature.

First, dogs recruited in our experiment were all initially naïve about the task and benefited from short daily training sessions over a six month period compared to three years of intensive training for Chaser (4-5 hours a day) (Pilley & Reid, 2011). Rico (another border collie) was taught this concept since he was 10 months old while the experiments were conducted as he reached 9 years of age (Kaminski et al., 2004). Similarly, Bailey (the Yorkshire Terrier) was reported to already know 120 object names by the day she participated in Griebel's study (Griebel & Oller, 2012). Thus, these dogs were already very familiar with the task and knew hundreds of object names when they enrolled the experiments. This also constituted a major asset because these dogs were tested on a much larger number of objects. In comparison, our dogs acquired between 2 and 8 object names in six months, which is much less likely to produce significant results. Indeed, the low number of trials performed by our dogs (since less objects were pitted against each other during testing) considerably reduced the power of the statistics. For example, dog subject N°2 demonstrated remarkable fetching scores (75% correct retrievals) but as he performed a single testing session which consisted of only four trials, his performances were unlikely to trigger a significant result.

Secondly, a notable difference stands in the age of acquisition of the task. Dogs from previous studies were all in their first year of life when they were initially taught object names (Chaser was 5 months old, Rico 10 months old and Bailey 2 months old). In contrast, the dogs recruited in our experiment ranged from 3 months to 8 years when they were first trained on this task. As with human infants, puppies might be more plastic to learning. Results from previous studies suggested that learning is similarly age-sensitive in dogs as it is commonly described in humans and

nonhuman primates or rodents (e.g. Chan et al., 2002; Head, Mehta, Hartley, & Kameka, 1995; Milgram et al., 1999; Wallis et al., 2016). Moreover, laboratory dogs are considered as relevant animal models for human aging diseases (e.g. Alzheimer disease) as they develop similar age-related neuropathologies, and exhibit similar decline with age in sensorimotor abilities, selective attention, executive functions, etc. (e.g. Adams et al., 2000; Head et al., 1995; Milgram et al., 1994). Thus, according to the literature, cognitive functions in dogs seem to be similarly sensitive to development and aging as it is for humans. Our results also support this hypothesis since four out of the five well-performing dogs of our experiment were puppies. The fifth was one of the youngest adults of the sample (18 months olds). Moreover, we demonstrated that only the group of **puppies performed significantly above chance** whereas the **group composed of adults performed at chance**. The difference between puppies and adults also arose out of the number of objects in competition during retrieval, which reflected the difficulty of the task. We showed that adults had on average significantly fewer objects placed beside each other compared to puppies. A generalized linear mixed model that included the difficulty of the task as a co-variable revealed that **puppies were indeed significantly more accurate at completing the task than adults**.

Altogether, our results, in addition to previous findings, strongly suggest that adult dogs encounter heavy difficulties in learning novel concepts, tasks or complex information. This finding has substantial implications for dog trainers as it indicates that teaching new tricks, tasks, general obedience, etc. may only be really efficient if done during puppyhood, which may represents a *sensitive period* for learning in dogs.

Nevertheless, puppyhood is actually difficult to define. Sexual maturity is often considered as the criterion that distinguishes a puppy from an adult. However, dog breeds have heterogeneous life expectancies and thus reach sexual maturity at different ages (in months). Previous studies that tested the effects of aging on the acquisition of novel tasks in pet dogs split their dog subjects into 4 or 5 age groups (late puppyhood, adolescence, early adulthood, middle age and late adulthood) (Milgram et al., 1994; Wallis et al., 2016) according to Siegal and Barlough's nomenclature which aimed to reflect the main developmental periods of dogs (Siegal & Barlough, 1995). Originally, we aimed to transpose these age categories to our study. Unfortunately, as explained in the methods, almost half of the recruited dogs left the experiment before its end and too few dogs conducted our comprehension tests to be split into so many age groups. Each age group would have only consisted of a few dogs which would have been statistically inappropriate. Therefore we divided our dog sample into two age groups. The cutoff was established at one year old as sexual maturity stands around one year for most of the breeds recruited in our experiment. Nevertheless we are aware that the high variability in age within the adult group ([14-96 months])

constitutes a limit of interpretability of our results. Future experiments investigating cognitively demanding tasks would be helpful to bring more insight about age effects in dogs.

Why did adults fail to learn the task?

Nonetheless, the results give rise to another important question: **why do dogs, and especially adult dogs, have such difficulty in learning this novel task?** At a behavioral level, most adult dogs presented **strong stereotyped behaviors** and **reduced flexibility** to adapt their responses according to prior feedback. More specifically, almost all adult dogs manifested a **preference for either an object or a side position** or even for both. In the majority of the cases, the preference was for a soft object (rarely for a plastic toy), and/or for the object located at the extreme left or at the extreme right position. When the favorite object was removed from the sample because the dogs were making too many errors, these dogs shifted their preference to another object. Therefore, for some dogs, we had to divide the set of objects into two subsets during testing; one composed of the favorite objects and the other of the neglected ones. Note that this only happened for adult dogs. Previous studies found similar side-bias, object-bias and/or perseverative responding in older dogs that contributed to their deficits to acquire a task (e.g. Prichard et al., 2018; Wallis et al., 2016). One could argue that more neutral stimuli (e.g. basic shapes stimuli) or apparatus (e.g. touchscreen) could have prevented these stereotypical behaviors from happening and future experiments would be required to investigate this hypothesis. Another explanation is that older dogs may require more time or more trials to understand a novel task. Previous studies showed that older dogs reached learning criteria after a substantially higher number of trials than younger dogs (Head et al., 1995; Milgram et al., 1994). In our experiment, the practice frequency as well as the duration of the training phase might not have been long enough for adults to comprehend the task. It would have been interesting to continue the teaching for a few more months but this option was unfortunately not conceivable for owners.

Another potential influencing parameter is related to the teaching of the task per se. Dogs were recruited from owners who volunteered to participate and were mainly trained by them at home. Thus it was impossible to control **owners' assiduity and regularity to teach the task** although weekly practice sessions in front the dog trainer enabled us to verify the precision of their teaching method. Nevertheless, this parameter could not be evaluated as a possible explanation for heterogeneity, and we are aware that it might have accounted for dogs' performances.

Following on from this idea, **dogs' own assiduity, concentration and motivation** during training presumably also played a role in their comprehension of the task. We observed a **high variability in dogs' temperament** (e.g. excited dogs, fussy dogs, nonchalant dogs, etc.) even between dogs

of the same breed and sometimes even between two sessions for the same dog. This factor could not be objectively measured to be considered as a variable in the analyses. Moreover, **dogs were rewarded differently** according to their character. For each dog, the type of reward was kept throughout the training phase. Rewards consisted of play, food treats and/or praise. At first glance, the use of different reward types might be questionable. Nevertheless it appears unlikely to explain the inter-subject differences in performances. Indeed, an fMRI study conducted on fifteen awake dogs showed that dogs have analogous neural mechanisms for preference and reward as humans and demonstrated roughly equal or greater brain activation to praise versus food (Cook, Prichard, Spivak, & Berns, 2016). This suggests that **praise was almost as efficient as food** to provide influential feedback of accuracy in our experiment. Likewise, Chaser was only rewarded with praise whereas Bailey received a mix of food, play and petting as positive reinforcements (Griebel & Oller, 2012; Pilley & Reid, 2011).

Finally, **intrinsic factors relative to genetics** might also have accounted for the inability of several dogs to acquire the task. In this experiment we asked whether dogs coming from any breed could be capable of learning word-object pairs. However, dogs represent one of the most morphologically diverse species on Earth with **exceptional systematic variation between breeds** (e.g. Byosiére, Chouinard, Howell, & Bennett, 2017; Wayne, 1986). Heterogeneity between breeds is reflected by a **diversity of morphologies, behaviors and cognitive aptitudes**. Thus, some breeds may be more better at comprehending complex cognitive tasks than others. In fact, breeds that were not genetically selected to solve problems and/or to comprehend human vocalizations may be more exposed to difficulties when confronted to these types of situations. Previous studies on word comprehension by dogs mainly reported cases of border collies. Apart from them, only a Yorkshire Terrier and a mongrel dog have shown similar word learning abilities. Another study investigated brain activity for human words processing after teaching the same word-referent protocol to a heterogeneous group of dogs (Prichard et al., 2018). Dogs were expected to learn two object names with an arbitrary performance criterion set at 80% correct retrievals for one object and at least 50% for the other. Such performances, as well as the number of objects learned by these dogs, are far from providing convincing evidence that they acquired the word-referent concept. Nevertheless, dogs reached the learning criteria set by the authors after a period of 2 to 4 months of training, which suggests that dogs originating from various breeds appear to be able to discriminate between two cross-modal associations. Intriguingly, in our experiment, border collies or mixed breed border collies were not the best performing subjects. The dog that exhibited the highest level of performance according to accuracy and number of objects placed beside each other was a mixed breed hunting dog essentially composed of the English Pointer breed. A Beagle also performed significantly above chance. Thus we provide the

first evidence that hunting dogs appear also capable of learning object names. This is not surprising as these breeds have been selected to cooperate with humans during hunting and to respond to human signals.

Conclusion about the acquisition of the word-referent task

In conclusion, the hypothesis that perhaps **any dog would be skilled enough to acquire words as object referents after a minimum training seems doubtful**. This reveals the **extreme difficulty of this task** while dogs have been shown remarkably talented in numerous other paradigms (cf chapter state-of-the-art). Several factors might have played a significant part in this result but it was challenging to isolate and investigate the role of each potential parameter. Amongst them, **age appears to be a critical factor** that could to explain a great part the discrepancy of performances between dogs. Dogs seem less flexible in learning a complex task after puppyhood. However, it is very much likely that cognitive skills - like learning a novel task - depend on the interactions between environmental, genetic (including age and breed) and social factors, in dogs as it does in humans (e.g. Baltes, 1987).

Lastly, and also interestingly, dogs seem **more gifted to learn words for action commands than for object names**. In an unpublished study, Ramos et al. submitted dogs to training procedures similar to those used to teach object names and indicated that while dogs were all able to acquire correct responses to action commands, most failed to master word-object associations (Ramos & Mills, 2009). We observed a similar pattern in an analogous experiment conducted on a few dogs (not reported in this thesis). From a dog's perspective, learning words as verbal commands for actions may be more ecologically valid than learning them as labels for objects.

Long-term memory of word-referent pairings

The second aim of this experiment was to examine whether dogs not only encoded but also **stored this knowledge in memory**. We hoped to understand if humans' distant animal relatives also display long-term memory skills for complex information. Thus, we tackled the question of long-term memory formation by dogs on the basis of this cognitively demanding task. We also investigated whether some parameters underlying long lasting memory formation in humans are similarly shared by dogs.

▪ *Long-term recognition of the target objects*

To these purposes, retention tests were carried out after a **3-month intervening period with no rehearsal of the material to remember** during this time delay. A first discrimination test

demonstrated that dogs could very **accurately recognize the objects that were positively reinforced during training** when dispatched among distractors (i.e. dogs' toys) that had not been seen for the same amount of time. Older dogs were as accurate as puppies to solve this test. These results demonstrate that **discrimination based on perceptual sensory modalities** (e.g. vision) in dogs is **robustly maintained after a long time interval independently of age**. A previous study investigated aging effects on cognitive functions in dogs and provided similar results (Wallis et al., 2016). Dogs were trained to discriminate clip art pictures in a forced two choice procedure (S+ and S-) involving a touchscreen apparatus and were tested after a six month break. Results indicated that long term memory was maintained into old age, with no difference in performance in any of the five age groups. It can be reasonably claimed that this type of **memory which only involved one sensory modality** (e.g. the visual modality) **is more likely to withstand aging**. But are dogs equally gifted in memorizing more complex information, like bimodal information, for the same amount of time?

- *Long-term retention of the target names*

In order to assess if dogs also remembered the names of these reinforced objects, we conducted the same stringent **word-object associative tests** used after the end of the training phase. Results showed that **dogs as a whole performed significantly above chance levels**. This finding suggests that dogs consolidated and **stored the entire information** that they could later retrieve upon command. This is consistent with the impressive ability of previous dogs shown as being able to retrieve hundreds of objects by name (e.g. Kaminski et al., 2004; Pilley & Reid, 2011). Although their memory performances had not been directly investigated by the authors, memory processes must have been partly responsible for their remarkable skills. For instance, Chaser, who acquired a high number of object names (>1,000), got only monthly rehearsals of the associations since her owner tested her on each object only once-a-month (Pilley & Reid, 2011). Her high rate of performance during these tests confirms the assumption that she stored this knowledge in memory. There is little other evidence of retention of higher-order cognitive information by unrelated species after a considerable amount of time reported in the literature. A sea lion was shown capable of remembering abstract problem-solving strategies after one year and up to ten years for a relational concept (Reichmuth Kastak & Schusterman, 2002). Another study revealed that horses can remember stimuli they encountered during discrimination and categorization learning tasks six years to a decade earlier (Hanggi & Ingersoll, 2009). The horses also recalled a relative size concept rule that they had learned seven years prior to the test. Together, our findings in addition to the previous evidence suggest that **once an animal has learned a conceptual task**

and stored this information in memory or learned sensory inputs from uni- or multi-modalities, the knowledge later remains stable for a long period of time.

Next, when focusing on the performance at an individual level, only one dog showed significant retention over three months (dog N°19). This dog was also the best of the cohort to acquire the task, according to the number of objects in competition and the performances during the comprehension tests. First, it is not surprising that dogs who did not learn the task and failed during the comprehension tests would be equally unsuccessful three months later. Furthermore, as for the comprehension tests, the small number trials made it difficult to obtain significant results in some dogs. For instance, three dogs mastered the task with high scores of accuracy (75% correct retrievals for dog N°2, N°10 and N°20) but as the memory test consisted of only four trials for each of these dogs, the statistical tests were unlikely to be significant. For the same reason, we did not analyze the evolution of the performance of each dog individually, but it appeared that few enhanced their fetching scores after a three month break (9 dogs), whereas the performances of others declined (7 dogs) or remained stable (5 dogs). One explanation for the increase in performance for a few dogs could reflect the lack of rehearsal during a long time period. Several dogs were appeared to be tired with the repetitive frequency of training on the same task and lost motivation to complete the comprehension tests. Their arousal and concentration was noticeably superior after the long intervening period. In contrast, dogs who were keen to do the task efficiently during training were, in some cases, very excited to interact with the objects again after the long break. Thus, they were more willing to fetch a random object to play with it behind the black panel than to complete the task appropriately. This constituted a limit of our design. In a future experiment, it would be interesting to use more neutral stimuli or a more standardized apparatus in order to explore the question of long term memory abilities by dogs. For example, a touch screen apparatus might be more appropriate and less subject to interfering variables.

We also aimed to know if age had similarly accounted for the memory performances as it did for the acquisition of the task. To address this question we analyzed the performances of both age groups (puppies versus adults) independently. We found that **puppies as a group performed significantly above chance levels** whereas **performance of adults as a group were not significantly different from chance**. A generalized linear mixed model that included the difficulty of the task as a co-variable (i.e. number of competing objects) demonstrated that **puppies were significantly more accurate in remembering object names than adults**. In contrast to visual recognition paradigms, this **cross-modal task appeared to be sensitive to age**. It could be that **age effects are better detected in more complex tasks**. Indeed, the word-object testing procedure is particularly challenging cognitively as it requires working-memory abilities

(maintaining the name of the target object pronounced by the owner actively in memory until the fetching act), selective attention toward each object, sensory skills (i.e. the sensory modalities used by dogs must be unimpaired to solve the task (e.g. vision, olfaction, taste)). As previously established by a few studies, old dogs can develop similar age-related deficiencies to humans, especially for sensorimotor abilities, selective attention, working-memory and executive functions (e.g. Adams et al., 2000; Head et al., 1995; Milgram et al., 1994). Therefore, it is not surprising to observe aging effects for this type of memory task that are not evident in easier paradigms. In humans, aging effects are task or information dependent. Indeed, memory either improves (especially for semantic information), remains stable or declines with age (especially for episodic memories) (e.g. Baltes, 1987).

Conclusion about long-term memory abilities by dogs on the word-referent task

To conclude on this second part of the study, we provide here the first evidence of dogs' ability to **remember higher-order cognitive information over an extended period**. Our results indicate that dogs successfully **recognized items based on their perceptual properties** (color, shape, texture, etc.) and seemed also skilled in remembering **the names of the objects** that have been seen for the last time three months earlier. Nevertheless, the low number of trials performed by each individual dog constitutes a limit in making unequivocal conclusions. **Some factors responsible for learning and memory formation** in humans appear to be similarly shared by dogs, notably **age**. However, the principles and mechanisms underlying cognitive processes - like learning and memory formation – in dogs, are poorly documented and warrant further investigations into this aspect.

CHAPTER VIII.

MOKA, A CASE STUDY TO UNDERSTAND THE LEARNING PRINCIPLES UNDERLYING LONG- TERM MEMORY FORMATION OF WORD-OBJECT PAIRS IN DOGS

General introduction

The previous study resulted in a small number of dogs who reliably demonstrated their aptitude to retrieve objects by name during comprehension tests that occurred six months after initial training. Amongst them, Moka, the third youngest dog of the cohort was the only one capable of correctly retrieving eight objects in a row upon verbal instruction with very few errors. This constitutes the first evidence that a **hunting dog is also able to acquire words as object referents**. Next, once a meaningful concept is established, it is interesting to understand if and how it is remembered over a long time scale.

In this second experimental project, we aimed to **characterize the precise conditions required for long-term memories to be formed in dogs**. Are they comparable to those required by humans? More precisely, we questioned whether we can detect equivalents in the efficiency of different learning strategies between humans and this unrelated species. Therefore, in a series of exploratory experiments, we attempted to **explore the learning principles underlying memory formation for word-object pairs** in our best performing dog subject. First, we intended to examine whether **memory decays as a curve of forgetting** according to different retention intervals. In a second experiment, we directly addressed the question of the **minimum number of exposures** required by this dog to reliably remember the label of a novel object. We also explored the **effects of different learning strategies** on the acquisition and consolidation of newly learned objects. Specifically, we compared the efficiency of (i) **ostensive labeling** alone, (ii) **ostensive labeling coupled with retrievals followed by a feedback of accuracy**, and (iii) **retrievals followed by a feedback of accuracy alone**, on the retention of object names. Memory tests were

carried out after two time intervals: two weeks and one month. In a last experiment, we examined whether Moka was able to **generalize his knowledge** to novel instances in order to evaluate to what extent his learning is flexible and semantic-like.

Before getting to the heart of the matter, we would like to emphasize that this second experimental part of the *dog project* is a single case study. All learning and testing sessions were carried out in a neutral room at the ENVT and were scattered in time over a one year period according to the availability of Moka's owner. Although we put a lot of effort in controlling all parameters and in designing stringent protocols to investigate these questions, we are aware that the following studies are first of all exploratory as they involve a unique dog subject, and sometimes suffer from a shortage of replication (or trials) due to time constraints. When we started this project on long-term memory in dogs, we had hoped that at least a third of the dogs we recruited would actually acquire the word-object associative concept to continue the research on this second and more interesting experimental part of the project. Initially, our inclusion criterion for being enrolled in this second stage of the study was to reliably retrieve at least 10 objects by names without (or with very few) errors. Unfortunately, none of the dogs who underwent the six months training phase reached this learning criterion. Therefore we reduced our inclusion criterion to 8 objects and extended the project with only one dog. Thus, we are conscious that only partial interpretations can be drawn from the following studies and that future experiments will be required to consolidate our preliminary findings.

2 SECOND STUDY. *Memory maintenance across time*

2.1 Introduction

Despite several successful demonstrations of concept or complex task acquisition by animals, subsequent memory formation of that knowledge has not been well studied. The first reason comes from the **difficulty to assess memory in animals deprived of language**. Hence, researchers cannot simply apply experimental paradigms developed in human memory research to the study of animal behavior. Protocols have to fit with the purposes of investigation of a specific type of memory in animals (e.g. episodic-like memory, working-memory, spatial memory, etc.). Researchers also need to adapt the experimental designs of each animal model to its ecological living conditions, by reproducing those conditions as appropriately as possible. However, they also need to isolate and control specific parameters in order to investigate their effects on the resulting behavior. The second reason is related to the **time constraints needed to perform longitudinal studies**. Animals have to be kept at the laboratory for the required time period, or if

they belong to private individuals, then the owners have to be available again in the distant future. Moreover, the animals have to avoid interfering protocols or tasks during the whole time interval which is not easy to control for animals that are not hosted at the laboratory. Furthermore, researchers have to ensure that no rehearsal of the material to remember was provided during the time delay.

Nevertheless, since previous studies brought significant insights into some animals being able to recall abstract concepts or reinforced stimuli after considerably long time intervals (Burdyn et al., 1984; Hanggi & Ingersoll, 2009; C. K. Johnson & Davis, 1973; T. L. Patterson & Tzeng, 1979; Reichmuth Kastak & Schusterman, 2002), the assumption that the first domesticated species - already shown to be gifted for higher-order cognitive skills - might also excel in memory tasks appears to be a rational hypothesis that deserves attention. Dogs already demonstrated **memory abilities for recognizing stimuli based on their perceptual properties after six months** (Wallis et al., 2016) and for **remembering a behavior instructed one month earlier** (Demant et al., 2011a). In our previous study, we showed that, overall, dogs performed significantly better than chance to **remember names of objects encountered three months earlier**. Altogether, this evidence suggests that long-term memory abilities for sensory information are probably not restricted to humans. In contrast, this cognitive trait is perhaps more widely distributed within the animal kingdom than originally thought. If so, **are the mechanisms responsible for memory formation in animals similar to those established in humans?** Moreover, does memory require similar conditions and rules to be formed and maintained in animals? To date, animal research typically lacks the advanced technologies needed to investigate these questions at a neural level, even though the first fMRI experiments on awake dogs have recently been launched. In the meantime, memory formation in dogs can be scrutinized at a behavioral level.

This first experiment addressed the question of **memory maintenance over time in dogs**. More precisely, we aimed to **investigate whether memory strength declines as a curve of forgetting** in dogs as it does in humans (e.g. Conway, Cohen, & Stanhope, 1991; Rubin & Wenzel, 1996). This experiment involved the retention of complex cross-modal stimuli (word-object pairs) and was addressed to the dog subject, Moka.

Moka, who acquired eight object names during a six-month training period underwent a prolonged training phase of one and a half years to learn more object names. Concurrently, his memory for object names was tested after **various time intervals**, specifically **immediately, two weeks, one month, three months and six months** (ongoing for the latter). If memory strength for stimuli involving two sensory modalities and associative bindings degrades with time in dogs, we

expect to observe diminished performances after longer time intervals. In contrast, if memory continues to be stable, no significant changes in memory recall should be measured.

2.2 Materials and methods

2.2.1 *Subject*

The subject is a one year old hunting dog belonging to a veterinary student of the veterinary school of Toulouse, France. This dog, Moka, was taught object names during an initial six month period and pursued training for another year and a half after the first memory tests was conducted. Moka knew 24 object names (20 soft toys, 3 plastic toys and 1 knotted rope) when the experiment ended.

2.2.2 *Stimuli*

Stimuli were 3D objects, particularly toys for children or for dogs, as described in the previous chapter. Stimuli consisted of 8 objects randomly chosen among Moka's set of objects. Objects were associated with pseudo-words composed of two dissimilar syllables, two identical syllables or consonant-vowel-consonant constructions (cf section "*materials & methods*" of previous chapter for more details).

2.2.3 *Procedure*

The entire experiment was conducted in a neutral room of the Vet school. Memory tests took place after various time intervals, precisely immediately, two weeks, one month, three months and six months after last training. For the first four time delays, two memory tests involving eight objects each were conducted. For now, only one memory test has been performed after a six-month time interval. Time delays were randomly assigned across the experimental year. In other words, the memory tests did not follow the gradual increase of the latencies of retention in any particular order (e.g. 1st memory test occurred after a three month break, 2nd after a two weeks break, 3rd immediately after a training session, etc.).

No rehearsal was provided during the intervening periods. Neither further experiments nor training to retrieve objects by names were conducted during the intervening periods, to avoid interferences and to exclude the possibility of an internal rehearsal of the information or of the concept. However, the teaching process of novel object names continued outside of these restricted periods (i.e. the only time the teaching process did not occur was during the break periods).

Each memory test consisted of retrieving eight objects upon verbal instruction **without replacement of the objects**. The eight objects were randomly chosen among the set of objects known by the dog at the time of the experiment (e.g. the 1st memory test involved the eight sole objects known by Moka, the 2nd involved eight objects randomly selected among fifteen objects Moka had already acquired, etc.). Few objects were involved in different memory tests. Thus each memory test consisted of objects with heterogeneous training and testing histories (i.e. some being more anciently acquired by the dog than others).

Testing followed the same tightly controlled procedure as described in section *Materials & Methods* of the previous chapter.

2.2.4 Analyses

For each latency, performance corresponded to the mean proportion of correct retrievals discarding the last trial (100% chance of being correct) over the two testing sessions (or a single testing session for the 6 months latency). Performance was compared to levels expected by chance, calculated on the basis of a hypergeometric distribution, since objects were not replaced after retrieval (refer to section *Materials & Methods* of the previous chapter for a detailed description). Adapted binomial tests were used to assess significant differences between observed and expected data. A generalized linear model was used to assess potential differences in the performance between the five time intervals.

2.3 Results

Moka's ability to remember object names was tested after various latencies. One corresponded to an immediate test, performed at the end of a practice session while others took place after different intervening periods with no rehearsal of the task in between. For each latency, except for the six month time interval (ongoing testing), performance consisted of the mean proportion of correct retrievals calculated across two testing sessions. The 95% confidence interval (CI95%) was given by the binomial tests, considering the two (or sole) test sessions.

Table 11 - Table indicating Moka's performance according to different delays between last training and testing. After each latency, two memory tests involving 8 objects placed beside each other were performed, except for the 6-month interval (ongoing experiment). For each time delay, the table reports the total number of trials performed, the number of successes, the mean expected chance level (calculated on the basis of a hypergeometric law, i.e. no replacement of the objects), the p-values, the dog's mean performance (in %), the 95% confidence interval and the Z score and p-value calculated with a generalized linear model (GLM).

Latency	N trials	N successes	Mean chance level%	P-values (binom test)	Mean performance%	CI95%	GLM Z score	GLM p-value
Immediate	14	11	25	3.25e ⁻⁶	79	0.49-0.95	1.89	-
15 days	14	9	25	0.002	64	0.35-0.87	-0.45	0.65
1 month	14	9	25	0.002	64	0.35-0.87	-0.45	0.65
3 months	14	9	25	0.002	64	0.35-0.87	-0.45	0.65
6 months	7	3	25	ns	42	0.09-0.81	-1.33	0.18

Moka performed **significantly above levels expected by chance** when tested **immediately after practice or after delays of two weeks, one month or three months** (see [Table 11](#) and [Figure 76](#)). Only one memory test had already been conducted after a six month interval. Although performance measured during this test appeared not to be significantly different from chance levels, interpretation of this result will only be possible when all the data has been collected (a second memory test is scheduled). However, for two objects mistakenly retrieved at first trial during this test, Moka automatically corrected his response on the second chance that was given to him (i.e. he was accurate on the second trial (not included in the analysis)). Similarly, errors made at first trial during all other memory tests, were systematically rectified by Moka, either on the second trial (44%, i.e. for 8 objects) or third trial (44%, i.e. for 8 objects) (not counted in the analysis). In only 11% of the times (i.e. 2 objects), he needed more than three trials to find the target item.

Next, a generalized linear model (GLM) compared the performance obtained at the different time delays (including the six month interval), considering the mean performance of the immediate latency as the baseline. The model did not establish any significant difference in performance between the delays of retention.

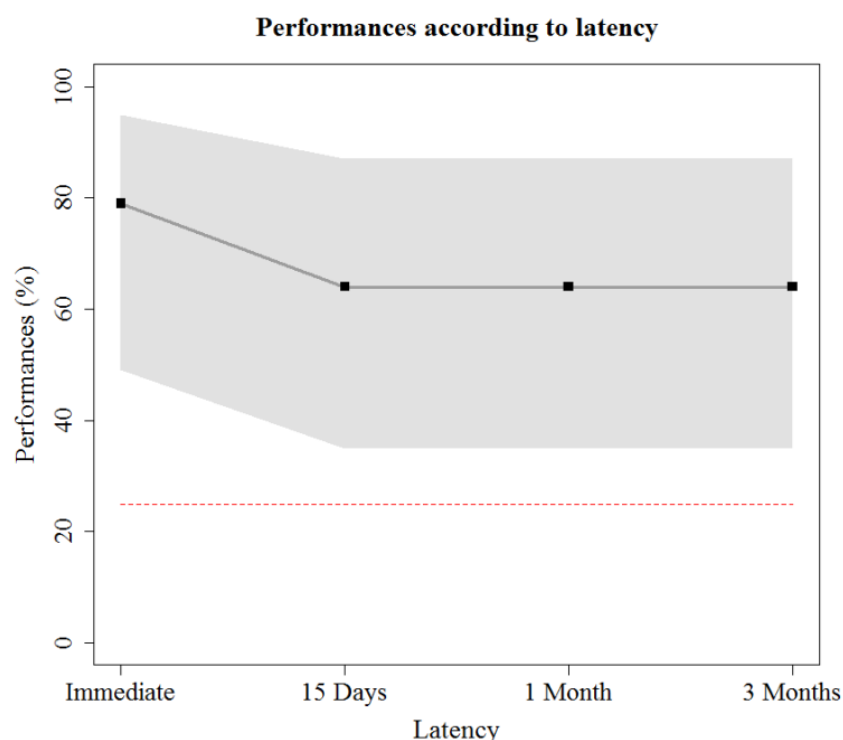


Figure 76 - Graphical showing the mean performance (squared) obtained by Moka during two testing sessions that occurred either immediately after last practice (“immediate”) or after delays of two weeks (“15 Days”), “1 month” and “3 months”. The grey shadow represents the 95% confidence interval given by the binomial tests. The mean expected chance level is shown by the dotted red line.

2.4 Discussion

In this experiment, we attempted to evaluate Moka's **memory strength for word-object associations after various delays without rehearsal**. Moka underwent stringent memory tests during which he was required to retrieve objects by name upon verbal command.

Results demonstrate Moka's **high level of accuracy** to complete the task **irrespective of the delay between last rehearsal and testing**. Firstly, Moka's high performance rate measured directly at the end of training sessions confirms that he reliably comprehended the task. Furthermore, the findings indicate that Moka's knowledge about object names persisted **unaltered for a period of at least three months**. Since no rehearsal was provided during the intervening period, it can be reasonably assumed that to solve this experiment, the information had to be stored and retrieved from long-term memory.

Currently, no serious interpretation can be drawn from the memory score obtained after a break of six months, given the small number of trials performed. A second memory test following this time delay is scheduled. Nevertheless, the first results are promising given that Moka correctly retrieved almost half of the objects on the first trial and accurately adjusted his response when another chance was given to him. Overall, the results demonstrated by Moka in this experiment furthered the findings of previous experiments on long-term memory in dogs that were restricted to visual discrimination tasks. This current finding promotes the hypothesis that this dog has created a **whole memory** by incorporating visual and auditory sensory inputs and by durably associating them together.

However, a non-trivial parameter affected the performance recorded during the delayed tests. Indeed, to prevent the information from being recalled during the time intervals, Moka had neither visible nor physical access to the objects, and no rehearsal of the task was performed with other objects. Thus, during the various delayed testing sessions, Moka was very cheerful to interact with his toys again as well as to perform the fetching task again. Undoubtedly, the longer the time interval was, the more the dog was excited. Observations made from the videotaped recordings revealed that in many cases, Moka ran straight to the first object he saw behind the panel and played with it without looking at the other objects. This was even more noteworthy for the most recent objects he acquired. His owner was instructed to wait at least five seconds before repeating the verbal command, since the dog sometimes required time to analyze each object before making his choice. After a substantial waiting time, his owner repeated the name of the target object, but Moka often simply brought the object he was playing with, which was logically counted as an

error. The use of stimuli that are emotionally strengthened presumably conducted Moka to this behavior which constitutes a limit in our design.

Considering the encouraging results obtained after several months without rehearsal, future experiments would be required, first to reinforce the present results and second to test Moka's memory after longer delays. Unfortunately, this was not feasible within the time allocated to this thesis but it would be interesting to conduct similar retention tests after a latency of a year or even more. It is possible that the absence of significant differences between the different latencies applied in this experiment arose from the robustness of the memory, which may only decrease after longer time scales. If so, it can be suggested that in dogs, once information is learned, it can be stored in memory and recalled months later without significant decay.

3 THIRD STUDY. *Efficiency of various learning strategies in learning and memorizing novel word-object pairs*

General Introduction

Do dogs learn and retain information as humans do? This question has not yet been directly investigated, but is of major importance to foster our understanding of a potential expansion of these cognitive processes to unrelated species. This second experimental project conducted with Moka was aimed at **understanding how a dog learns and retains novel information such as object names**. More precisely, we wanted to know which word learning strategy would be the most appropriate for this dog, to lead to a successful comprehension and retention of a novel object name.

As demonstrated in the previous part of this thesis and in numerous studies from the literature, young children rely on different strategies to acquire vocabulary (e.g. Markson & Bloom, 1997). For instance, they learn and retain new words that are ostensibly taught by an adult (e.g. Schafer & Plunkett, 1998; Woodward, Markman, & Fitzsimmons, 1994). They can also track co-occurrence regularities across multiple ambiguous situations (Smith & Yu, 2008). Finally, they have also been shown skilled to incorporate new vocabularies to their lexicon when facing a problem that involved reasoning abilities to logically exclude items by deduction, i.e. "fast-mapping" (e.g. Carey & Bartlett, 1978). Moreover, this thesis, in accordance with previous studies from the literature, showed that the efficiency of those strategies is age-related and inclined to evolve during development (e.g. Bion, Borovsky, & Fernald, 2013, Zosh et al. 2013). But **would**

these learning strategies be also appropriate for dogs to learn and retain novel object names?

This second study carried out with Moka consisted of a **series of short experiments** designed to shed light on those issues. Remember that the following experiments are merely exploratory; further replications and investigations would be required in order to build upon our pioneering findings. Moreover, for ease of reading, this section does not follow the formal format of an article but will rather report our work, findings and interpretations in a more informal manner.

First, we addressed the question of whether **ostensive labeling alone** could induce learning in our dog subject, Moka. We also examined whether he was able to **fast-map new words to their corresponding referents by using a deductive strategy** (i.e. “referent selection”). To examine these questions, novel object names were taught to Moka using either ostensive labeling or referent selection by exclusion and acquisition of that knowledge was tested during **subsequent comprehension tests**.

Next, we aimed to explore if the **learning conditions** required for memory to be formed in humans are relatively equivalent for dogs. Does long-term memory formation in dogs similarly require an initial memorization phase, during which **memory strength would increase linearly with the number of presentations as we showed in young children**? Does memory formation also primarily rely on repetitions in dogs as it does in humans? If so, **how many exposures** to the novel information would be enough to form a memory trace that can last in time? We tackled these questions by rehearsing learning during successive sessions until Moka reached a criterion of accuracy which would indicate the number of exposures he required to reliably incorporate these new object names in his “lexicon”. Once acquired, these objects were then integrated as any other well-known object in the memory tests performed in the “*memory maintenance across time*” protocol described previously.

Nonetheless, this experiment cannot tell whether **learning was primarily attributable to ostensive labeling or to the reinforcing consequences of the positive feedbacks** received after retrieval during the comprehension trials performed at the end of each session. Thus, to foster our understanding about Moka’s learning skills, we also examined the **effects of three learning strategies on the retention** of object names at long-term. We compared the efficiency of (i) **ostensive labeling alone**; (ii) **ostensive labeling coupled with the reinforcing consequences of positive feedback** during subsequent comprehension tests; and (iii) **learning from the positive reinforcement** received during comprehension tests alone, on the retention of object names. Memory tests were carried out after two time intervals: **two weeks and one month**.

3.1 Does ostensive labeling induce word learning in a dog?

3.1.1 *Introduction*

The remarkable abilities of Rico and Chaser to acquire hundreds of object names and to infer the referent of a novel word by exclusion caused intense debate in the scientific community (Paul Bloom, 2004; Markman & Abelev, 2004). Among them, the actual comprehension that words can refer to objects by dogs was questioned. Some authors speculated whether or not dogs simply formed basic associative learning following an extensive training, which would have little to do with word learning per se. Bloom argued that children can learn words from overheard speech, even if nobody is trying to teach them (Paul Bloom, 2004). He wondered whether dogs would be able to learn a new word by simply being shown an object and hearing a person name it. Current knowledge about word learning by dogs is far from bringing a robust reply to Bloom's issue. Neither is the following study trying to answer this deep question. Nevertheless, one could claim that if a dog actually comprehends the abstract concept that words correspond to objects rather than a mere learning of isolated sound-item associations, he should be able to acquire new object names simply from ostensive labeling as young children do. In other terms, once a dog acquired the word-referent concept, **is he able to learn new object names without the need of positive reinforcements through operant conditioning?** Here we aimed to understand if our dog subject, Moka, who showed skill on the word-object pair concept, would be able to learn new words without extended training or reinforcement but merely from basic ostensive labeling during playing activities. We also intended to know if he could use other strategies to solve the word-object task, such as logical exclusion.

3.1.2 *Material and methods*

3.1.2.1 *Subject*

The subject is a one year old hunting dog, Moka, living with his owner in Toulouse, France. Moka was trained during a six month period to learn word-object associations and knew around ten object names when this experiment was launched.

3.1.2.2 *Stimuli*

Stimuli were novel objects, especially toys, as described in the section *Materials & Methods* of the previous chapter. Stimuli used in this experiment were scrupulously chosen with Moka's owner to ensure that none of them resembled Moka's toys at home. Moreover, each object had to be equally salient for the dog. Thus, the owner firstly had to assess the attraction his dog would

develop towards each object. The experimenter then randomly chose objects that were rated by the owner as similarly salient for the dog. Objects were associated with pseudo-words that were very different from the labels already known by the dog (cf. section *Materials & Methods* of previous chapter for more details).

3.1.2.3 Procedure

The entire experiment took place in a neutral room at the ENVT. The experiment involved investigating the effect of different learning strategies on the acquisition of novel object names by Moka. This experiment focused on the following learning strategies:

- *Ostensive labeling*

In this case, a novel object was given to the dog and the owner ostensively labeled this object while holding it and while his dog was playing with it. The owner repeated the pseudo-word associated to this object a few times in a row (maximum three times). Since the dog was deeply engaged in the playing activity, and probably less or even not at all focused on the corresponding label, the owner was instructed to throw the object away and to ask his dog to fetch it by incorporating the novel label in an utterance, such as “Moka, go fetch [*label*]” or “Moka, bring me [*label*] back”, as he would naturally do during a playing game [Figure 77]. The owner repeated this throwing game twice in a row. Finally, the owner hid the object alone without other stimuli behind the panel and similarly asked his dog to find it and to bring it back to him upon similar verbal commands. This was repeated twice and it was essential for Moka to understand that this novel item was now part of his set of known objects rather than a simple toy. Indeed, at the beginning of the six month training phase, the dogs’ own toys were considered as distractors (refer section *Materials & Methods* of previous chapter) and were never positively reinforced when dogs mistakenly selected those objects. Thus, dogs rapidly learnt to ignore their toys during the informal tests performed at home or at the vet school. Pilot experiments revealed that newly introduced objects were systematically considered as distractors by the dog (i.e. considered as basic toys) if they were not displayed at least once in the configuration of the formal tests (i.e. behind the black panel) and positively reinforced at least once after retrieval in this configuration. In the current experiment, in total, the name of the object was repeated between five to ten times, and the dog interacted with the object for about five minutes. Reward was only provided when the dog brought the object back to his owner, as the owner would usually reward his dog for general obedience.

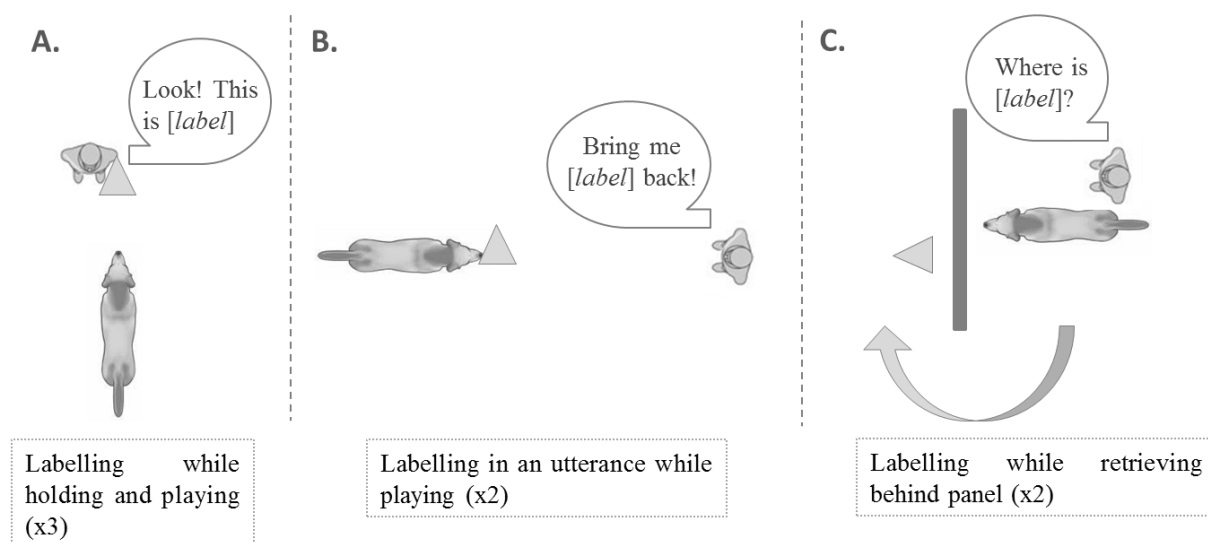


Figure 77 - Figure illustrating the ostensive labeling procedure completed by Moka's owner when introducing a novel object. He first held the object in front of his dog and labelled it three times in a row while Moka was allowed to play with it (A); he then initiated a throwing game: he threw the object twice and asked his dog to fetch it (B) and finally he hid the object alone behind the panel and instructed Moka to bring it back to him upon command (C).

Ten to twenty minutes elapsed between this labeling procedure and the subsequent comprehension test. During this time delay, Moka laid-down to get some rest, went outside for a short walk or played short non-interfering games with his owner. Moka then completed a comprehension test to assess whether he correctly mapped the previous unknown label to the novel object introduced to him. For this test, the previously introduced object was placed together with three familiar objects already known by Moka (i.e. randomly chosen among the set of Moka's objects that already have names) behind the panel. In the first trial, the owner always asked Moka to bring a familiar item and in the second or third trial he asked for the newly introduced object using the same verbal instruction as for the familiar ones. This way of testing controlled that Moka's response was not due to an inherent saliency preference toward the recently introduced object in comparison to the familiar ones. Moreover, for these test trials, objects were always replaced behind the panel after retrieval, so that the probability of success remained stable at 0.25 for each trial (i.e. chance level set at 25%). Finally, the recently introduced object was requested only once during the testing session since the positive feedback that followed Moka's retrieval would have reinforced the mapping, and thus strengthened the learning. This would have constituted a bias if Moka were to have been tested again on that word-object association. This protocol was replicated five times (each time with different items) spaced over five sessions scheduled over one year. Performance was analyzed using binomial tests.

- *Referent selection by exclusion (i.e. “fast-mapping”)*

To assess whether a dog is able to map a word to its referent by exclusion, the referent selection procedure involves a unique choice trial in which the subject is provided a novel name and is expected to choose the sole unnamed object arranged amongst a set of familiar objects that already have names. Moka’s ability to infer a novel word to its referent by exclusion was examined by implementing Kaminski’s pioneer procedure (Kaminski et al., 2004) but with a restricted number of competing objects since Moka knew much less object names than Rico by the time he performed this task. Moreover, as we aimed to directly compare ostensive labeling and referent selection by exclusion, the novel object was similarly introduced to Moka prior to the test except that no label was provided. Precisely, the owner was instructed to similarly play throwing and fetching games with that object during a five minute period with the difference that no label accompanied the verbal commands (same procedure than shown on [Figure 77](#) without the label). The referent selection test was similarly carried out after a ten to twenty minutes break. Likewise, the novel unnamed object was pitted against with three familiar objects behind the panel. Moka was first requested to bring familiar objects upon command (on the first or first and second trials) before being asked for the novel object. Once again, this procedure controlled that a correct retrieval of the novel item was not merely due to novelty or saliency preference. Objects were always replaced behind the panel after retrieval, so that chance level was equally set at 25% for each trial. The protocol was replicated 9 times (each time with a novel object) spaced in time over a one year period. Performance was analyzed using binomial tests.

3.1.3 *Results and discussion*

Moka’s ability to grasp the associative link between a word and its referent after being only briefly introduced to the novel object was examined. In this condition, the novel object was ostensively labeled while Moka interacted with it during playing and fetching games. A comprehension test carried out about ten to twenty minutes after learning involved three familiar objects (i.e. objects that already have names) and the recently introduced one.

The results indicate that Moka accurately retrieved the target object upon request during testing in four out of the five replications of the experiment (binomial test, $p < 0.01^{**}$) [[Table 12](#) & [Figure 78](#)]. Moreover, Moka almost never failed on the familiar objects that were requested on the first or first two trials. This excludes the possibility that the correct retrievals for recently introduced objects were due to an impulsive behavior caused by personal attraction preferences.

Table 12 - Table recapitulating Moka’s performance to remember the names of recently labeled objects (ostensive labeling and ostensive labeling + distractor) and to disambiguate referential situations (referent selection). Table indicates the number of replications of each condition (only one test trial performed for each novel object); the number of successes (i.e. number of correct trials); the expected chance level; the p-values, Moka’s performance and the 95% confident interval given by the binomial test.

Conditions	N replication (1trial/object)	N successes	Expected chance level (%)	P-value	Performance %	CI95%
Ostensive labeling	5	4	25	0.016	80	28-99
Referent selection by exclusion	9	5	25	0.049	56	21-86
Ostensive labeling + distractor	7	2	20	0.63	29	4-71

Nevertheless, it cannot be advanced that this result was mediated by the ostensive labeling procedure itself. The option that Moka logically excluded the objects that already had names to solve the task cannot be discarded. To test this hypothesis, Moka’s ability to map a novel word to a referent by deduction was also tested.

To allow comparison with the ostensive labeling method, objects were similarly introduced to Moka during playing activities, except that no labeling was provided. Hence, during the subsequent “referent selection” trials, the targets had the same “novelty history” as the objects used in the ostensive labeling condition.

Results reveal that Moka **accurately retrieved the unnamed objects by exclusion** in five out of the nine replications of the experiment which is **slightly significantly above chance** (mean accuracy: 55%, $p=0.049$) [Table 12 & Figure 78]. This confirms that **Moka is capable of logical reasoning skills to map a novel word to its referent by deduction**. No significant difference could be established between the results obtained here and the previous ostensive labeling condition (GLM, $Z=0.892$, $p=0.37$). This suggests that Moka may not have benefitted from the ostensive labeling procedure that occurred in the previous condition but rather that he achieved the former tests by using a logical exclusion strategy.

To verify this hypothesis, we carried out a **third condition** which involved **two novel objects**. Both were introduced to the dog as described in the methods but **only one of them was labeled**. Note that this procedure replicates traditional word learning paradigms developed with children (e.g. Woodward, Markman, & Fitzsimmons, 1994). Precisely, for the unlabeled item (called “distractor”), the owner used the same verbal commands without incorporating any name (e.g. “go fetch”, “bring back”, etc.). The owner interspersed the presentation of the two objects in a random order (e.g. obj1-obj2-obj2-obj1-etc.). In total, the dog spent about 5 minutes with each object. We also controlled that Moka had no innate preference for one object which would have biased the following test. To this end, both objects were simultaneously displayed on the ground in front of

the dog and we scrutinized his behavior toward each object during one minute. If he similarly and alternatively interacted with both, we continued the experiment; otherwise we discarded both objects and started the experiment again with a novel set of two items. Once again, ten to twenty minutes elapsed between the last interaction with the objects and the test. Testing consisted of placing the newly named object (target), the novel unnamed distractor and three familiar objects that already had names beside them behind the panel [Figure 78]. In this configuration, “referent selection” by exclusion was unlikely to occur since two objects were “novel”. The five objects were pitted against each other and replaced after retrievals, which sets the chance level at 0.2. Again, the first or first two trials involved familiar objects and the second or third asked for the newly labeled one. The whole procedure was replicated seven times scattered in time; each time with novel sets of objects.

For this third condition, our hypotheses were the following: if Moka learnt the name of the object that was ostensibly labeled, he should accurately retrieve it during testing. Conversely, if ostensive labeling did not constitute an efficient learning strategy, Moka should randomly retrieve one of the two newly introduced objects when requested for the target item.

Results reveal that Moka’s performance to retrieve the target object upon command was not significantly different from chance in this condition (binomial test, $p > 0.05$, see Table 12 & Figure 78). This confirms that **Moka did not learn any linkage between the label that was pronounced during learning and the object that was presented at the same time**. It also suggests that about **five tokens of the label** and a **five minute exposure to the object** were **not sufficient for him to properly integrate this new mapping** into his “lexicon”. It should be noted that very few errors were made when Moka was asked for familiar objects, putting aside the possibility that the low performance obtained in this last condition was merely reflecting a lack of comprehension of the task as a whole.

Originally, if significant results had been reached, we had intended to foster our comprehension about Moka’s reasoning abilities by examining whether he would have retrieved the distractor when using a novel word. Would he have attributed an unknown word to the sole object that had never been labeled before? Unfortunately, since both newly introduced objects seemed not to have been attributed a label from Moka’s perspective, it made no sense to conduct this further investigation.

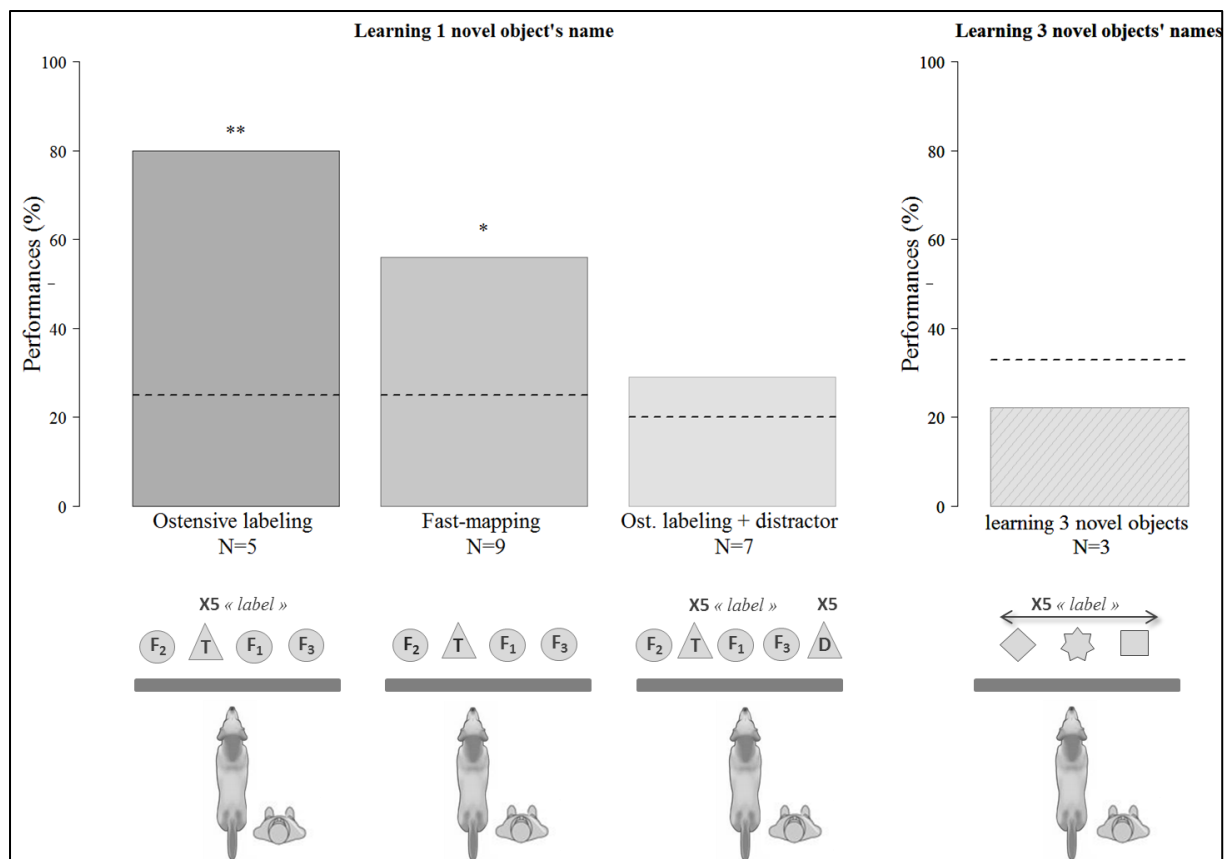


Figure 78 - Illustration of the testing configurations of the four conditions and their corresponding results. The first condition (left) involved a single newly introduced object (T=target) labeled approximately 5 times prior to the test, and placed beside three familiar objects (F₁, F₂ and F₃). The experiment was replicated 5 times over independent sessions scattered in time (N=5). Each target object was requested only once. Performance consisted of the percentage of correct retrievals over the 5 sessions. The second condition (“fast-mapping”) corresponded to a replication of the first condition except that no labeling was provided during the former presentation of the target object. The experiment was replicated 9 times, and consisted of one test trial per target. The third condition (Ost.labeling + distractor) involved two novel objects, one being labeled approximately 5 times prior to the test (T) whereas the other was not (D=distractor). The experiment was replicated 7 times, with different sets of objects each time. Each target was requested only once. The last condition (right) examined Moka’s ability to learn three novel objects at the same time. Each object was labeled approximately 5 times before testing. All three objects were placed beside each other during testing. The experiment was replicated 3 times with three different sets of objects, each target being requested twice in a testing session. Dotted lines represent the levels expected by chance.

Finally, one could have argued that this type of testing configuration could also be solved by extended logical reasoning abilities. Indeed, in case of significant results, the alternative explanation that the dog first excluded the objects that already had names and then excluded the one that had never been named would have constituted another logical interpretation. Aware of this potential confounding explanation, in parallel to the three other conditions, we conducted a **fourth condition** which consisted of teaching **three novel object names during a single learning session**. Note that this configuration reproduces the experiment we developed with children and somehow allows comparison between both models. The three objects were alternatively introduced to Moka - as described in the *Material & Methods* of this study - and each object was attributed a label. Moka heard the three novel labels about five times each during

learning. Learning lasted approximatively 15 minutes, namely a total of 5 min per object (1 min/presentation interspersed). After a break of about 10-20 minutes, a testing session was carried out and **only involved the newly learned objects** that were pitted against each other behind the panel [Figure 78]. In this configuration, the probability of retrieving the correct object by chance was established at 33%. On each trial, the owner randomly requested one object (a generator randomly designated the object to request at each trial), and each object was requested twice. This 4th condition was replicated three times, each time with three novel objects. Overall, Moka's performance was not significantly different from chance (mean performance: 22%, binomial test, $p > 0.05$, see Table 12), indicating that he was **not able to learn three novel object names during a single learning session that used ostensive labeling**. This 4th condition of the experiment confirms that **brief exposures with novel objects that are ostensively labeled are not sufficient to induce word learning in our dog, Moka**.

This result is **far from what was previously demonstrated in 2-year-olds**. Remember that in study N°2, each of the three objects was presented to the participants during a unique 15s video clip and labeled three times in a row using an ecologically valid carrier phrase. Nevertheless, making **direct comparisons would be unwise** given substantial differences in (i) the **apparatus utilized** (movies displayed on a touchscreen vs. real objects), the **number of participants** (twenty-three 2-year-olds vs 1 dog) and the **length of experience with this abstract concept**. As reviewed earlier, children usually acquire the word-referent concept during their first year of life. Consequently, the 2-year-olds we recruited had at least a one year experience with the concept and thousands of words in their receptive repertoire when our protocol began. In contrast, Moka reliably comprehended the mapping concept between words and objects only few months before this experiment started and had a receptive "lexicon" of barely more than ten words.

3.1.4 *Conclusion*

To conclude on this experiment, we showed that Moka was not capable of learning the names of novel objects through few ostensive labeling repeats provided during basic playing activities. Even when only one novel object among two was attributed a label, Moka did not manifest retention of this association after a delay of about half an hour. On the other hand, when being taught a single novel object, significant retention was evidenced. We showed that this finding was probably due to Moka's ability to map a word to his referent by using an elimination process. In sum, our dog subject does not seem to have the ability to learn the name of an object following ostensive naming, comforting Bloom's assessment that only children can learn words from overheard speech, even if nobody is trying to teach them (Paul Bloom, 2004).

3.2 Is memory formation related to the number of exposures in dogs?

3.2.1 *Introduction*

This leads us to one central issue at the heart of the M4 project, namely the question about the **minimal number of exposures that are required for a subject to form a memory trace that will last in time**. Would a dog be capable of remembering any situation, person or sensory stimulus a long time period after they encountered it for the last time? If so, how many exposures to the information would be required for them to remember it at long term? Unfortunately, since **animals are deprived of language**, it is **challenging to explicitly demonstrate memory recall in nonhuman animal models**. As discussed earlier, laboratory research has to develop tasks and protocols that fit with the animal model and that are replicable for statistical constraints. In most of the cases, the animal model will firstly have to be trained on a specific task. It has been shown that the number of trials needed to acquire a novel task is inherent to each task and to each subject (e.g. young subjects learn novel tasks more rapidly than older subjects). But **once a task is properly understood** by an animal, **how many repeats of novel instances related to the task** would be required for the animal to recall those **specific sensory inputs** after a substantial delay? This would constitute a reliable and replicable way to examine the minimal number of exposures underlying long-term memory formation for sensory inputs in animals. To date, and to the best of our knowledge, there is no clear evidence related to this question in dogs.

In the context of the word-referent concept, one should disentangle visual from auditory sensory inputs. Remember that the participant must encode the physical features that compose the referent and the auditory sound pattern that refers to this particular referent. Accordingly, with this specific paradigm, the number of exposures corresponds to the number of times the animal hears the label and the number of times he sees its corresponding item. Previously, we showed that Moka was unable to retrieve a novel object by name, if it was ostensibly labeled about 5 times during five brief presentations that occurred in the same session. Personal explorations suggested that increasing the number of repetitions of the label during the learning session did not lead to successful retention. It might be that for dogs, learning has to be distributed in time in order to induce a robust memory. Pilley indicated that each time he gave Chaser a novel name to learn, he first held the object and repeated the associated new label a few times (Pilley & Reid, 2011). Then, Chaser engaged in several playing periods during which the experimenter repetitively verbalized the name of the object 20-40 times each session. Next, the object was placed on the ground amongst objects that already had names and Chaser was tested on her ability to retrieve the newly introduced item by name. Pilley indicated that Chaser underwent this daily labeling and testing rehearsals over a period of 2-4 weeks. Unfortunately, Pilley gave no indication about

Chaser's accuracy when completing the informal tests performed at home all along this training period. Therefore, this only gives us an imprecise idea of the amount of training and repetitions Chaser needed to reliably incorporate the name of a novel object into her "lexicon". In another study, authors investigated how long it took to a female Yorkshire terrier, with a "lexicon" of over 120 objects names, to learn the name of a novel object (Griebel & Oller, 2012). To do this, they taught the dog the names of two novel items during numerous learning sessions and tested her comprehension of those names during a two-choice identification task at the end of each session, as well as during formal exclusion test trials until the cumulative total of correct trials was significantly better than chance. In total, the dog had been given more than 150 trials of informal two-choice retrieval tests in addition to considerable formal tests with the items over a period of more than a month. This supports the idea that dogs require a huge number of trials to form a memory about the associative link between novel bimodal stimuli. It should also be noted that here, authors investigated the learning rate this dog needed to learn only two novel word-object pairs. It would have been interesting to rehearse the protocol with more items. Furthermore, in this study, the dog exhibited a strong preference toward one of the two objects which biased the results of most of the tests. Therefore, only a rough interpretation of this finding can be established.

On the other hand, Kaminski et al. demonstrated that Rico remembered the names of half of the objects he successfully mapped during a single selection by exclusion trial performed one month earlier (Kaminski et al., 2004). Although the testing design developed by those authors raised several criticisms (e.g. performance could have been attributable to the reinforcement Rico received after initial retrieval (Markman & Abelev, 2004) or to a basic ability to process extended exclusions (Griebel & Oller, 2012)), results remain impressive and unique, and still strongly suggest that a memory can be formed after very few, or even a single visual and verbal exposure, by a dog.

In the following experiment, we decided to continue the teaching process of the objects involved in the 4th condition of the previous experiment over **multiple ostensive labeling sessions spaced in time until a learning criterion was reached**, as assessed by comprehension tests that followed each learning session. To achieve this, Moka underwent a **3AFC** (alternative forced-choice) testing procedure at the end of each learning session. Our **criterion was set at 66% accuracy**, which corresponds to double the accuracy expected by chance.

3.2.2 *Material and methods*

3.2.2.1 *Stimuli*

Stimuli consisted of the three triplets of objects introduced to Moka in the 4th condition of the previous experiment and their associated pseudo-words.

3.2.2.2 *Procedure*

Learning sessions were spaced in time according to the availability of Moka's owner, basically one to three learning sessions per week. Learning consisted of ostensive labeling as described in the previous section [Figure 77]. Each object was always presented and labeled in isolation, but the owner interspersed the order of presentation of the three objects (e.g. obj3-obj2-obj1-obj2-obj3-etc.). In total, one learning session consisted of 5x1min visual exposure to each object, and 5-10 verbal exposures to the associated label. Each session involved one triplet of objects (fatigue would have impaired learning if two or three triplets of objects would have been taught during the same session). About 10-20 minutes after learning, a 3AFC comprehension test was run. The three objects previously labeled were placed beside each other behind the panel. A generator randomly designated which object was the target at each trial. Moka received a feedback of accuracy after each retrieval. If he didn't find the target object at first trial, another (or a maximum of two other) chance(s) was given to him, but correct responses on second or third trial were not included in the current analyses. Objects were replaced after retrieval. Thus, Moka had a 33% chance to bring the correct object by chance at each trial. Binomial tests were used to assess accuracy according to a probability of success set at 0.33.

3.2.3 *Results and discussion*

Moka's knowledge about the names of newly introduced objects was measured during 3AFC comprehension tests performed at the end of each learning session. As already demonstrated at the end of the previous experiment (4th condition), the **first testing session led to inconclusive demonstration of comprehension** of the recent knowledge (mean performance: 22%, $p>0.45$, see Table 13 & Figure 79). None of the nine objects (3 triplets of objects) was retrieved without error at the end of the first testing session, except one that was clearly Moka's favorite object among the triplet. Although the systematic pre-tests of preference did not demonstrate any specific preference bias, it is possible that Moka established this strong preference from the first test trial. Indeed, when tested on that triplet, the dog systematically brought this particular object at first on each test trial. We did not exclude this result from the analysis but we discarded this triplet of objects from the experiment thereafter.

Table 13 - Table indicating the number of triplets of objects taught and tested during each session (in brackets the number of test trials that involved each object is specified). For each session (1st, 2nd and 3rd), the table also reports the total number of trials performed and the number of successes achieved. The expected chance level, p-values, mean performance and 95% confident intervals are also reported

Teaching + testing sessions	N triplets (N test trials/object)	Nb of overall trials during testing	Nb of overall successes	Expected chance level %	P-values	Mean performance %	95% CI
1 st session	3 (2)	18	4	33	0.45	22	6-48
2 nd session	2 (3)	18	10	33	0.047	56	31-78
3 rd session	1 (1)	3	2	33	0.25	66	9-99

Since Moka did not exhibit overall learning of the mappings, a **second learning session was scheduled** for the two remaining triplets of objects. It followed the **same ostensive labeling method** as used during the first learning session. Similarly, comprehension of the mappings was tested about 10-20 minutes after the last labeling trial. At the time of this second test, Moka had a cumulative number of about 15-20 verbal repetitions of the label (i.e. including the repetitions provided during the first test session) and 10-15min visual exposure to the object (i.e. including the visual exposure provided during the first test session). In this second test session, each object was requested three times in a random order. Moka was reinforced after retrievals. Moka's mean **performance was higher than for the first test session and slightly significantly above chance** ($M=56\%$, exact binomial $p=0.047^*$, see [Table 13](#) & [Figure 79](#)). However, as he didn't reach the learning criterion that we established (mean performance $\geq 66\%$), a **third learning session was scheduled**.

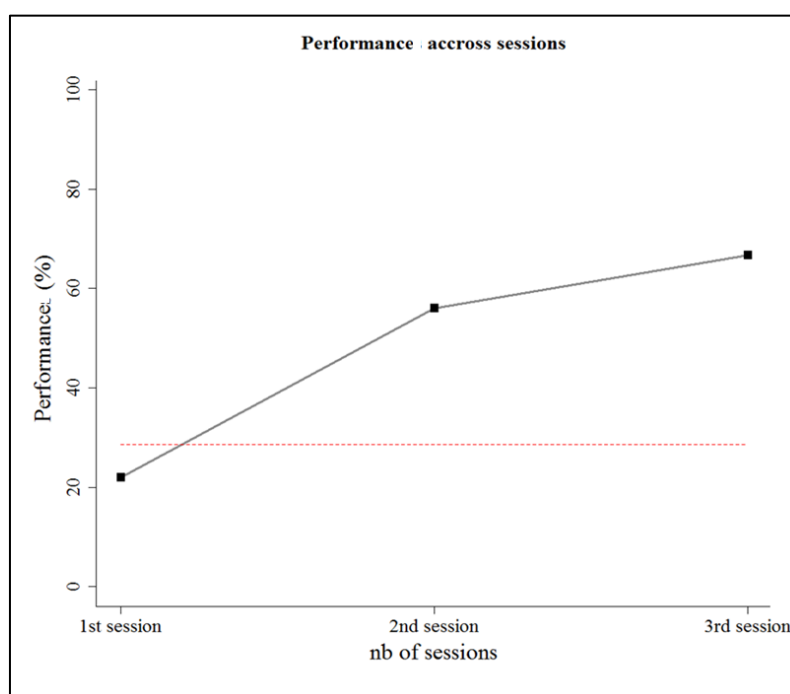


Figure 79 - Graphical showing Moka's performance during the 3AFC comprehension tests performed at the end of each learning session. The red dotted line indicates chance level.

Because of experimental errors, one of the two triplets of objects couldn't be further utilized and the third learning session only involved the remaining triplet of objects. Each of these three objects was similarly labeled by implementing the same method as in the previous learning sessions. Testing was similarly carried out after a break of the same length of time. Each object was requested only once. This time, performance reached our learning criterion of 66% accuracy (i.e. Moka correctly retrieved two out of the three objects) [Table 13 & Figure 79]. **These preliminary results suggest that about 25-30 verbal utterances of the label and 20-25min visual exposure to its associated referent, distributed over three sessions (including both learning and testing), allowed learning to be demonstrated.**

Obviously, this result does not yet bring a robust demonstration of the minimum number of exposures needed to form a memory that will last in time. Firstly there were a number of experimental biases, such as a varying number of triplets involved during each test session, a varying number of trials performed during each test session and the strong object preference Moka developed toward one specific object. Secondly, the small number of items encompassed in this experiment lowers the impact of these results. Future replications of this protocol will be required to build conclusions upon these preliminary findings.

Despite this, our results bring an interesting matter of reflection. First of all, it confirms the idea that the **number of (visual and verbal) exposures** required to form a reliable mapping between a word and its referent is **subject-dependent**. Indeed, results from this experiment, in addition to our personal observations about how long it took for Moka to reliably acquire each of his novel word-object pairs, are far from the hundreds of trials reported by Griebel et al. (Griebel & Oller, 2012). Nonetheless, it is also far from the hypothesis that a single trial could be enough for a consistent encoding of such cross-modal stimuli. At least, it seems not to be the case for objects acquired through an ostensive labeling learning method. At first sight, this finding would effectively rule in Bloom's favor, as he argued that only human children are able to learn words from overheard speech, even when adults are not trying to teach them those words (Paul Bloom, 2004). This is actually true for children who have a wide lexical repertoire. But currently, as far as we know, and as a direct implication of our work on young children, there is no proof that an infant provided with a ten to twenty receptive vocabulary lexicon is able to successfully acquire new words after a single (or even very few) exposure(s) without extended external cues. Indeed, in their **initial stage of receptive vocabulary learning, infants require dozens of ostensive repetitions accompanied with non-verbal cues** such as object manipulation, eye gaze or pointing to the object, in order to show comprehension on immediate tests (Gurteen et al., 2011c; Hollich et al., 2000; Oviatt, 1980; Schafer & Plunkett, 1998b; Woodward et al., 1994b). **Moka's**

receptive vocabulary “repertoire” is roughly equivalent to the one of a 6-9 month old infant, which corresponds to the beginning of the comprehension of the word-referent concept by human infants (e.g. Fenson et al., 1994; Oviatt, 1980). For our dog, word learning appears to be a time-consuming process, exactly as is the case for infants in their initial stages of language acquisition. However, there is no doubt that young children with larger vocabulary repertoires comprehend new words from overhead speech and without the need of repetition (e.g. as a direct illustration of this claim, see our findings on Part 1, study N°2). The question is whether this would also be the case for expert dogs provided with a receptive “lexicon” of thousands of words? It would be interesting to rehearse this protocol with an expert dog. But for now, there is undoubtedly still a major step before claiming that dogs understand human words in a human way.

Next, this study, in addition to Pilley’s or Griebel’s previous reports about the word learning process their dogs had to undergo to acquire the names of novel object, suggests that **learning has to be spaced in time to be efficient**. Apparently, a single ostensive labeling learning session (with multiple verbal and visual exposures to the material to learn) is not sufficient, even if the label is extensively repeated during this session. It might be that, to optimize the acquisition of novel information, learning has to be split into short sessions spaced in time. In a previous study, laboratory dogs were trained on a traditional obedience task and were divided into four groups that were differentiated in frequency and duration of the training sessions (Demant et al., 2011). Results demonstrated that dogs trained 1-2 times per week had significantly better acquisition rates than daily trained dogs. Moreover, a daily training session rehearsed over a three day period led to significantly better acquisition rates than three training sessions performed in a row by those dogs. Thus, the authors argued that **spaced training was better than massed training for dogs** to acquire a task, which is similar for humans. Indeed, **rest periods** and especially **sleep** are undeniably of great importance and probably the **principal mechanism behind this spacing effect** (e.g. Axelsson et al. 2018; Born, Rasch, & Gais, 2006; Friedrich, Wilhelm, Born, & Friederici, 2015; Wilhelm, Born, Friederici, & Friedrich, 2015). Indeed, **spaced training** offers more **possibility for subjects to mentally replay or rehearse the information** outside of the learning sessions and **involuntarily during sleep**. Sleep has been identified as a state that **optimizes consolidation of newly acquired information**. Therefore, it can be reasonably assumed that when a night separates two learning sessions, consolidation of the information acquired the day before already occurred in dogs (as it does in humans), which leads to enhanced performance compared to two training sessions massed on one day. Interestingly, training schedule established by Demant et al.’s did not affect long-term retention of the learned task as all of their dog groups remembered the exercise well after a four-week break.

Finally, the current results did not truly answer our initial question related to a hypothetical correlation between memory strength and number of presentations during learning. Previous studies suggested that once a behavior (or a specific sensory stimulus) is accurately learned by a dog subject, his memory for that knowledge will remain relatively unimpaired even after a long time period (e.g. Demant et al., 2011; Wallis et al., 2016). But would this also have been the case if only a small number of repetitions were provided during learning? Originally, we had planned to measure Moka's long term retention skill for the object names used in this experiment. Unfortunately, considering the low number of objects that were left when the learning criterion was reached, and since it was impossible to replicate the protocol with new sets of objects because of time constraints, we were unable to pursue our initial objectives. However, without further extensive training, the three objects that remained were integrated into Moka's set of known objects and served in other experiments, such as in the memory tests performed after different latencies shown previously. Therefore, we did not report Moka's memory accuracy according to the length of experience with each object, but we nevertheless observed that most of the errors were made on the most recent objects. Specifically, we observed that Moka had a tendency to attach a preference toward the most recent objects he learnt, perhaps because they appeared more salient to him, which biased the results for a few trials. As these mistaken retrievals were not positively reinforced during the testing, Moka often decided not to retrieve these items anymore, and thus failed when these objects became the targets.

Conclusion

To conclude on that experiment, we found that learning the names of novel objects has to be spaced in time to be efficient, at least for a dog subject provided with a small receptive "repertoire". Successful retention was achieved after about 20-30 verbal exposures to the label and 20-25min visual exposures to the object, which is a very encouraging preliminary result given the reports of most previous studies from the literature on ostensive labeling. Nonetheless, it is far from the very few (or even single) learning events children require to demonstrate accurate retention of object names. A replication of this protocol with an expert dog would bring more insight in this field.

3.3 Which learning strategy is the most appropriate to induce a long lasting memory of an object name?

3.3.1 *Introduction*

The previous experiment involved an ostensive labeling learning method and revealed that at least three learning sessions spaced in time, each followed by a testing session, allowed an acceptable acquisition of the material to learn by our dog subject, Moka. But what accounted for the enhancement of the performance we observed from one session to another? Did Moka mainly **learn and consolidate the knowledge from the ostensive labeling learning sessions** or essentially **from the comprehension tests as feedback on his performances were delivered to him**? Results from our previous work strongly suggested that Moka did not learn anything from the very first ostensive labeling session. But in spite of this, does it mean that ostensive labeling was totally inefficient and that learning only arose from the reinforcements Moka received? It would be very interesting to understand the origin of the learning and the conditions required for memory formation in dogs. Finally, it would also help in understanding if Markman and Abelev's reluctance to accept Rico's fast-mapping abilities as compelling because of the reinforcements he received after the initial retrieval of the novel objects, is warranted (Markman & Abelev, 2004).

As the first step towards elucidating this issue, we developed an experiment that disentangled these two parameters. Specifically, **pairs of novel objects - one associated with a label while the other remained nameless** - were divided into three groups. Objects from the first group were taught to Moka using ostensive labeling and were then tested during a **comprehension test (L+T = Labeling + Testing)**. Objects from the second group were similarly taught to Moka but no tests followed the labeling sequence (**L = Labeling**). Objects from the third group were not taught to Moka but were directly incorporated within a comprehension test (**T = Testing**). The overall procedure was **repeated four times**. For each group, a **final test** assessed Moka's accuracy to map the labels to their corresponding items.

Our hypotheses were the following: if ostensive labeling does not induce learning at all, we expect Moka to perform at chance during the final test that involves the objects of the second group (*Labeling* group). On the other hand, if learning is only attributable to the reinforcements delivered during testing or to the testing effect per se (see Roediger & Butler, 2011 for a review), we expect the mean performance involving the objects of the third group (*Testing* group) to be equivalent to the one of the first group (*Labeling + Testing* group). Finally, if Moka is learning from both – ostensive labeling and reinforcements received during testing - we should observe a

higher performance for the final test that involves the objects of the first group since they were implicated in both learning strategies.

Next, we intended to know which learning strategy would be the most appropriate for the formation of a robust memory trace. Hence, for the three groups, we carried out retention tests after a delay of two weeks and one month.

3.3.2 *Material and methods*

3.3.2.1 *Stimuli*

Stimuli were novel objects, never seen and never labeled before. Objects were grouped in pairs according to Moka's owner appraisal about the "affinity" his dog would develop for each object. Specifically, objects belonging to the same pair had to be equally salient for the dog with no obvious biases during a pre-test of preference. During this test, the two objects were displayed on the ground for one minute. Moka spontaneously interacted and played with them. We scored the time Moka spent with each object and the number of interactions toward each object. If a clear preference for one object was noticeable, we discarded this pair of objects and started again with two novel objects.

3.3.2.2 *Procedure*

During an initial learning phase (t_0), Moka was introduced to six pairs of objects. For each pair, only one object was associated with a label (T=Target) while the other remained nameless (D=Distractor). The object that was attributed a label was randomly chosen among the two by a third person. The six pairs were not introduced to the dog during the same session because it would have been too tiring for him. Thus, these initial learning phases were scheduled over different days. During the initial learning phase, objects of each pair were presented to Moka five times each in an interspersed manner [Figure 80].

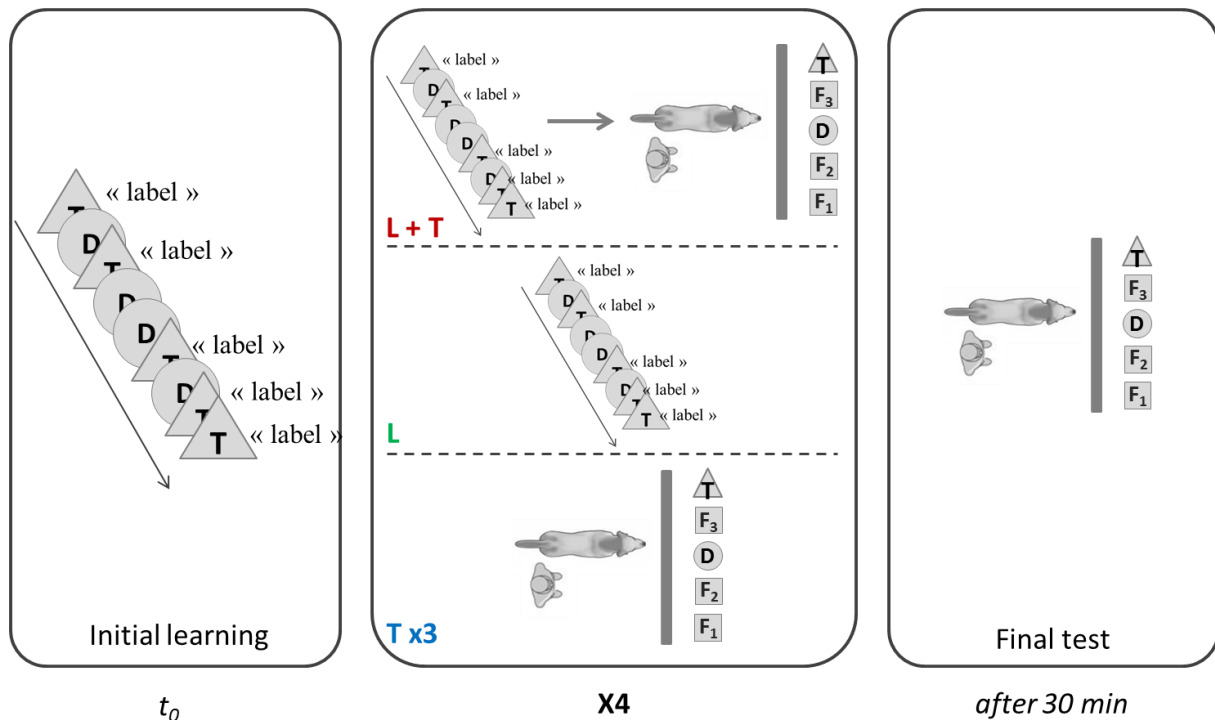


Figure 80 - Illustration of the experimental set up developed in this experiment. It consisted of an initial learning phase (t_0), where two novel objects were introduced to Moka in an interspersed manner. Only one object – randomly chosen by a third person - was associated with a label (T=Target), while the other was similarly introduced to Moka but was not labeled (D=Distractor). The six pairs of objects were divided into three groups: L+T= Labeling + Testing (i.e. learning rehearsal followed by a comprehension test (T requested three times during a test)); L=Labeling (i.e. learning rehearsal but no subsequent testing); T=Testing (i.e. no learning rehearsal, but only comprehension tests (T requested three times during a test)). For each learning condition, the procedure was implemented four times, spaced in time. A final test performed 30 min after last practice assessed Moka's accuracy of the mappings.

Then, the pairs were randomly assigned to one of the following groups (2 pairs per group):

- **L+T= Labeling + Testing** group: each learning session consisted of five presentations of the two objects of a given pair interspersed, during which only the target was accompanied by its label. Learning was systematically followed by a comprehension test, which occurred after a break of about 10-20min. Comprehension tests involved both objects along with three familiar items randomly chosen among Moka's set of known objects. During testing, among requests of familiar objects, Moka was instructed three times to bring the labeled object. Objects were replaced behind the panel after each trial. Moka received feedback of accuracy after each retrieval. If he mistakenly brought an untargeted object, he got another (or maximum two other) chance(s) to find the correct object.
- **L= Labeling** group: objects of each pair were presented to the dog five times each in an interspersed manner (again, only the target was labeled) as during the initial learning phase. In this condition, no testing followed learning.

- **T= Testing group:** objects from this group were never presented to the dog again and the target was never ostensively labeled anymore. The sessions directly started with comprehension tests that involved both novel objects and three familiar objects. Testing was carried out exactly as described above.

The overall procedure was repeated four times and a final test was carried out 30min after the last comprehension test (or after the last teaching trial for the *Labeling* group). The final comprehension test was conducted for each pair separately and was equivalent to the comprehension tests performed during training (though with different familiar objects).

Next, to evaluate whether Moka formed a robust memory of those word-object mappings, long-term retention tests were carried out after a delay of two weeks and one month without rehearsal between the time intervals [Figure 81]. The retention tests replicated those implemented during the final comprehension test but only one test trial was performed in this case, since it would have generated a rehearsal of the mappings.

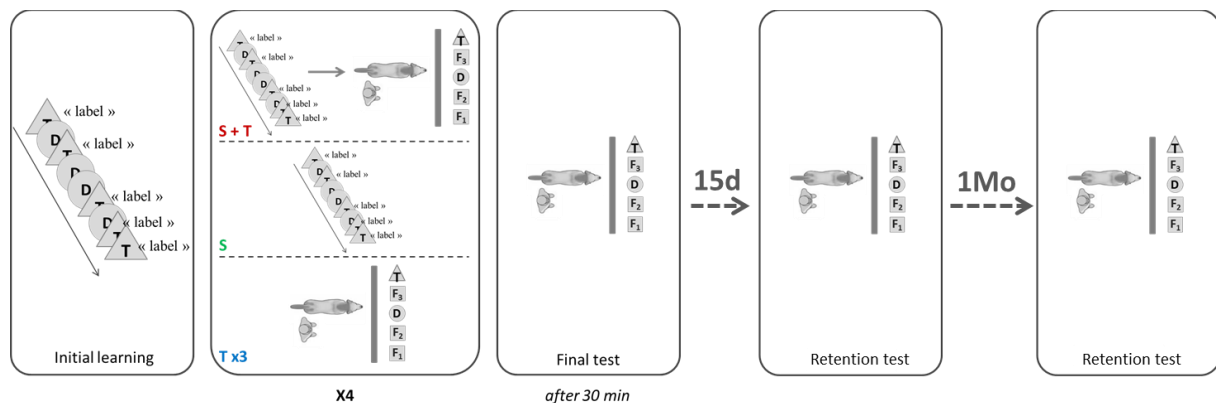


Figure 81 - Illustration of the whole experimental set up including the two retention tests, one carried out after a delay of two weeks (15d) and the second after a delay of one month (1Mo) from this last test. During these retention tests, the target was requested only once.

3.3.3 Results and discussion

Moka underwent three learning conditions, each involving two pairs of objects. Although results seemed at first glance relatively similar for the three groups (Mean(L+T)=66.66%; Mean(L)=49.5%; Mean(T)=50%, see Table 14 & Figure 82), the highest percentage of correct retrievals involved the objects from the *Labeling + Testing* group, which was the only group to reach significant results (exact binomial, $p=0.017^*$). Even if it is too preliminary to draw convincing interpretations, from this finding it can be speculated that Moka benefitted from both learning strategies to acquire the novel object names.

Table 14 - Table detailing the number of trials, number of successes, expected chance level, p-values and mean performance (in %) during the tests conducted after last practice (i.e. 30min), after a delay of 15days and after a month delay. The table reports the results for the three learning strategies (L+T=*labeling + testing*; L=*labeling*; T=*testing*).

Learning strategies	Latency	N trials	N successes	Expected chance level	P-values	Mean Performance%
L+T	30 min	6	4	20	0.017	66.66
	15d	2	2	20	0.04	100
	1Mo	2	2	20	0.04	100
L	30 min	6	3	20	0.09	49.5
	15d	2	0	20	1	0
	1Mo	2	0	20	1	0
T	30 min	6	3	20	0.09	50
	15d	2	1	20	0.36	50
	1Mo	2	2	20	0.04	100

However, since the objects of this first group were submitted to comprehension tests that provided feedback of accuracy, one could argue that Moka merely learnt the associations from the reinforcements themselves. This hypothesis was not clearly established since he correctly brought the targets especially on the first test trials (out of the three test trials). Nevertheless, future replications of this protocol would be required to strengthen these preliminary findings.

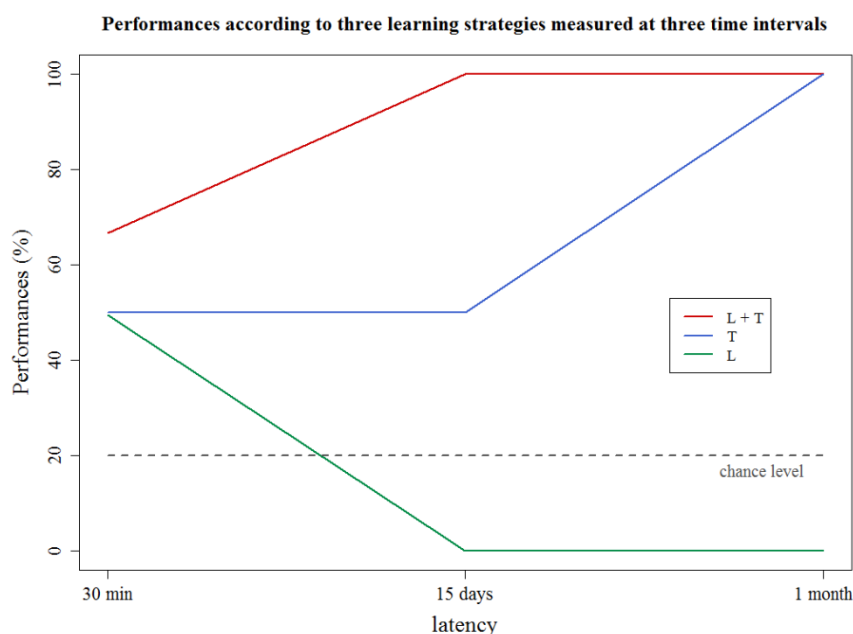


Figure 82 - Moka's mean accuracy in recalling the names of the objects measured after three latencies (i.e. 30min, 15d and 1Mo). Objects were taught using one of these three learning strategies: L=*labeling* (green curve); T=*testing* (blue curve); L+T=*labeling + testing* (red curve). Dashed line indicates the chance level.

When focusing on Moka's performance after longer delays, there are some points that are particularly interesting. While each learning strategy seemed to have at least slightly facilitated Moka's comprehension of the mappings when tested after a relatively short time interval after last practice (i.e. 30min), noteworthy differences arose after longer delays. Indeed, Moka failed to remember the names of the objects that were exclusively taught with ostensive labeling. After a

break of two weeks, his failure was not attributable to a higher preference for the distractor, because he rather fetched a familiar object in both cases. Interestingly, after a break of one month, analyses of the video recordings revealed that his choices went for the distractor. Hasty interpretations of this finding would be unwise, but it suggests at least that Moka did not always mix the target up with the unnamed distractor but also sometimes with objects that already had names. This finding is interesting given that some authors have pointed out the difficulty of choosing an appropriate chance level when facing this kind of testing configuration (Griebel & Oller, 2012). Indeed, they discussed the fact that, when the target is requested, the dog should ignore the well-known objects and only choose between the remaining options. Thus, in our case, should the probability of success by chance include the familiar items as potential competitors? Or should we apply a more conservative chance level value of $\frac{1}{2}$? This interrogation remains open since the mistakes did not always concern the distractor.

In contrast to Moka's incapacity to remember at long term the names of the objects that belonged to the *Labeling* group, Moka successfully retrieved one of the two objects taught exclusively with positive reinforcement after a two-week delay (*Testing* group), and both objects without error after a one month delay. He also never failed to retrieve the objects that were ostensibly labeled and positively reinforced during learning (*Labeling* + *Testing* group).

But did Moka truly retain the names of those objects? The criticisms of Kaminski et al.'s and Pilley et al.'s testing procedures could also apply in this case (Griebel & Oller, 2012). Indeed, nothing excludes the possibility that Moka reasoned by extended exclusion to complete this task. He could have solved the task by excluding the familiar objects at first because they already have names, and then the object that was not reinforced during the previous test trials. Moreover, as in children's word learning paradigms, the alternative option that he only retrieved the one of the two objects which was given special treatment cannot be ruled out. To put this alternative explanation to rest, it would have been interesting to pit the targets of each group against each other and to randomly ask Moka to retrieve each of them (see Grassmann, Kaminski, & Tomasello, 2012 for related methodology). Unfortunately, we didn't progress to this control test. We considered that the objects were, from that moment, properly acquired and we incorporated them into Moka's set of well-known objects that served for other protocols (e.g. "memory across time", "playbacks", etc.). We did not notice that Moka had greater difficulty with those objects compared to older ones. Thus, our personal observation and intuition goes in favor of an actual memory formation of these pairs but further replications with stringent controls are still necessary.

Conclusion

Taken together, from this experiment, it can be proposed that a dog with a small “vocabulary lexicon” like Moka, does not take advantage of an ostensive labeling learning method, at least not to form a long lasting memory of a novel word-object pair. On the other hand, the sole reinforcing consequences of the positive feedback received after retrievals (or simply the retrieval effect) had no evident impact at a “relative short term” (~30min) on the acquisition of novel word-object associations. However, reinforcing the correct behaviors during learning seems to be a key component for long term memory formation. Further replications of this experiment in addition to supplementary investigations about possible extended reasoning by exclusion strategies, should enlarge our understanding of this phenomenon. This would constitute a major advance for dog trainers and dog owners. Finally, these preliminary findings do not challenge Rico’s and Chaser’s long term memory abilities to remember the mappings they acquired by logical exclusion one month earlier. These expert dogs might have formed the associative links in one trial even without the need of the reinforcements that were provided. Needless to say that replicating our design with expert dogs would undoubtedly help better understand the mechanisms underlying such cognitive processes.

4 FOURTH STUDY. *Generalization of Moka’s knowledge to novel stimuli*

4.1 Introduction

One poignant argument in favor of actual word learning abilities by dogs would come from the capability in generalizing their knowledge to novel instances of voices or objects. Bloom stressed the possibility that Rico’s referential understanding was limited to specific routines and to distinct stimuli (Paul Bloom, 2004). Moreover, he argued that if a dog really comprehends word-object mappings, it should not matter who the speaker is. Few studies started to bring light to these issues.

First, Pilley & Reid provided evidence that their dog, Chaser, was not only learning labels as proper nouns to designate specific referents, but that she could also learn labels for common nouns, like categories (Pilley & Reid, 2011). She understood at least three labels for categories, namely “toy” (i.e. objects she was allowed to play with, in contrast to objects she was forbidden to play with), “Frisbee” and “Balls” (i.e. two subcategories of her toys determined by specific physical features, like shape). Thus, she was able to map one label onto many objects that she

could discriminate based on common physical properties or functionality (for the “toy” category). Moreover, it illustrates that she could also learn up to three different labels for the same object that referred to different levels of categorization (e.g. the name of a proper Frisbee (i.e. item-unique level), the word “Frisbee” (i.e. basic level), and the word “toy” (i.e. subordinate level)). Chaser’s demonstration of one-to-many and many-to-one word-object mappings was a clear proof of her extensibility skills, which contradicts Bloom’s comment about dogs’ having learning abilities restricted to an item-unique level (Paul Bloom, 2004). Moreover, it has been shown that in order to generalize their knowledge to novel items, dogs presumably rely on size properties when only briefly familiarized with a novel word-object pair and on texture when familiarized with a pair for a longer time (Van Der Zee et al., 2012). More recently, another study demonstrated that a dog could also generalize their knowledge of words to novel voices (Griebel & Oller, 2012). The authors reported the case of a dog who successfully retrieved objects by names pronounced by novel speakers. One speaker was a female with a German accent while the other was a male, native American English speaker with a Californian accent. This finding illustrated that this dog was capable of extensive verbal recognition, as she recognized words in her repertoire even when pronounced by novel voices (see also the results of Grassmann et al., 2012 for familiar labels requested by an unfamiliar person). This research lends weight to the groundbreaking findings of Kaminski et al. and Pilley & Reid (Kaminski et al., 2004; Pilley & Reid, 2011). It also directly replied to Bloom’s reluctance to accept Rico’s word learning abilities as compelling because he only received verbal commands from his owner.

In the following study, we wanted to investigate to what extent a dog provided with a small word-object “lexicon” would be able to generalize his knowledge to novel voices or novel objects. Moreover, a dog’s ability to recognize words pronounced by natural and artificial voices recorded beforehand and delivered through loudspeakers has never been studied. It would be very interesting to know if a dog can recognize words originating from various speakers, contexts, and devices (like for example from a TV clip). This would tell us a lot about his actual word learning skills. Furthermore, if he is accurate with such a device, it would strongly suggest that neither visual nor auditory cues are provided by the owner to help him solve the task.

To explore those questions, we compared Moka’s accuracy to retrieve objects by name when the verbal instructions were given (i) by the owner, (ii) by a playback of the owner’s voice, (iii) by a playback of a female’s voice and (iv) by artificial playbacks generated by synthetic voices of females and males. Next, we also examined his ability to generalize his knowledge about word-object mappings to novel objects.

4.2 Materials and methods

The dog subject was Moka, a hunting dog of one year old, provided with a “lexicon” of about 15-20 word-object pairs when this experiment started.

Moka’s performance to retrieve objects by name was measured using a tightly controlled method as described in the previous sections. For each *audio generalization* condition, five familiar objects were placed beside each other and Moka was randomly requested to bring each object upon command (each object was requested only once). A generator randomly selected the target object at each trial. Objects were not removed from the set after each trial (i.e. with replacement). Thus, the probability of success by chance was 1/5 on each trial. The experiment was carried out during several succeeding sessions, and sessions involving different conditions in an interspersed fashion. In this way, the task did not gradually increase in difficulty over the sessions, which would have biased the results since Moka would have become accustomed to the playback apparatus. In the case of the *object generalization* condition, we replicated the experiment twice, one involving four novel objects and the other three. Thus, respectively four and three novel objects were pitted against each other. As they all had the same history of novelty, the probability of success by chance was 1/4 (for four trials) and 1/3 (for three trials), thus 0.27 on average. .

The five testing conditions were as following:

- Verbal instructions provided by the owner. The owner stood next to his dog and verbally requested the object designated by the random generator at each trial. Five replications of the experiment were carried out, involving different objects each time. The mean performance obtained through this condition corresponded to Moka’s baseline of accuracy (i.e. familiar label baseline).
- Verbal instructions provided by playbacks of the owner’s voice. The voice of Moka’s owner was recorded prior to the experiment. An audio file for each object label was generated. At each trial, the generator randomly selected a target and launched the corresponding tape recorded playback. The transcribed verbal request originated from the loudspeakers connected to the computer. The owner stood next to his dog in order to preserve a similar testing configuration. At each trial, Moka was expected to bring an object back into his owner’s hand. As usual, the owner gave feedback of accuracy. This experimental condition was replicated four times, involving different sets of objects each time.
- Verbal instructions provided by playbacks of a female’s voice. A female native French speaker (unfamiliar to Moka), with a southern French accent recorded herself uttering the names of the objects. The same procedure as described above was applied: the owner stood next to Moka, and at

each trial the generator randomly selected the target and launched the corresponding audio file. Four replications of this experiment were performed, each time with different sets of objects.

- Verbal instructions provided by artificial playbacks of synthetic females' and males' voices. Artificial playbacks were generated with the multilingual vocal synthesizer 'Hear it first' software, hosted by Microsoft. The male and female voices were accent-less and with a robotic-like consonance. The experiment was carried out as described for the previous conditions. Six replications of the experiment took place, 3 involving male synthetic playbacks, and 3 female synthetic playbacks.
- Novel objects provided with resembling physical features. In this condition, seven novel stimuli that resembled Moka's original objects were found. These copies shared at least one physical feature with the original objects: size, shape, color or texture. Only copies were placed beside each other during testing. Thus, Moka had never seen either of them before the test. Two replications of the experiment were carried out, one involving four objects and the other the remaining three objects. Each object was requested only once. Moka's mean accuracy on this task was compared to the one he reached with the original objects. This corresponded to the baseline of accuracy (i.e. familiar objects baseline).



Figure 83 - Illustration of the original items (A, B, etc.) and their corresponding replicas (A', B', etc.) used from the *object generalization condition*. Replicas that were correctly selected at first trial are ticked with a green check mark. Replicas incorrectly retrieved at first trial are marked by a red cross. Replicas that were correctly retrieved at second trial were C' and E'.

4.3 Results

Moka's accuracy to map labels to their corresponding referents was highly significant when the verbal commands were uttered by his owner (exact binomial, $p=2.66e^{-7***}$, **Table 15** & **Figure 84**). This control condition served as a baseline of accuracy. Moka similarly performed highly when the requests originated from transcribed playbacks of his owner's voice (exact binomial, $p=1.85e^{-6***}$, **Table 15** & **Figure 84**). No significant difference was established between these

two conditions (GLM, $Z=0.14$, $p>0.05$). However, Moka's mean accuracy drastically dropped when provided with playbacks of either a novel female voice (♀Pb) or of synthetic voices (SynPb) [Table 15 & Figure 84]. In both cases, the mean performance was significantly below Moka's baseline of accuracy (GLM, $Z_{\text{♀Pb}}=-2.16$, $p=0.03^*$; $Z_{\text{SynPb}}=-2.5$, $p=0.02$) and did not significantly differ from chance level (exact binomials, $p_{\text{♀Pb}}=0.098$ and $p_{\text{SynPb}}=0.17$ [Table 15 & Figure 84]).

Table 15 - Moka's performance according to four *audio conditions* and two *object conditions* are reported. For each audio type and each object type, the total number of trials, total number of successes, expected chance level, p-value, mean performance (in %) and 95% confident intervals are detailed.

Conditions	N trials	N successes	Expected chance level	P-value	Mean performances%	CI95%
Owner (male)	25	17	20	2.66e^{-7}	68	46-85
Owner's playbacks	20	14	20	1.85e^{-6}	70	46-88
Unknown female's playbacks	20	7	20	0.098	35	15-59
Synthetic playbacks	30	9	20	0.17	30	15-49
Original	15	14	27	1.2e^{-7}	93	68-99
Copy	7	3	27	0.39	42	9-81

Next, Moka's generalization ability was also measured for novel objects. Moka was highly accurate to find the original objects upon command ($M=93\%$, $p=1.2\text{e}^{-7***}$, Table 15 & Figure 84). When tested with copies that resembled the original items, his mean performance significantly declined from this baseline (GLM, $Z=-2.27$, $p=0.02^*$). In this case, Moka's mean accuracy did not significantly differ from the mean level expected by chance (adapted binomial test, $p=0.39$, Table 15). Note that this condition corresponded to only seven trials (since only 7 copies of the original objects were found) and that more trials should be conducted to truthfully conclude on this result.

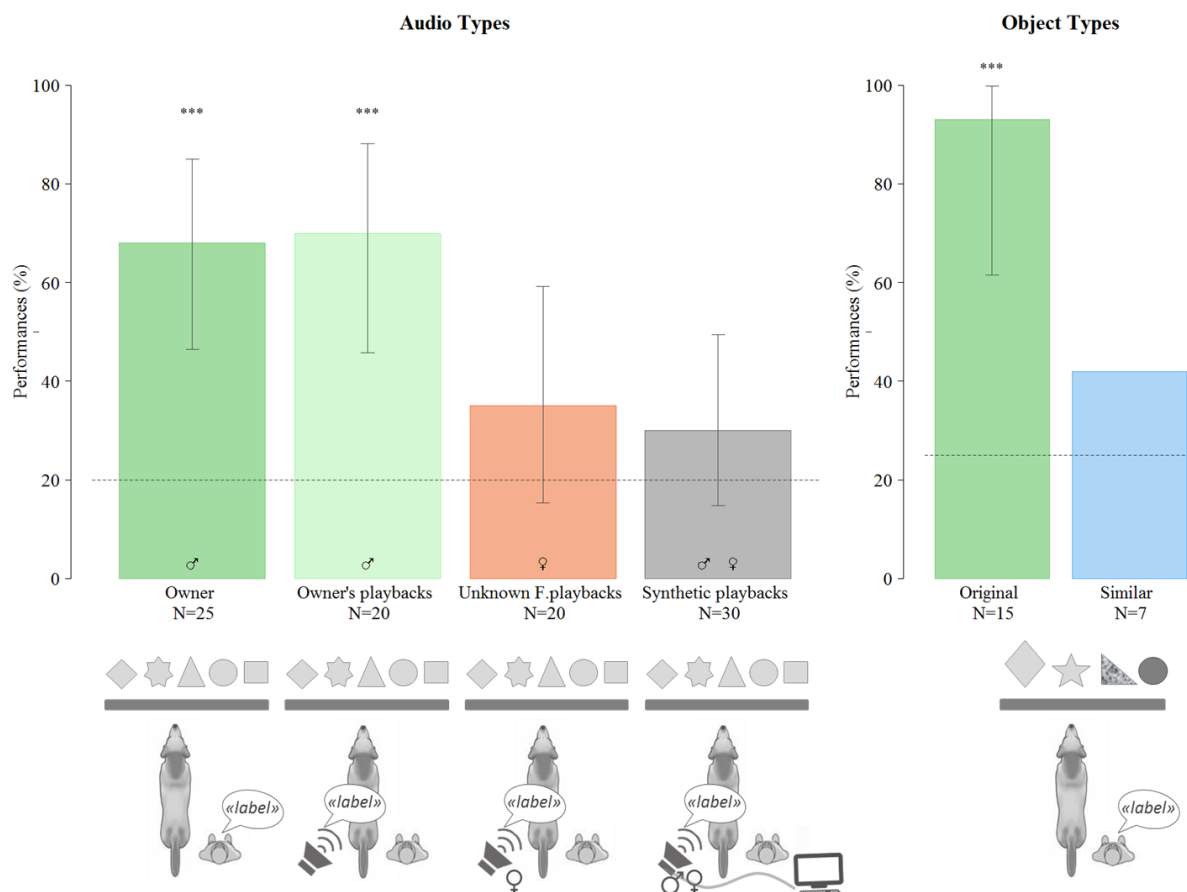


Figure 84 - Graphical showing Moka's mean performance \pm 95%CI in retrieving objects by name when the verbal commands were given by his owner (i.e. familiar label baseline) (left); by playbacks of his owner's voice; by playbacks of a novel female speaker's voice and by synthetic playbacks generated by a multilingual vocal synthesizer software. Moka's performance to generalize his knowledge to copies that resemble the original objects but that have never be seen before was measured and reported on the right side of the graphical (familiar objects baseline and copy condition). Illustrations of the experimental settings are drawn below each bar plot. Dashed lines represent the expected chance levels.

4.4 Discussion

This study aimed to better understand whether dogs are learning stimulus-specific associations or labels for larger categories of objects that share physical features. Moreover, this work intended to foster previous investigations about dogs' ability to extend their knowledge to sound patterns that share close auditory features with the original words.

Results from this experiment demonstrate that Moka performed just as well when confronted with playbacks of his owner's voice than with his owner himself. This finding entirely rules out the possibility that the owner was providing Moka with subtle auditory cues when asking him to fetch the objects. Indeed, in this testing configuration, Moka's trainer was admittedly present but remained completely neutral and silent. A previous study evaluated the effect of varying non-verbal features, like the distance of location of the trainer, the posture of the trainer or the eye contact and it showed that it moderated dogs' responsiveness to the commands and sometimes

affected their willingness to complete a well-known task (Fukuzawa et al., 2005). Considering these findings, we kept the testing configuration that Moka was familiar with (i.e. the owner standing next to his dog). Nevertheless, since the experimenter located the objects behind the panel himself and since the generator randomly selected and launched the auditory playbacks at each trial, the owner was totally blind about which target object would be chosen and where it was situated. This attests that Moka could only rely on his mental representations of the mappings to achieve the task.

However, when hearing playbacks from a novel speaker or from artificial voices, Moka failed to recognize the words. His mean performance was similar to the level expected by chance and significantly below the score obtained in the baseline condition. Our results are in agreement with previous studies that used tape-recorded verbal commands and that similarly established a significant decline of the performance for a well-known task (e.g. Fukuzawa et al., 2005). One explanation might arise from the device itself. Such apparatus involving loudspeakers and transcribed voices probably goes beyond ecologically valid conditions, which might have disturbed our dog subject. Personal observations suggest that Moka tried to locate where the sound came from and seemed disturbed by the unusual configuration of the task. It can be speculated that the absence of certain key non-verbal cues (e.g. lip movements) affects vocal perception in dogs as it does in humans (e.g. Moore, 1989). In addition, certain features of the glottal source signal are absent in a tape recorded voice, resulting in a sound that will not have the same frequency composition, or harmonics, and will lack resonance compared to a human-generated vocal signal (e.g. Howard & Angus, 1996). Nevertheless, as Moka provided convincing aptitudes when exposed to playbacks of his owner's voice, a second more likely explanation would suggest that he truly did not recognize the words. The intonation of the labels provided by both the female and the synthetic playbacks was clearly different from the pronunciation of Moka's owner. He might not have distinguished the words and therefore attributed a random object to them. It is worth remembering that Moka's word-object repertoire is greatly below those of previous dog cases reported in the literature. It is possible that a dog requires a longer experience with the concept and a larger "lexical repertoire" to develop such generalization skills. Authors asserted that cognitive abilities depend upon previous learning experiences and/or past exposures to demanding training regimes (Herman & Wolz, 1984). The dog that showed aptitudes to retrieve objects by name verbalized by novel speakers had a 120 words "lexicon", acquired in twelve years (Griebel & Oller, 2012). Originally, we did not plan to test an equivalent condition involving an unknown live speaker, but considering our current results, it would have been interesting to carry out this condition as well. This would help understanding if, despite the current results, Moka still has auditory generalization abilities. Undoubtedly, it would also be

interesting to replicate this whole experiment with an expert dog in order to uncover if the absence of convincing generalization aptitudes established here are primarily attributable to the short history of familiarization with the word-referent concept that Moka suffered from.

In parallel, we also examined his ability to generalize his knowledge to novel objects that resembled the original ones. Since most of the objects were obtained from second-hand trades, gathering copies that shared at least one physical property was challenging. We managed to collect seven objects that resembled the ones Moka was trained with for months. Thus, testing consisted of two sessions, one involving four copies and the other three copies. Moka successfully retrieved three out of these seven copies at first trial. For two objects incorrectly retrieved, he automatically adjusted his response on the second trial he was offered. Together, even so Moka did not demonstrate conclusive object generalization abilities, he succeeded in finding five out of the seven novel objects at first or second trial. This result is very encouraging considering the notable difficulty of the task. Indeed, in this configuration, Moka was facing four (and three) novel objects he had never encountered before. It means that at each trial, he had to figure out which object best matched his mental representation of the targeted association. Although it has been shown that dogs are skillful at using life-sized or miniature replicas to identify the object a human wanted them to fetch (Kaminski, Tempelmann, Call, & Tomasello, 2009), such a design had never been used before and deserves to be replicated with more objects to fortify our preliminary results.

Additionally, if this ability could be straightforwardly confirmed, it would also clearly demonstrate that dogs do not rely on odors to solve this kind of task but well and truly on visual features. It has not been clearly established yet that dogs from previous studies (as well as Moka) relied on visual features rather than on scent cues when completing their tasks (see Kaminski, Tempelmann, Call, & Tomasello, 2009's discussion for a related discussion). Although a considerable increasing number of studies assessing canine cognition utilize paradigms that are heavily reliant on vision, little is known about how dogs visually perceive their external environment. In general, **the visual system of dogs appears to be worse than that of humans** (see Byosiere, Chouinard, Howell, & Bennett, 2017 for a review). First, dogs appear to be more scotopic than humans, since they are especially adapted to dim light conditions and **less sensitive to color perception**. Dogs have two types of cone photoreceptor cells (against three for humans) that typically correspond to short-wave (a spectrum of blue) and long-wave (a spectrum of yellow) sensitivities. Moreover, 3% of dogs' retina cells are cone photoreceptor cells in contrast to roughly 5% in humans (Byosiere et al., 2017). Nevertheless, dogs' ability to distinguish different colors remains controversial and studies often provide conflicting results. As evidence, some

studies have suggested that dogs have dichromatic color vision and have trouble perceiving differences between green, yellow and red color cues (Neitz, Geist, & Jacobs, 1989), although others have claimed appropriate identification of all primary colors by dogs (e.g. Rosengren, 1969; Tanaka, Watanabe, Eguchi, & Yoshimoto, 2000). Secondly, it has also been estimated that dogs' visual acuity is around 20/75, meaning that from 20 feet away, a dog could perceive an object that a human being could differentiate from 75 feet away (Miller & Murphy, 1995). Such diminished visual acuity possibly reflects a higher difficulty in discriminating fine details by dogs. Currently, there is no consensus about the perceptual cues that dogs utilized to differentiate between objects, but being accurate when faced with a sample of copies of objects never encountered before, can reasonably discard the assumption that they relied on odor cues.

Conclusion

This study shows that **Moka is successful on this task when confronted to his owner's voice or to recordings of his owner's voice**, suggesting that **no visual nor auditory cues influence Moka's decisions when solving the task**. Nonetheless when confronted with **unknown recorded voices**, Moka **failed to recognize the associations**, which is likely due to the apparatus or to a default of generalization abilities to novel auditory instances. Finally, he demonstrated **encouraging generalization abilities to novel visual instances** that shared at least one visual property with the trained object. Further replications of this finding would confirm that dogs are capable of relying on visual cues despite their presumed lower visual acuity.

5 GENERAL DISCUSSION

In the last 20 years, a myriad of research has been conducted and uncovered a variety of behaviors and abilities that make dogs an important model for studying cognition. This can be essentially attributed to their unique social cognitive aptitude and comprehension of human communicative intentions. For instance, dogs outperformed all other non-human animal species in locating hidden food based on human-given cues (e.g. Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Hare et al., 2002; Kaminski, Tempelmann, Call, & Tomasello, 2009). Dogs have also been shown to understand human words, a complex cognitive trait that served as the basis for our research on canine long-term memory formation. Since very little research has examined if and how animals form long lasting memories, the central issues of this thesis were 1) to **investigate if a non-primate species do create memories for complex stimuli that can last in time**; 2) to **characterize the precise conditions required for a long-term memory to be formed** for this animal model and for this specific cross-modal paradigm and 3) to understand **whether these**

conditions and learning strategies are similar to those currently established in human adults.

To address these questions, we recruited 40 dogs that were trained on the word-referent concept during a six month period. Only about half of them passed the training phase and underwent the comprehension and retention tests. Then, the best performing dog of the cohort was subjected to a series of exploratory experiments which started to shed light on those issues and truly constituted an avenue for future research.

Overall, our results revealed that **the word-object concept is a complex task for dogs**, since only a very few of them reliably acquired it. Acquisition of such a demanding task appeared to be **age-sensitive** in our dog sample, which was reflected both **in terms of accuracy and number of assimilated stimuli** during the allocated time. Nevertheless, we provided evidence that **dogs are capable of creating long-term memories for this knowledge**, as they significantly remembered the mappings after a delay of three months without rehearsal in between. This underpins previous findings showing that animals can form a sensory memory or a memory for an abstract task that can last in time (e.g. Burdyn et al., 1984; Demant et al., 2011b; Johnson & Davis, 1973; Patterson & Tzeng, 1979; Reichmuth Kastak & Schusterman, 2002; Wallis et al., 2016). However, our work fosters our understanding about memory formation in animals, since it demonstrates that a non-primate animal can store and retrieve an entire knowledge such as “semantic-like” information. Indeed, dogs not only recognized the objects and the auditory labels, they also remembered the relational links between objects and words, suggesting that they **created internal representations of this knowledge to solve the task**, a point that will be discussed later on. Nevertheless, as for humans, we observed an **age-dependent decline in remembering the information in the long term**. This suggests that dogs share intrinsic mechanisms underlying memory formation and recall with humans.

Next, we showed that once a dog is familiar with a task, learning novel instances related to this task, is **reliant on the number of repetitions**. In contrast to humans, a dog may not be capable of acquiring “semantic-like” information after a single learning trial. **In dogs, learning seems to be a repetition-driven process, at least regarding arbitrary cross-modal information**. Our study also supports the idea that **repetition must be distributed in time for learning to be efficient** and for a memory trace to be formed. We did not directly test whether the memory strength increased linearly with the number of presentations encountered during learning, but basically the longer the dog was familiar with an item, the better he retrieved its associated name during the subsequent memory tests.

Additionally, our results strongly suggest that **a dog with a relatively small word-object “lexicon” is not learning from ostensive labeling**. However, repeated sessions of ostensive labeling spaced in time induce acquisition of novel information but not the creation of robust memory traces lasting in time. Conversely, our pioneering experiments propose that the **consolidation process primarily relies on the reinforcements provided during the retrieval exercises and/or to the retrieval effect itself**.

Finally, we showed that a dog with a fairly recent history of familiarization toward such an abstract concept is already **able to demonstrate reasoning by exclusion skills** and to exhibit encouraging capacities of generalizing his knowledge to novel stimuli that share similar visual features. However, he did not seem capable of extending his knowledge to novel voices that were tape-recorded and generated by loudspeakers. Dogs from previous studies who were greatly accustomed to this concept demonstrated overall larger generalization skills (e.g. Griebel & Oller, 2012; Pilley & Reid, 2011; Van Der Zee et al., 2012). Thus it can be assumed that **generalization is mostly possible when the subject is provided with a substantial experiential background**, in other words with an important expertise in the task domain.

In order to generalize, an animal should be **able to decontextualize the knowledge** he acquired during training and to extend this knowledge to novel situations, stimuli, contexts, etc. If so, it can be claimed that the given animal **developed a semantic-like memory** of that information and that he **created an internal representation of that information**.

In this discussion, I would like to debate these two assumptions, taking account of the word-object pair concept. First, I questioned myself whether word learning by dogs reflects a real processing of words in a human sense and with a semantic-like component. And second, I wondered whether dogs create internal representations of object names which would inform us about the underlying type of memory.

In 1984, Herman et al. defined a word as “a unique, independent semantic entity; entities are agents, objects, action or modifiers of place and direction” (Herman & Wolz, 1984). Findings from previous studies in addition to the preliminary empirical evidence provided by our work represent a giant leap in the understanding of word learning by dogs. It already replied to some of Markman & Abelev’s and Bloom’s concerns (Paul Bloom, 2004; Markman & Abelev, 2004) but future investigations are still needed to better comprehend the mechanisms underlying word learning by dogs. **Do dogs learn the meaning of a specific word on the basis of inherent cognitive processes, of low-level associative processes induced by the reinforcements, or of a combination of both?** Intensive training combined with systematic reinforcement may produce

the desired behavior in the end but maybe not any understanding of language with a sense-making component. If so, succeeding in mapping words to items or to actions would have, at first glance, little to do with word learning.

Nevertheless, in an admittedly more natural way, a newborn human will also be similarly exposed countless times to the objects of their environment that will be labeled in a temporal contiguity before grasping the sense of this word and the fact that the two agents are related. As reviewed in the previous part of this thesis, in preverbal infants, initial word learning is characterized by a slow process that requires multiple repetitions combined with co-occurring non-verbal gestural cues (e.g. pointing to the object, holding the object, gazing at the object, etc.) and implicit reinforcement (e.g. mother's smile). During their 6th to 8th month of life, infants acquire the referential understanding that words can refer to objects, actions, people, properties, etc. Thereafter, what makes the main difference, as well-described by Bloom, is the wide range of vocabularies children acquire and the amount of novel words they learn per day, even from overhead speech or when nobody is trying to teach them (Paul Bloom, 2004). Obviously, even an expert word-trained dog will never reach such a level of aptitude. However, Bloom acknowledged the possibility that **Rico's learning of object names may be qualitatively similar to that of a child, differing only in degree, not in kind.**

To illustrate this idea, let's take the example of a farmer who trains his dog to work with him in herding sheep. When let loose with the sheep herd for the first time, the dog will spontaneously initiate behaviors. On the basis of positive or negative vocal feedback given by the farmer, he will progressively learn to infer correct behaviors. Additionally, the farmer will label the expected behaviors each time his dog will perform them (e.g. "turn right", "stop here", etc.) so that in the course of time, the dog will achieve a distinct behavior upon command even without the farmer in view. Consequently, the dog learned merely from daily experience without intensive explicit training. Thereafter, he may also appropriately respond to the words in various contexts and perhaps also if pronounced by different people. This greatly suggests that the dog gradually and naturally assigned a semantic-like meaning to each distinct word, exactly as defined by Herman.

To promote this assumption, one empirical study demonstrated a dog's combinatorial understanding of words referring to objects, namely a preposition object, a verb and a direct object (e.g. "to ball take Frisbee") (Pilley, 2013). Since the dog was similarly successful with objects never encountered before, performance could not be explained by rote memorized strings of words. In contrast, it clearly demonstrates that this female dog could appropriately extend her knowledge to novel instances and contexts without additional learning.

Furthermore, **inferential reasoning abilities by dogs go also beyond basic associative processes**. Indeed, selecting a referent by exclusion corresponds to “the selection of the correct alternative by logically excluding other potential alternatives” (Aust, Range, Steurer, & Huber, 2008). Thus, such choice responses cannot be based on mere associative learning mechanisms because the name and the referent haven’t been presented together in this single trial prior to the inferential selection. Alternatively, it clearly suggests that dogs are endowed with some inherent word learning principles to succeed, as described previously with children. For instance, they must at least have acquired the “**whole object**”, “**mutual exclusivity**” and “**novel-name-nameless**” principles (Baldwin & Markman, 1989; Grassmann et al., 2012; Mervis & Bertrand, 1994).

Lastly, dogs’ integration of the word-object referential concept seems **facilitated by an ease of interpreting verbal and gestural information from humans**. For instance, dogs are sensitive to co-reference of pointing and naming in order to determine the referent of a spoken word, but only if pointing does not conflict with their previous knowledge of a given word (Grassmann et al., 2012). This skill is close to referential understanding of the co-occurrence of multimodal referring cues observed in babies (e.g. when a mother points to an object and says “look, this is a cup”). Similarly, dogs seem to understand the representational nature of objects since they successfully decode new iconic communicative signs provided by humans to identify desired objects, often from the first trial. Since apes require extended training to accomplish this type of task, it is likely that only dogs are able to grasp the communicative frames of humans which explains the flexibility and rapidity with which they acquire new abstract tasks (Kaminski et al., 2009). This strengthens the hypothesis that word learning by dogs is perhaps not merely reflected by a relatively simple instrumental response but that some inherent cognitive processes (probably ancestral or inherited via the domestication process) drives their exceptional comprehension of human language.

In sum, in my opinion, dogs learn the word-referent concept mainly by **means of positive and negative reinforcement** (either implicitly for dogs working daily with humans or explicitly for specific purposes) but are **yet equipped with probable ancestral or inherited intrinsic principles** that drive their behavior and **facilitate the acquisition of such abstract concepts**. I also suggest that **once a dog is capable of generalizing and decontextualizing**, the knowledge they have been trained on acquires a **semantic-like component**. In my opinion, this is only possible if the **animal developed internal representations** of the abstract concepts and their related contents.

This leads us to the second point I wanted to discuss here, namely the possibility that dogs create mental representations of sensory information.

For nonhuman animals, concept learning is considered as the highest degree of abstraction attainable (Reichmuth Kastak & Schusterman, 2002). Concepts can be centered on common perceptual, relational or associational properties. Perceptual concepts include categories of stimuli that are perceptively similar, relational concepts are based on common abstract relationships shared by sets of stimuli (e.g. sameness or oddity concepts), whereas associative concepts are formed when disparate arbitrary stimuli are linked together on the basis of common relations (Reichmuth Kastak & Schusterman, 2002). A conceptual behavior is demonstrated when an animal appropriately responds to novel problems or situations on the basis of the common properties/relations gained via prior experience (either implicitly or explicitly as discussed earlier). In order to overcome such novel situations efficiently and more rapidly, one could claim that the animal must have relied on internal representations of these common properties/relations.

Fortunately, one study examined this proposal. To test this hypothesis, researchers have modified methods originally developed for research on preverbal infants to investigate mental representations in dogs. As described in the previous part of this thesis, studies demonstrated that human infants would look longer at unexpected stimuli compared to familiar ones if displayed simultaneously. In a similar dog-designed paradigm, Adachi et al. presented a photograph either of the dog's owner face or of an unfamiliar person's face, after playing back a voice that either matched (congruent condition) or mismatched (incongruent condition) the face stimulus (Adachi, Kuwahata, & Fujita, 2007). The results demonstrated that when the displayed face contradicted the auditory stimulus, dogs spent significantly more time looking at this incongruent photograph compared to the looking time toward the face in the congruent pairings. The authors suggested that dogs had expectations about visual events based on auditory information. Dogs may have activated internal representations of their owner's face when they heard their owner's voice. If so, specific brain regions allocated to internal representations of sensory information could exist.

In humans, the neural regions responsible for detecting auditory novelty and for discriminating between competing auditory sensory stimuli (studied by means of oddball paradigms) include the primary auditory cortex along with a broader attentional network (e.g. Cacciaglia, Escera, Slabu, Grimm, & Sanjuán, 2015; Linden et al., 1999). Moreover, a typical ERP (event-related potential) component called the mismatch negativity (MMN), a negative deflection that peaks approximately 160-220ms post novel stimulus onset, is believed to be related to discrimination of unattended stimuli (e.g. Näätänen, Paavilainen, Rinne, & Alho, 2007 for a review). Using a minimal invasive procedure (multi-unit intracranial recordings), one study demonstrated a

negative peak of the ERP at 180ms after infrequent deviant stimuli in dogs, but not after standard stimuli (Howell, Conduit, Toukhsati, & Bennett, 2012). This result suggests that dogs detected a difference between the tones and processed the discriminable changes in auditory stimuli with relatively similar mechanisms as humans. Furthermore, one fMRI study revealed a hemispheric bias for praise words versus neutral words, suggesting for the first time, semantic-like processing of auditory stimuli by dogs (although there was a subsequent correction in which right and left were reversed) (Andics et al., 2016).

Very recently, researchers probed the neural basis for auditory discriminations, and more specifically for **internal representations of trained words**, using an **awake-fMRI approach** (Prichard et al., 2018). This study is particularly influential for our understanding of representations of object names by dogs given that dogs were trained to select objects based on their associated labels. Dogs had to appropriately discriminate between two trained objects (with 80% correct responses for one object and 50% or above for the other) before being tested using an MRI scan. The MRI protocol consisted of an event-based design that included four trial types: (1) trained label-expected object; (2) trained label-unexpected object; (3) pseudoword-novel object and (4) random reward trials. The authors described **auditory novelty detection areas in the domain of human speech**; specifically greater activation in the **right parietotemporal cortex** to novel pseudowords relative to trained words. The authors argued that such novelty detection may be explained either by less frequent exposures to the pseudowords (i.e. similarly to the oddball paradigm) or to a lack of meaning associated to those pseudowords (i.e. lexical processing). In humans, meaningful words typically result in higher activity within the so called “semantic network” relative to non-evocative pseudowords (Binder, Desai, Graves, & Conant, 2009 for a review). This semantic network’s anatomical location is ill-defined in humans and is still a matter of debate since no consensus has been reached (see the previous part of this thesis). A meta-analysis of 120 functional neuroimaging studies revealed that this left-lateralized network comprises at least seven regions, namely the posterior inferior parietal lobe, the middle temporal gyrus, the fusiform and parahippocampal gyri, the dorsomedial prefrontal cortex, the inferior frontal gyrus, the ventromedial prefrontal cortex and the posterior cingulate gyrus (Binder et al., 2009). The entire network is sometimes called the temporoparietal junction (TPJ) or temporal-parietal-occipital cortex (TPO). In the fMRI study in dogs, follow up MVPA analyses identified regions involved in the discrimination of the two trained words, namely the **posterior thalamus**, the **amygdala**, the **left TPJ**, the **left dorsal caudate nucleus** and for seven dogs, the **left temporal cortex**. The authors concluded that such areas are remarkably similar to the **conceptual component of the “semantic network” in humans**. Although these regions were of marginal statistical significance, they could correspond to potential **sites for receptive word processing** in

dogs. If so, this finding goes in favor of semantic representations in dogs. Unfortunately, the authors did not include a condition in which a trained object followed a pseudoword, or a condition in which a trained object was mismatched to a trained word. This would have provided supplementary neuroimaging evidence for violation of expected semantic content and would have furthered the possibility that dogs do create internal representations of sensory information.

To conclude, I think that these innovative findings lend weight to the hypothesis that the remarkable abilities of dogs to know tens, hundreds or thousands of object names **truly depend on mental representations** of the arbitrary mappings. It also suggests that dogs have access to that information when needed by reactivating their internal representations. If so, very interesting hypotheses can be advanced about the underlying type of memory and underlying mechanisms. Indeed, if dogs really create mental representations and if the long-term memory performance we found can truly be explained by an access to these representations, this would appeal to structures **analogous to the human medial temporal lobe, rather than brain regions involved in procedural memories**. In other words, this would fortify the assumption that retrieving objects by name from memory is not the consequence of a low-level association between an item and a reward generated across repeated trials, but truly the **formation of unitary integrated concepts**. If so, it can be claimed that arbitrary word-object pairs are perhaps similarly mediated by **highly selective “concept neurons” within medial temporal structures**. Indubitably, the substantial improvement of neuroimaging techniques will truly open a fruitful avenue for research on memory in dogs which should help in probing the exact underlying neural mechanisms. For us, humans, the validity of long-term memory can be empirically established as we do recall things or events over many decades (Larzabal et al., 2017). It is difficult to draw a parallel in animal research to substantiate the claim for such long-term memory. Meanwhile, the pioneering and promising demonstrations of episodic-like and now semantic-like memory in nonhuman animals enlightens the possibility of applying human memory models to animals. It also suggests that the cognitive gap between humans and distantly related species is not the chasm it has originally believed to be.

GENERAL CONCLUSION, LIMITS AND PERSPECTIVES

The two objectives of this thesis were to examine if long-term memories for complex stimuli can be established in humans early in life and in a non-primate animal model, and to investigate some of the conditions that are required for these memories to be formed. In order to address these questions, this thesis especially focused on four of the M4 claims that have already been investigated in adult humans, at a behavioral level and in the specific context of the word-referent paradigm.

Using this paradigm, we showed that young children and dogs could recognize cross-modal associations that they have not experienced for months (at least 1 month for 4-year-old children and at least 3 months for puppies) [claim 1] and that were not reactivated during the delay [claim N°2], but we were not able to demonstrate that very few presentations were enough to sustain these memories [claim N°4], neither that the memory trace could survive extreme long delays without being reactivated.

Indeed, for humans early in life, we established that from 2 years of age, a single learning event was sufficient to observe retention after a 30-min delay, but when re-tested after a prolonged delay up to 6 months, older children (preschoolers) failed to recognize the pairings even when five learning trials were provided during learning. Moreover, for domestic dogs, we showed that information recall was possible after a delay of 3 months provided that the knowledge was repeated several times during learning. Indeed, dogs were exposed to the associations almost daily during a 6-month training period since there were inexperienced with the word-referent concept when learning began.

The absence of retention after very long delays with only a few repeats can be explained by the difficulty of the task and the nature of the stimuli we used - stimuli involving two sensory modalities – that depend on one of the most complex structure of the brain. In the general introduction of this thesis, we reviewed the work of colleagues who found that only 1 to 3 exposures to drawings were sufficient for adult participants to recognize these drawings among novel ones a decade later, indicating that few exposures are enough for adults to form long-lasting sensory memories (Larzabal et al., 2018). The stimuli used by these authors were images while here we used word-object associations. Thus, the paradigm implemented in this thesis was much more demanding, since it not only required the creation of a representation of a visual stimulus; it also required the creation of a representation of a sound pattern, and linking the two representations together. As mentioned earlier in this manuscript, linking information from different sensory modalities together in order to create a whole concept is probably only possible

in brain regions at the top of the hierarchy (i.e. in multimodal association areas and medial temporal lobe structures) that are amongst the latest structures to reach maturity in humans. Therefore, it might be that a certain level of brain maturation is required for such complex representations to be maintained higher up in the hierarchy. These arguments are at least in accordance with the phenomena of infantile and juvenile amnesia described in the context of episodic memories and that depend on such medial temporal lobe structures.

However, our work does not exclude the possibility that young children remember things well with a few repeats. It might be that retention after a substantial delay following very few exposures to unimodal sensory inputs that do not require higher-up structure to bind inputs together, is possible in young children and in dogs too. Lower-level brain structures mature earlier, and thus may support longer-term memories in unimodal domains. This is something we would like to test in future work. Our finding that almost all dog subjects were excellent at recognizing the objects after a break of 3 months while many failed to recall the associations stresses how challenging our task was and evidences that long-term memories in one modality should be longer preserved.

Secondly, throughout this thesis, we focused on a parameter that is at the heart of the M4 project and that is reported to play a key role in the establishment and maintenance of a memory at long-term as described by Hebbian and STDP rules: repetition. Our results indeed suggest that beyond the maturational stage of the brain structures and the difficulty of the task, the number of times organisms encounter the stimuli during learning seems to be a critical parameter to form durable memories. Indeed, we were able to highlight that the strength of the memory was dependent on the number of presentations provided during learning [claim 3]. Our results are in line with the idea that early in life, children (and regarding our findings, probably puppies too) rapidly acquire novel information but if not rehearsed, the information is also relatively rapidly lost. This is consistent with the fact that synaptic plasticity is increased early in life, allowing young organisms to rapidly learn a huge amount of information. Perhaps, all that is needed for young organisms to maintain information at long-term is to regularly reactivate the memory traces via repetitions distributed in time. It might be that without repetitions distributed in time and without the possibility for children or animals to internally rehearse the newly acquired knowledge, the information is not accessible anymore (or at least difficult to access later). In this thesis, due to experimental constraints, the repetitions occurred over the course of a single learning session in children and three learning sessions in dogs (Moka). In future work, it would be interesting to explore the effect of scattering the repetitions in time (i.e. one repetition a day during several days scattered in time). Since it is now well-established that sleep is crucial for the consolidation

process, we hypothesize that few repetitions distributed in time may be enough to maintain such complex representations in memory at very long-term during the early life of organisms.

Overall, our findings that emphasize a repetition-driven learning fit with the theories at a neuronal level described in the introduction of this thesis. Indeed, it can be proposed that already early in life, neurons can become selective to their preferred sensory stimuli and probably do not fire at all until the stimulus is present again. However, in contrast to adults for whom the neural pathways are already well-organized, the circuits that are poorly active as well as the synapses that are not sufficiently reinforced during learning are pruned during early childhood. Therefore, even if a neuron preserves its pattern of selectivity for highly integrated information (e.g. a “concept” neuron or even a “grandmother” neuron for a complex representation such as the name of an object), it might be that it is “the access” to that neuron that is lost during childhood. Repetitions, especially if distributed in time, would reinforce the pathway as well as the synaptic weights of the recruited neurons, and thus allow an access to the higher-order neurons that are selective to the concept as a whole. As mentioned just above, in this thesis, since we were using cross-modal stimuli, we were tapping into neurons that integrate concepts as a whole (e.g. the “rivot” concept) and that are localized on the top of the processing hierarchy. However, since no single unit recordings have never been performed in a child or a dog brain it is still unclear whether “rivot” neurons similar to the Jennifer Anniston’s neuron discovered in adult humans can be formed early during development or in another animal species.

Finally, in the context of word learning, we also highlighted that providing repetitions was not the only “ingredient” to promote the creation of a memory: we showed that the attentional resources of an individual during learning as well as the learning strategies also played an important role. For instance, we showed that for a similar number of exposures, when using an ostensive labeling method, 2-year-olds were capable of recognizing the name of newly learned objects, while when using a reasoning learning strategy, retention was not observed in children of the same age. Likewise, in dogs, we found that for a similar number of exposures to the pairings, ostensively naming objects seemed to be less efficient than implementing test trials with feedback of accuracy and reward, to induce a memory trace that will last in time (at least 1 month). Furthermore, our results indicate that young children and dogs do not rely on the same learning strategies to form a memory trace of the name of a novel object. Thus, it might be that the methods implemented for acquiring novel information may vary from one species to another, but to be established and preserved, memories may depend on similar mechanisms (e.g. neuronal selectivity, synaptic plasticity, etc.). Obviously, in this thesis we are far from having enough evidence to claim that the mechanisms underlying the formation of long-term memories may be “universal”, but our results,

at least, do not go against this idea. However, as discussed earlier, our findings suggest that the rapidity with which a memory is formed as well as the durability of that memory in time is enhanced in mature and more complex brain systems. According to our opinion, there might be an evolutionary aspect regarding the complexity, rapidity and durability of the information an organism can acquire.

Limits

We are aware that this thesis contains many weaknesses. First, we would like to mention that it was the first time that experiments involving dogs were carried out in our lab and only a very few number of studies at CerCo included children in their protocols in the past. Therefore, it was very challenging to run experiments with these two models of research. We first had to set up collaborations and partnerships with specialists from both study models. Then, we had to find facilities to carry out the experiments with dogs. Finally, we had to overcome the heavy procedures to get all our ethical agreements. It took us more than a year to get our studies approved by the CPP (Comité de Protection des Personnes), which was vital to start our experiments with young children at CerCo. This is the reason why some of our studies encompassing children are unfortunately not totally completed. But we are aware that the main flaw comes from our experiments with dogs. We did not expect so few dogs to acquire the word-referent concept. Indeed, in the literature, apart from the few impressive dog cases, there was no large-scale study investigating this concept on a large range of dog species and dog ages (the fMRI study from Prichard et al. was only published in 2018 and the criteria for learning were not very stringent). Thus, by the time we started our experiments, there was no evidence that this concept would be so tough for dogs to acquire. Originally, we expected that at least a third or even a half of our dog subjects would reliably know the meaning of about a dozen objects at the end of the 6-month training period. Thus, we had initially planned to conduct the series of experiments on long-term memory following various learning strategies with a sample of about fifteen to twenty dogs, which would have been more relevant for statistical purposes. Unfortunately, as explained in this manuscript, only a very few dogs and particularly one was consistently accurate during the fetching exercises. Even with this unique dog subject, we had to downgrade our initial objectives that consisted of implementing the same protocols used with children, although the learning support and environment would anyway have been different (e.g. touchscreen vs real objects; neural verbal encouragements between trials vs food reward, etc.). For example, we had hoped to teach Moka three novel word-object pair a session and test him on that knowledge 30min after learning. As shown in this manuscript, a single learning session was far from enough for him

to reliably incorporate the name of three novel items, probably because his degree of expertise with the task was not yet sufficient.

Perspectives

There are many perspectives to the work done in this thesis. First of all, we obviously need to complete our experiments 4, 5 and 6 by including more child and adult participants. Secondly, it would naturally be very interesting to examine the formation of (very) long-term memories for complex sensory inputs in other animal species, perhaps genetically more distant from humans or less keen to communicate with humans. Next, as already mentioned several times throughout this thesis, neuroimaging techniques would provide objective evidence about the neural mechanisms and brain regions that support the formation of cross-modal sensory memories early in life as well as in other animal species.

Furthermore, as mentioned above, it would also be interesting to investigate whether the results we obtained are specific to the word-referent paradigm or generalizable to other types of stimuli. For instance, we could replicate our experiments with novel faces associated with invented names, since it is well-known that early in life, babies are predominantly attracted by faces. One could argue that using stimuli with such a positive emotional valence would increase the strength of the memory. As also said previously, in future work we would like to replicate the RSVP experiment conducted on adults in children, in order to better characterize if repetition-based learning during childhood is enhanced when only one modality is involved. We would also like to record ERPs in order to underlie the brain correlates during the detection of repetitions in children.

REFERENCES

- Adachi, I., Kuwahata, H., & Fujita, K. (2007). Dogs recall their owner's face upon hearing the owner's voice. *Animal Cognition*, 10, 17–21. <http://doi.org/10.1007/s10071-006-0025-8>
- Adams, B., Chan, A., Callahan, H., Siwak, C., Tapp, D., Ikeda-douglas, C., ... Milgram, N. W. (2000). Use of a delayed non-matching to position task to model age-dependent cognitive decline in the dog. *Behavioural Brain Research*, 108, 47–56.
- Adlam, A. R., Vargha-khadem, F., Mishkin, M., & De Haan, M. (2005). Deferred Imitation of Action Sequences in Developmental Amnesia. *Journal of Cognitive Neuroscience*, 17(2), 240–248.
- Aggleton, J. P., Dumont, J. R., & Warburton, E. C. (2011). Unraveling the contributions of the diencephalon to recognition memory : A review. *Learning & Memory*, 384–400.
- Agus, T. R., Thorpe, S. J., & Pressnitzer, D. (2010). Article Rapid Formation of Robust Auditory Memories : Insights from Noise. *Neuron*, 66(4), 610–618. <http://doi.org/10.1016/j.neuron.2010.04.014>
- Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., & Mills, D. (2016). Dogs recognize dog and human emotions. *Animal Behaviour*, (12). <http://doi.org/10.1098/rsbl.2015.0883>
- Andics, A., Gábor, A., Gácsi, M., Faragó, T., Szabó, D., & Miklósi, Á. (2016). Neural mechanisms for lexical processing in dogs. *Science*, 353(6303), 1030–1032. <http://doi.org/10.1126/science.aaf3777>
- Andics, A., Gacsi, M., Farago, T., Kis, A., & Miklosi, A. (2014). Report Voice-Sensitive Regions in the Dog and Human Brain Are Revealed by Comparative fMRI. *Current Biology*, 24, 574–578. <http://doi.org/10.1016/j.cub.2014.01.058>
- Andrillon, T., Kouider, S., Agus, T., & Pressnitzer, D. (2015). Perceptual Learning of Acoustic Noise Generates Report Perceptual Learning of Acoustic Noise Generates Memory-Evoked Potentials. *Current Biology*, 25(21), 2823–2829. <http://doi.org/10.1016/j.cub.2015.09.027>
- Aravind, A., Villiers, J. De, Pace, A., Valentine, H., Golinko, R., Hirsch-pasek, K., ... Wilson, M. S. (2018). Fast mapping word meanings across trials : Young children forget all but their first guess. *Cognition*, 177, 177–188. <http://doi.org/10.1016/j.cognition.2018.04.008>
- Atir-sharon, T., Gilboa, A., Hazan, H., Koilis, E., & Manevitz, L. M. (2015). Decoding the Formation of New Semantics : MVPA Investigation of Rapid Neocortical Plasticity during Associative Encoding through Fast Mapping. *Neural Plasticity*, 1–17.
- Aust, U., & Huber, L. (2001). The role of item- and category-specific information in the discrimination of people versus nonpeople images by pigeons. *Animal Learning and Behavior*, 29(2), 107–119.
- Aust, U., Range, F., Steurer, M., & Huber, L. (2008a). Inferential reasoning by exclusion in pigeons , dogs , and humans, 587–597. <http://doi.org/10.1007/s10071-008-0149-0>
- Aust, U., Range, F., Steurer, M., & Huber, L. (2008b). Inferential reasoning by exclusion in pigeons , dogs , and humans. *Animal Cognition*, 11, 587–597. <http://doi.org/10.1007/s10071-008-0149-0>
- Axelsson, E. ., Churchley, K., & Horst, J. S. (2012). The Right Thing at the Right Time : Why Ostensive Naming Facilitates

- Word The right thing at the right time : why ostensive naming facilitates word learning. *Frontiers in Psychology*, 3(88), 1–8. <http://doi.org/10.3389/fpsyg.2012.00088>
- Axelsson, E. L., & Horst, J. S. (2013). Testing a word is not a test of word learning. *Acta Psychologica*, 144, 264–268. <http://doi.org/10.1016/j.actpsy.2013.07.002>
- Axelsson, E. L., Swinton, J., Winiger, A. I., & Horst, J. S. (2018). Napping and toddlers ' memory for fast-mapped words. *First Language*, 38(6), 582–595. <http://doi.org/10.1177/0142723718785490>
- Bachevalier, J. (2015). Commentary: The development of hippocampal-dependent memory functions: Theoretical comments on Jabe`s and Nelson review (2015). *International Journal of Behavioral Development*, 39(4), 310–314.
- Bachevalier, J., Brickson, M., & Hagger, C. (1993). Limbic-dependent recognition memory in monkeys develops early in infancy. *NeuroReport*, 4, 77–80.
- Bachevalier, J., & Mishkin, M. (1984). An Early and a Late Developing System for Learning and Retention in Infant Monkeys. *Behavioral Neuroscience*, 98(5), 770–778.
- Backhaus, J., Hoeckesfeld, R., Born, J., Hohagen, F., & Junghanns, K. (2008). Immediate as well as delayed post learning sleep but not wakefulness enhances declarative memory consolidation in children. *NEUROBIOLOGY OF LEARNING AND MEMORY*, 89, 76–80. <http://doi.org/10.1016/j.nlm.2007.08.010>
- Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: a first step. *Child Development*, 60(2), 381–398.
- Ball, G., Aljabar, P., Zebari, S., Tusor, N., Arichi, T., Merchant, N., & Robinson, E. C. (2014). Rich-club organization of the newborn human brain. *PNAS*, 111(20), 7456–7461. <http://doi.org/10.1073/pnas.1324118111>
- Baltes, P. B. (1987). Theoretical Propositions of Life-Span Developmental Psychology : On the Dynamics Between Growth and Decline. *Developmental Psychology*, 2(5), 611–626.
- Barr, R., Dowden, A., & Hayne, H. (1996). Developmental changes in deferred imitation by 6- to 24-month-old infants. *Infant Behavior and Development*. [http://doi.org/10.1016/S0163-6383\(96\)90015-6](http://doi.org/10.1016/S0163-6383(96)90015-6)
- Bates, E., Thal, D., Finlay, B., & Clancy, B. (1992). Early Language Development And Its Neural Correlates. *Handbook of Neuropsychology*, 6, 2nd edi.
- Bauer, P. J. (1996). What do infants recall of their lives? Memory for specific events by one- to two-year-olds. *American Psychologist*, 51, 29–41.
- Bauer, P. J. (2004). Getting explicit memory off the ground: Steps toward construction of a neuro-developmental account of changes in the first two years of life. *Developmental Review*. <http://doi.org/10.1016/j.dr.2004.08.003>
- Belin, P., Fecteau, S., Charest, I., Nicastro, N., Hauser, M. D., & Armony, J. L. (2014). Human cerebral response to animal affective vocalizations. *Proceedings of the Royal Society B: Biological Sciences*, 275(1634), 473–481. <http://doi.org/10.1098/rspb.2007.1460>
- Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, 6(2), 183–200. <http://doi.org/10.1017/S0305000900002245>
- Bensky, M. K., Gosling, S. D., & Sinn, D. L. (2013). *The World from a Dog ' s Point of View : A Review and Synthesis of*

- Dog Cognition Research. Advances in the Study of Behavior* (Vol. 45). Elsevier Inc. <http://doi.org/10.1016/B978-0-12-407186-5.00005-7>
- Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253–3258. <http://doi.org/10.1073/pnas.1113380109>
- Berns, G. S., Brooks, A. M., & Spivak, M. (2014). Scent of the familiar: An fMRI study of canine brain responses to familiar and unfamiliar human and dog odors. *Behavioural Processes*. <http://doi.org/10.1016/j.beproc.2014.02.011>
- Berns, G. S., Brooks, A., & Spivak, M. (2013). Replicability and Heterogeneity of Awake Unrestrained Canine fMRI Responses, 8(12). <http://doi.org/10.1371/journal.pone.0081698>
- Bi, G., & Poo, M. (2001). SYNAPTIC MODIFICATION BY CORRELATED ACTIVITY : Hebb ' s Postulate Revisited. *Annual Review of Neuroscience*, 24, 139–66.
- Bichler, O., Suri, M., Querlioz, D., Vuillaume, D., Desalvo, B., & Gamrat, C. (2012). Visual Pattern Extraction Using Energy-Efficient “ 2-PCM Synapse ” Neuromorphic Architecture. *IEEE TRANSACTIONS ON ELECTRON DEVICES*, 59(8). <http://doi.org/10.1109/TED.2012.2197951>
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. (2009). Where Is the Semantic System ? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cerebral Cortex*, 19(December), 2767–2796. <http://doi.org/10.1093/cercor/bhp055>
- Bion, R. A. H., Borovsky, A., & Fernald, A. (2013a). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*. <http://doi.org/10.1016/j.cognition.2012.08.008>
- Bion, R. A. H., Borovsky, A., & Fernald, A. (2013b). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, 126(1), 39–53. <http://doi.org/10.1016/j.cognition.2012.08.008>
- Bion, R. A. H., Borovsky, A., & Fernald, A. (2013c). Fast mapping , slow learning : Disambiguation of novel word – object mappings in relation to vocabulary learning at 18 , 24 , and 30 months. *Cognition*, 126(1), 39–53. <http://doi.org/10.1016/j.cognition.2012.08.008>
- Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT press.
- Bloom, P. (2004). Can a Dog Learn a Word ? *Science*, 304(5677), 1605–1607. <http://doi.org/DOI: 10.1126/science.1099899>
- Booth, A. E. (2009a). Causal supports for early word learning. *Child Development*, 80(4), 1243–1250. <http://doi.org/10.1111/j.1467-8624.2009.01328.x>
- Booth, A. E. (2009b). Causal Supports for Early Word Learning. *Child Development*, 80(4), 1243–1250.
- Booth, A. E., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-Pragmatics and Attention: Contributions to Gesturally Guided Word Learning in Toddlers. <http://doi.org/10.1080/15475440802143091>
- Booth, A. E., McGregor, K. K., & Rohlfing, K. J. (2008). Socio-Pragmatics and Attention: Contributions to Gesturally Guided Word Learning in Toddlers. *Language Learning and Development*, 4(3), 179–202. <http://doi.org/10.1080/15475440802143091>

- Born, J. A. N., Rasch, B., & Gais, S. (2006). Sleep to Remember. *Neuroscientist*, 12(5). <http://doi.org/10.1177/1073858406292647>
- Bräuer, J., Kaminski, J., Riedel, J., Call, J., & Tomasello, M. (2006). Making Inferences About the Location of Hidden Food : Social Dog , Causal Ape. *Journal of Comparative Psychology*, 120(1), 38–47. <http://doi.org/10.1037/0735-7036.120.1.38>
- Brown, M. W., & Aggleton, J. P. (2001). RECOGNITION MEMORY : WHAT ARE THE ROLES OF THE PERIRHINAL CORTEX AND HIPPOCAMPUS ? *Nature Neuroscience*, 2, 51–61.
- Bullock Drummey, A., & Newcombe, N. (1995). Remembering versus Knowing the Past: Children's Explicit and Implicit Memories for Pictures. *Journal of Experimental Child Psychology*, 59, 549–565.
- Burdyn, L. E., Noble, L. M., Shreves, L. E., & Thomas, R. K. (1984). Long-term memory for concepts by squirrel monkeys. *Physiological Psychology*, 12(2), 97–102.
- Byosiére, S., Chouinard, P. A., Howell, T. J., & Bennett, P. C. (2017). What do dogs (*Canis familiaris*) see ? A review of vision in dogs and implications for cognition research. *Psychonomic Bulletin & Review*, 25(3).
- Cacciaglia, R., Escera, C., Slabu, L., Grimm, S., & Sanjuán, A. (2015). Neuropsychologia Involvement of the human midbrain and thalamus in auditory de- viance detection. *Neuropsychologia*, 68, 51–58. <http://doi.org/10.1016/j.neuropsychologia.2015.01.001>
- Candland DK. (1993). *Feral children and clever animals*. New York: Oxford University Press.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, 15, 17–29.
- Carver, L. J., & Bauer, P. J. (1999). When the Event is More Than the Sum of its Parts : 9-month-olds ' Long-term Ordered Recall. *Memory*, 7(2), 147–174. <http://doi.org/10.1080/741944070>
- Carver, L. J., & Bauer, P. J. (2001). The Dawning of a Past : The Emergence of Long-Term Explicit Memory in Infancy. *Journal of Experimental Psychology*, 130(4), 726–745. <http://doi.org/10.1037/0096-3445.130.4.726>
- Castellucci, V., Pinsker, H., Kupfermann, I., & Kandel, E. R. (1970). Neuronal mechanisms of habituation and dishabituation of the gill-withdrawal reflex in *Aplysia*. *Science*, 167(3926), 1745–1748.
- Chan, A. D. F., Nippak, P. M. D., Murphey, H., Ikeda-douglas, C. J., Muggenburg, B., Head, E., ... Al, C. E. T. (2002). Visuospatial Impairments in Aged Canines (*Canis familiaris*): The Role of Cognitive – Behavioral Flexibility. *Behavioral Neuroscience*, 116(3), 443–454. <http://doi.org/10.1037/0735-7044.116.3.443>
- Chugani, H. T., Phelps, M. E., & Mazziotta, J. C. (1987). Positron Emission Tomography Study of Human Brain Functional Development. *Annals of Neurology*, 22(4), 487–497.
- Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language*, 15(2), 317–335.
- Clayton, N. S., & Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395, 272–274.
- Cohen, N. J., & Squire, L. R. (1980). Preserved Learning and Retention of Pattern-Analyzing Skill in Amnesia : Dissociation of Knowing How and Knowing That. *Science*, 210, 207–210.

- Collier, K., Bickel, B., van Schaik, C. P. Van, Manser, M. B., & Townsend, S. W. (2014). Language evolution : syntax before phonology ? *Proceedings of the Royal Society B: Biological Sciences*, 281(1788).
- Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology Gen*, 120(4), 395–409. <http://doi.org/10.1037/0096-3445.120.4.395>
- Cook, P. F., Prichard, A., Spivak, M., & Berns, G. S. (2016). Awake canine fMRI predicts dogs ' preference for praise vs food. *Social Cognitive and Affective Neuroscience*, (August), 1853–1862. <http://doi.org/10.1093/scan/nsw102>
- Coutanche, M. N., & Thompson-Schill, S. L. (2014). Fast Mapping Rapidly Integrates Information Into Existing Memory Networks. *Journal of Experimental Child Psychology: General*, 143(6), 2296–2303.
- Craik, F. I. M., & Salthouse, T. A. (2008). *The handbook of Aging and Cognition: Third Edition* (Third Edit). New York: Oxford University Press.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 934–945.
- Cuaya, L. V, Hernández-pérez, R., & Concha, L. (2016). Our Faces in the Dog s Brain : Functional Imaging Reveals Temporal Cortex Activation during Perception of Human Faces. *PLoS ONE*, 11(3), 1–13. <http://doi.org/10.1371/journal.pone.0149431>
- Dapretto, M., & Bjork, E. L. (2000). The Development of Word Retrieval Abilities in the Second Year and Its Relation to Early Vocabulary. *Society for Research in Child Developemnt*, 71(3), 635–648.
- De Baene, W., & Vogels, R. (2010). Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. *Cerebral Cortex*, 20(9), 2145–65. <http://doi.org/10.1093/cercor/bhp277>
- De Lahunta, A., Glass, E., & Kent, M. (2014). *Veterinary neuroanatomy and clinical neurology*.
- Dehaene-Lambertz, G., & Spelke, E. S. (2015). Perspective The Infancy of the Human Brain. *Neuron*, 88(1), 93–109. <http://doi.org/10.1016/j.neuron.2015.09.026>
- Demant, H., Ladewig, J., Balsby, T. J. S., & Dabelsteen, T. (2011a). The effect of frequency and duration of training sessions on acquisition and long-term memory in dogs. *Applied Animal Behaviour Science*, 133, 228–234. <http://doi.org/10.1016/j.applanim.2011.05.010>
- Demant, H., Ladewig, J., Balsby, T. J. S., & Dabelsteen, T. (2011b). The effect of frequency and duration of training sessions on acquisition and long-term memory in dogs. *Applied Animal Behaviour Science*, 133(3–4), 228–234. <http://doi.org/10.1016/j.applanim.2011.05.010>
- Diekelmann, S., Wilhelm, I., & Born, J. (2009). The whats and whens of sleep-dependent memory consolidation. *Sleep Medicine Reviews*, 13, 309–321. <http://doi.org/10.1016/j.smrv.2008.08.002>
- Diesendruck, G., & Markson, L. (2001). Children ' s Avoidance of Lexical Overlap : A Pragmatic Account. *Developmental Psychology*, 37(5), 630–641. <http://doi.org/10.1037/0012-1649.37.5.630>
- Dilks, D. D., Cook, P., Weiller, S. K., Berns, H. P., Spivak, M., & Berns, G. S. (2015). Awake fMRI reveals a specialized region in dog temporal cortex for face processing. *PeerJ*, 1–13. <http://doi.org/10.7717/peerj.1115>

- Dohen, M., Vilain, A., Loevenbruck, H., Rochet-Cappellan, A., & Gillet-Perret, E. (2016). The role of Manual Representational Gestures in Novel Word Acquisition in Children with Specific Language Impairments. In *7th Conference of the International Society for Gesture Studies (ISGS7)* (p. 166). Paris.
- Doupe, A. J., & Kuhl, P. K. (1999). BIRDSONG AND HUMAN SPEECH : Common Themes and Mechanisms. *Annual Review of Neuroscience*, 22(1), 567–631.
- Eckenhoff, F., & Rakic, P. (1988). Nature and Fate of Proliferative Cells in the Hippocampal Gyrus During the Life Span of the Rhesus Monkey Dentate. *The Journal of Neuroscience*, 8(8), 2729–2747.
- Eichenbaum, H., & Cohen, N. J. (2001). *From conditioning to conscious recollection: Memory systems of the brain*. New York: Oxford University Press.
- Eichenbaum, H., Schoenbaum, G., Young, B., & Bunsey, M. (1996). Functional organization of the hippocampal memory system. *PNAS*, 93, 13500–13507.
- Fabre-Thorpe, M. (2003). Visual categorization : accessing abstraction in non-human primates. *The Royal Society*, (June), 1215–1223. <http://doi.org/10.1098/rstb.2003.1310>
- Fagan, J. F. (1973). Infants ' Delayed Recognition Memory and Forgetting. *Journal of Experimental Child Psychology*, 16, 424–450.
- Fantz, R. L. (1956). A method for studying early visual development. *Perceptual and Motor Skills*, 6, 13–15.
- Farago, T., Andics, A., Devecseri, V., Kis, A., Gacsi, M., & Miklosi, A. (2014). Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. *Biology Letters*, 10(1).
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., ... Stiles, J. (1994). Variability in Early Communicative Development. *MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT*, 59(5).
- Fernald, A., Zangl, R., Portillo, A. ., & Marchman, V. A. (2008). Looking while listening : Using eye movements to monitor spoken language comprehension by infants and young children. In I. Sekerina, E. M. Fernandez, & H. Clahsen (Eds.), *Developmental Psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam.
- Fischer, J., Call, J., & Kaminski, J. (2004). A pluralistic account of word learning. *Trends in Cognitive Sciences*, 8(11), 6613. <http://doi.org/10.1016/j.tics.2004.09.007>
- Fiset, S., Beaulieu, C., & Landry, F. (2003). Duration of dogs ' (Canis familiaris) working memory in search for disappearing objects, 1–10. <http://doi.org/10.1007/s10071-002-0157-4>
- Frank, H. (2011). Wolves , Dogs , Rearing and Reinforcement : Complex Interactions Underlying Species Differences in Training and Problem-Solving Performance. *Behav Genet*, 41, 830–839. <http://doi.org/10.1007/s10519-011-9454-5>
- Friedrich, M., & Friederici, A. D. (2011a). Word Learning in 6-Month-Olds : Fast Encoding – Weak Retention. *Journal of Cognitive Neuroscience*, 23, 3228–3240.
- Friedrich, M., & Friederici, A. D. (2011b). Word Learning in 6-Month-Olds : Fast Encoding – Weak Retention, 3228–3240.
- Friedrich, M., & Friederici, A. D. (2015). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Developmental Science*, 20(2), 1–13. <http://doi.org/10.1111/desc.12357>

- Friedrich, M., Wilhelm, I., Born, J., & Friederici, A. D. (2015). Generalization of word meanings during infant sleep. *Nature Communications*. <http://doi.org/10.1038/ncomms7004>
- Fugazza, C., & Miklósi, Á. (2014). Deferred imitation and declarative memory in domestic dogs. *Animal Cognition*, 17, 237–247. <http://doi.org/10.1007/s10071-013-0656-5>
- Fujita, K., Morisaki, A., Takaoka, A., Maeda, T., & Hori, Y. (2012). Incidental memory in dogs (*Canis familiaris*): adaptive behavioral solution at an unexpected memory test. *Animal Cognition*. <http://doi.org/10.1007/s10071-012-0529-3>
- Fukuzawa, M., Mills, D. S., & Cooper, J. J. (2005). More than just a word : non-semantic command variables affect obedience in the domestic dog (*Canis familiaris*). *Applied Animal Behaviour Science*, 91, 129–141. <http://doi.org/10.1016/j.applanim.2004.08.025>
- Galibert, F., Quignon, P., Hitte, C., & Andre, C. (2011). Comptes Rendus Biologies Toward understanding dog evolutionary and domestication history. *Compte Rendus Biologies*, 334, 190–196. <http://doi.org/10.1016/j.crvi.2010.12.011>
- Gardner, R. A., & Gardner, B. T. (1969). Teaching Sign Language to a Chimpanzee. *Science*, 165(3894), 664–672.
- Gardner, R. A., Gardner, B. T., & Van Cantfort, T. E. (1989). *Teaching Sign Language to a Chimpanzee*.
- Gathercole, S. E., & Hitch, G. . (1993). Developmental changes in short-term memory: a revised working memory perspective. In S. . Collins, S. E. Gathercole, Conway, & P. . Morris (Eds.), *Theories of memory* (pp. 189–210). Hove, England: Erlbaum.
- Gathercole, S. E., Pickering, S. ., Ambridge, B., & Wearing, H. (2004). The structure of Working Memory from 4 to 15 years of age. *Developmental Psychology*, 20(2), 177–190.
- Gibson, J. M., Scavelli, S. A., Udell, C. J., & Udell, M. A. R. (2014). Domestic Dogs (*Canis lupus familiaris*) are Sensitive to the “ Human ” Qualities of Vocal Commands. *Animal Behavior and Cognition*, 1(3), 281–295. <http://doi.org/10.12966/abc.08.05.2014>
- Giedd, J. N., & Blumenthal, J. (1999). Brain Development during Childhood and Adolescence : A Longitudinal MRI Brain development during childhood and adolescence : a longitudinal MRI study. *Nature Neuroscience*, 2(10), 861–863. <http://doi.org/10.1038/13158>
- Gold, A. E., & Kesner, R. P. (2005). The Role of the CA3 Subregion of the Dorsal Hippocampus in Spatial Pattern Completion in the Rat. *Hippocampus*, 814, 808–814. <http://doi.org/10.1002/hipo.20103>
- Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition : rate, content, and the vocabulary spurt. *Journal of Child Language*, 17(01), 171–183. <http://doi.org/10.1017/S0305000900013167>
- Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young Children and Adults Use Lexical Principles to Learn New Nouns. *Developmental Psychology*, 28(1), 99–108. <http://doi.org/10.1037/0012-1649.28.1.99>
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14(1), 23–45. <http://doi.org/10.1017/S030500090001271X>
- Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The Role of Semantic Context and Memory in the Acquisition of Novel Nouns. *Child Development*, 69(5), 1330. <http://doi.org/10.2307/1132269>

- Grassmann, S., Kaminski, J., & Tomasello, M. (2012). How two word-trained dogs integrate pointing and naming. *Animal Cognition*, 15, 657–665. <http://doi.org/10.1007/s10071-012-0494-x>
- Greco, C., Rovee-collier, C., Hayne, H., Griesler, P., & Earley, L. (1986). Ontogeny of Early Event Memory : Forgetting and Retrieval by 2- and 3-Month-Olds. *Infant Behavior and Development*, 9, 441–460.
- Greve, A., Cooper, E., & Henson, R. N. (2014). Neuropsychologia No evidence that ‘ fast-mapping ’ bene fi ts novel learning in healthy Older adults. *Neuropsychologia*, 60, 52–59. <http://doi.org/10.1016/j.neuropsychologia.2014.05.011>
- Griebel, U., & Oller, D. K. (2012). Vocabulary Learning in a Yorkshire Terrier : Slow Mapping of Spoken Words. *PLoS ONE*, 7(2). <http://doi.org/10.1371/journal.pone.0030182>
- Gurteen, P. M., Horne, P. J., & Erjavec, M. (2011a). Journal of Experimental Child Rapid word learning in 13- and 17-month-olds in a naturalistic two-word procedure : Looking versus reaching measures. *Journal of Experimental Child Psychology*, 109(2), 201–217. <http://doi.org/10.1016/j.jecp.2010.12.001>
- Gurteen, P. M., Horne, P. J., & Erjavec, M. (2011b). Rapid word learning in 13- and 17-month-olds in a naturalistic two-word procedure: Looking versus reaching measures. *Journal of Experimental Child Psychology*. <http://doi.org/10.1016/j.jecp.2010.12.001>
- Gurteen, P. M., Horne, P. J., & Erjavec, M. (2011c). Rapid word learning in 13- and 17-month-olds in a naturalistic two-word procedure: Looking versus reaching measures. *Journal of Experimental Child Psychology*, 109(2), 201–217. <http://doi.org/10.1016/j.jecp.2010.12.001>
- Guyonneau, R., VanRullen, R., & Thorpe, S. J. (2005). Neurons Tune to the Earliest Spikes Through STDP. *Neural Computation*, 17(4), 859–879.
- Haber, S. N., & Knutson, B. (2009). The Reward Circuit : Linking Primate Anatomy and Human Imaging. *Neuropsychopharmacology*, 35(1), 4–26. <http://doi.org/10.1038/npp.2009.129>
- Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87(1), 23–34. [http://doi.org/10.1016/S0010-0277\(02\)00186-5](http://doi.org/10.1016/S0010-0277(02)00186-5)
- Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*. <http://doi.org/10.1016/j.cogpsych.2006.04.003>
- Hanggi, E. B., & Ingersoll, J. F. (2009). Long-term memory for categories and concepts in horses (Equus caballus). *Animal Cognition*, 451–462. <http://doi.org/10.1007/s10071-008-0205-9>
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The Domestication of Social Cognition in Dogs. *Science*, 298, 1634–1637.
- Hare, B., & Hare, B. (2007). From Nonhuman to Human Mind What Changed and Why ? *Current Directions in Psychological Science*. <http://doi.org/10.1111/j.1467-8721.2007.00476.x>
- Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs ? *Trends in Cognitive Sciences*, 9(9), 439–444. <http://doi.org/10.1016/j.tics.2005.07.003>
- Hartshorn, K., Rovee-collier, C., Gerhardstein, P., Bhatt, R. S., Wondoloski, T. L., Klein, P., & Campos-de-carvalho, M. (1998). The Ontogeny of Long-Term Memory over the First Year- and-a-Half of Life. *Developmental Psychology*, 32,

- Hayne, H., Boniface, J., & Barr, R. (2000). The Development of Declarative Memory in Human Infants : Age-Related Changes in Deferred Imitation. *Behavioral Neuroscience*, 114(1), 77–83. <http://doi.org/10.1037//0735-7044.114.1.77>
- Hayne, H., Greco, C., Earley, L., Griesler, P., & Rovee-collier, C. (1986). Ontogeny of Early Event Memory : Encoding and Retrieval by 2- and SMonth-Olds. *Infant Behavior and Development*, 9, 461–472.
- Hayne, H., MacDonald, S., & Barr, R. (1997). Developmental changes in the specificity of memory over the second year of lie. *Infant Behavior and Development*, 20(2), 233–245. [http://doi.org/10.1016/S0163-6383\(97\)90025-4](http://doi.org/10.1016/S0163-6383(97)90025-4)
- Hayne, H., Rovee-collier, C., & Perris, E. E. (1987). Categorization and Memory Retrieval by three-month-olds. *Child Development*, 58, 750–767.
- Head, E., Mehta, R., Hartley, J., & Kameka, M. (1995). Spatial Learning and Memory as a Function of Age in the Dog. *Behavioral Neuroscience*, 109(5), 851–858.
- Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley & Sons.
- Heffner, H. (1975). Perception of biologically meaningful sounds by dogs. *The Journal of the Acoustical Society of America*, 58(S1). <http://doi.org/10.1121/1.2001931>
- Herbert, J., & Hayne, H. (2000). The ontogeny of long-term retention during the second year of life. *Developmental Science*, 3(1), 50–56.
- Herman, L. M., & Wolz, J. P. (1984). Comprehension of Sentences of Bottlenosed Dolphins. *Cognition*, 16, 129–219. [http://doi.org/10.1016/0010-0277\(84\)90003-9](http://doi.org/10.1016/0010-0277(84)90003-9)
- Hill, W. L., Borovsky, D., & Rovee-collier, C. (1988). Continuities in Infant Memory Development. *Developmental Psychobiology*, 21(1), 43–62.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., ... Rocroi, C. (2000). Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*. <http://doi.org/10.1111/1540-5834.00090>
- Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128–157. <http://doi.org/10.1080/15250000701795598>
- Horst, J. S., Scott, E. J., & Pollard, J. A. (2010). The role of competition in word learning via referent selection. *Developmental Science*, 13(5), 706–713. <http://doi.org/10.1111/j.1467-7687.2009.00926.x>
- Houston-Price, C., Caloghris, Z., & Raviglione, E. (2010). Language experience shapes the development of the Mutual Exclusivity bias. *Infancy*, 15(125–150).
- Houston-Price, C., Plunkett, K., & Harris, P. (2005). Word-learning wizardry' at 1;6. *Journal of Child Language*, 32(1), 175–189. <http://doi.org/10.1017/S0305000904006610>
- Howard, D. M., & Angus, J. (1996). *Acoustics and Psychoacoustics*. (F. Press, Ed.). Oxford.
- Howell, T. J., Conduit, R., Toukhsati, S., & Bennett, P. (2012). Auditory stimulus discrimination recorded in dogs , as indicated by mismatch negativity (MMN). *Behavioural Processes*, 89(1), 8–13.

<http://doi.org/10.1016/j.beproc.2011.09.009>

- Hubel, D. H., & Wiesel, T. N. (1959). RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT ' S STRIATE CORTEX. *J. Physiol.*, 148, 574–591.
- Huttenlocher, P. R. (1979). SYNAPTIC DENSITY IN HUMAN FRONTAL CORTEX- DEVELOP- MENTAL CHANGES AND EFFECTS OF AGING. *Brain Research*, 163, 195–205.
- Jabès, A., & Nelson, C. A. (2015). 20 years after “the ontogeny of human memory: A cognitive neuroscience perspective,” where are we? *International Journal of Behavioral Development*, 39(4), 293–303. <http://doi.org/10.1177/0165025415575766>
- Johnson, C. K., & Davis, R. T. (1973). Seven-year retention of oddity learning set in monkeys. *Perceptual and Motor Skills*, 37(3), 920–922.
- Johnson, M. H., & De Haan, M. (2015). *Developmental cognitive neuroscience*. (4th Edition, Ed.) (Wiley Blac). Oxford, UK.
- Jones, E. J. H., Pascalis, O., Eacott, M. J., & Herbert, S. (2011). Visual recognition memory across contexts. *Developmental Science*, 14(1), 136–147. <http://doi.org/10.1111/j.1467-7687.2010.00964.x>
- Kahana, M. J. (2006). The Cognitive Correlates of Human Brain Oscillations. *The Journal of Neuroscience*, 26(6), 1669–1672. <http://doi.org/10.1523/JNEUROSCI.3737-05c.2006>
- Kaminski, J., Call, J., & Fischer, J. (2004). Word Learning in a Domestic Dog : Evidence for “ Fast Mapping .” *Science*, 304(June), 1682–1684.
- Kaminski, J., Fischer, J., & Call, J. (2008). Prospective object search in dogs : mixed evidence for knowledge of What and Where. *Animal Cognition*, 11, 367–371. <http://doi.org/10.1007/s10071-007-0124-1>
- Kaminski, J., Tempelmann, S., Call, J., & Tomasello, M. (2009). Domestic dogs comprehend human communication with iconic signs. *Developmental Science*, 12(6), 831–837. <http://doi.org/10.1111/j.1467-7687.2009.00815.x>
- Kaulfuß, P., & Mills, D. S. (2008). Neophilia in domestic dogs (*Canis familiaris*) and its implication for studies of dog cognition. *Animal Cognition*, 11(3), 553–556. <http://doi.org/10.1007/s10071-007-0128-x>
- Kendrick, K., & Baldwin, B. (1987). Cells in temporal cortex of conscious sheep can respond preferentially to the sight of faces. *Science*, 236(4800), 448–450.
- Kern, S., Langue, J., Zesiger, P., & Bovet, F. (2010). Adaptations françaises des versions courtes des inventaires du développement communicatif de MacArthur-Bates. *ANAE - Approche Neuropsychologique Des Apprentissages Chez l'Enfant*, 22(107–108), 217–228.
- Knornschild, M., Nagy, M., Metz, M., Mayer, F., & Von Helversen, O. (2009). Complex vocal imitation during ontogeny in a bat. *Biology Letters*, 6(2), 156–159.
- Kreiman, G., Koch, C., & Fried, I. (2000). Category-specific visual responses of single neurons in the human. *Nature*, 3(9), 946–953. <http://doi.org/10.1038/78868>
- Kucker, S. C., & Samuelson, L. K. (2012). The First Slow Step: Differential Effects of Object and Word-Form Familiarization on Retention of Fast-Mapped Words. *Infancy*, 17(3), 295–323. <http://doi.org/10.1111/j.1532-7078.2011.00081.x>

- Kuhl, P. K. (2004a). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*. <http://doi.org/10.1038/nrn1533>
- Kuhl, P. K. (2004b). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*. <http://doi.org/10.1038/nrn1533>
- Ladefoged, P. (2004). *Vowels and Consonants: An introduction to the sounds of language*. (Blackwell, Ed.) (2nd ed). Oxford, UK.
- Larzabal, C., Bacon-macé, N., Muratot, S., & Thorpe, S. J. (2017). Waking Up Buried Memories of Old TV Programs, *11*(April). <http://doi.org/10.3389/fnbeh.2017.00060>
- Larzabal, C., Tramon, E., Muratot, S., Barbeau, E. J., & Thorpe, S. J. (2018). Extremely long-term memory and familiarity after 12 years. *Cognition*, *170*, 254–262. <http://doi.org/10.1016/j.cognition.2017.10.009>
- Lavenex, P., & Amaral, D. G. (2000). Hippocampal-Neocortical Interaction: A Hierarchy of Associativity. *Hippocampus*, *10*, 420–430.
- Lavenex, P., Banta Lavenex, P., & Amaral, G. (2006). Postnatal Development of the Primate Hippocampal Formation. *Developmental Neuroscience*, 179–192. <http://doi.org/10.1159/000096222>
- Lavenex, P., & Lavenex, P. B. (2006). Spatial relational memory in 9-month-old macaque monkeys. *Learning & Memory*, *13*, 84–96. <http://doi.org/10.1101/lm.97606.venile>
- Lenroot, R. K., & Giedd, J. N. (2006a). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews*. <http://doi.org/10.1016/j.neubiorev.2006.06.001>
- Lenroot, R. K., & Giedd, J. N. (2006b). Brain development in children and adolescents: Insights from anatomical magnetic resonance imaging. *Neuroscience and Biobehavioral Reviews*. <http://doi.org/10.1016/j.neubiorev.2006.06.001>
- Leonardi, R. J., Vick, S.-J., & Dufour, V. (2012). Waiting for more: the performance of domestic dogs (*Canis familiaris*) on exchange tasks. *Animal Cognition*, *15*(1), 107–120. <http://doi.org/10.1007/s10071-011-0437-y>
- Linden, D. E. J., Prvulovic, D., Formisano, E., Völlinger, M., Zanella, F. E., Goebel, R., ... Goethe-universität, J. W. (1999). The Functional Neuroanatomy of Target Detection : An fMRI Study of Visual and Auditory Oddball Tasks. *Cerebral Cortex*, *9*(8), 815–823. <http://doi.org/10.1093/cercor/9.8.815>
- Lomo, T. (1971). Patterns of Activation in a Monosynaptic Cortical Pathway : The Perforant Path Input to the Dentate Area of the Hippocampal Formation. *Experimental Brain Res*, *12*, 18–45. <http://doi.org/10.1007/bf00234414>
- Lomo, T. (2003). *The discovery of long-term potentiation*. Philos. Trans. R. Soc. London B Biol. Sci 358.
- Markman, E. M., & Abelev, M. (2004). Word learning in dogs ? *Trends in Cognitive Sciences*, *8*(11), 8–10.
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.
- Markram, H., Lubke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of Synaptic Efficacy by Coincidence of Postsynaptic APs and EPSPs. *Science*, *275*(213–15). <http://doi.org/10.1126/science.275.5297.213>
- Markson, L., & Bloom, P. (1997a). Evidence against a dedicated system for word learning in children. *Nature*, *385*(6619),

- Markson, L., & Bloom, P. (1997b). Evidence against a dedicated system for word learning in children. *Nature*. <http://doi.org/10.1038/385813a0>
- Marsh, H. L., Vining, A. Q., Levendoski, E. K., & Judge, P. G. (2015). Inference by Exclusion in Lion-Tailed Macaques (*Macaca silenus*), a Hamadryas Baboon (*Papio hamadryas*), Capuchins (*Sapajus apella*), and Squirrel Monkeys (*Saimiri sciureus*). *Journal of Comparative Psychology*, 129(3), 256–267.
- Martinos, M. M., Yoong, M., Patil, S., Chin, R. F. M., Neville, B. G., Scott, R. C., & De Haan, M. (2012). Recognition memory is impaired in children after prolonged febrile seizures. *Brain*, 135, 3153–3164. <http://doi.org/10.1093/brain/aws213>
- Masquelier, T., & Thorpe, S. J. (2007). Unsupervised Learning of Visual Features through Spike Timing Dependent Plasticity. *PLoS Computational Biology*, 3(2). <http://doi.org/10.1371/journal.pcbi.0030031>
- Mather, E., & Plunkett, K. (2009). Learning words over time: The role of stimulus repetition in mutual exclusivity. *Infancy*. <http://doi.org/10.1080/15250000802569702>
- Mather, E., & Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language*, 38, 933–950.
- Mattsson, B., Maliniema-Piispanen, S., & Aaltonen, J. (2015). The lost mother tongue: an interview study with Finnish war children. *The Scandinavian Psychoanalytic Review*, 38(2), 128–139. <http://doi.org/10.1080/01062301.2015.1119612>
- Mayor, J., & Plunkett, K. (2010). A neural computational account of taxonomic responding and fast mapping in early word learning. *Psychological Review*, 117(1), 1–31.
- McClelland, J. L. (2013). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, 142(4), 1190–2013.
- McClelland, J. L., McNaughton, B. L., & Reilly, R. C. O. (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.
- McDonough, L., Mandler, J. M., Mckeet, R. D., & Squire, L. R. (1995). The deferred imitation task as a nonverbal measure of declarative memory. *PNAS*, 92, 7580–7584.
- Mckee, R. D., & Squire, L. R. (1993). On the Development of Declarative Memory. *Journal of Experimental Child Psychology*, 19(2), 397–404.
- McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*. <http://doi.org/10.1126/science.1144073>
- McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word Learning Emerges From the Interaction of Online Referent Selection and Slow Associative Learning. *Psychological Review*, 119(4), 831–877. <http://doi.org/10.1037/a0029872>
- Mervis, C. B., & Bertrand, J. (1994). Acquisition of the Novel Name-Nameless Category (N3C) Principle. *Child Development*, 65(6), 1646–1662.
- Mervis, C. B., Michnick Golinkoff, R., Bertrand, J., & Michnick, R. (1994). Two-Year-Olds Readily Learn Multiple Labels for the Same Basic-Level Category. *Source: Child Development*, 65(4), 1163–1177. Retrieved from

- Miklósi, Á., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z., & Csányi, V. (2003). A Simple Reason for a Big Difference : Wolves Do Not Look Back at Humans , but Dogs Do. *Current Biology*, 13(9), 763–766. <http://doi.org/10.1016/S>
- Miklosi, A., & Soproni, K. (2006). A comparative analysis of animals ' understanding of the human pointing gesture. *Animal Cognition*, 9, 81–93. <http://doi.org/10.1007/s10071-005-0008-1>
- Miles, H. L. W. (1991). Book review: Teaching sign language to chimpanzees. *International Journal of Primatology*, 12(3), 303–307.
- Milgram, N. W., Adams, B., Callahan, H., Head, E., Mackay, B., Thirlwell, C., & Cotman, C. W. (1999). Landmark Discrimination Learning in the Dog. *Learning & Memory*, 6, 54–61.
- Milgram, N. W., Head, E., Weiner, E., & Thomas, E. (1994). Cognitive Functions and Aging in the Dog : Acquisition of Nonspatial Visual Tasks. *Behavioral Neuroscience*, 108(1), 57–68.
- Miller, P. E., & Murphy, C. J. (1995). Vision in dogs. *Journal of the American Veterinary Medical Association*, 207, 1623–1634.
- Mills, D., Coffey-Corina, S. A., & Neville, H. J. (1993). Language Acquisition and Cerebral Specialization in 20-Month-Old Infants. *Journal of Cognitive Neuroscience*, 5, 317–334. <http://doi.org/10.1162/jocn.1993.5.3.317>
- Milner, B. (1968). F U R T H E R ANALYSIS OF THE HIPPOCAMPAL AMNESIC S Y N D R O M E : 14-YEAR FOLLOW-UP S T U D Y OF H . M . *Neuropsychologia*, 6, 215–234.
- Moore, B. C. J. (1989). *An introduction to the Psychology of Hearing* (Academic P). London.
- Morris, R. G. M. (2007). Theories of hippocampal function. In P. Anderson, R. G. M. Morris, D. G. Amaral, T. V Bliss, & J. O'Keefe (Eds.), *The hippocampus book* (pp. 581–713). Oxford University Press.
- Muller, C. A., Schmitt, K., Barber, A. L. A., & Huber, L. (2015). Report Dogs Can Discriminate Emotional Expressions of Human Faces Dogs Can Discriminate Emotional Expressions of Human Faces. *Current Biology*, 1–5. <http://doi.org/10.1016/j.cub.2014.12.055>
- Munoz, M., Chadwick, M., Perez-Hernandez, E., Vargha-Khadem, F., & Mishkin, M. (2011). Novelty Preference in Patients With Developmental Amnesia. *Hippocampus*, 21, 1268–1276. <http://doi.org/10.1002/hipo.20836>
- Murphy, C. M., & Messer, D. J. (1977). Mothers, infants, and pointing: A study of gesture. In H. R. Schaffer (Ed.), *Studies in mother-infant interaction* (pp. 325–354). London: Academic Press.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing : A review. *Clinical Neurophysiology*, 118, 2544–2590. <http://doi.org/10.1016/j.clinph.2007.04.026>
- Nadel, L., & Moscovitch, M. (1997). Memory consolidation , retrograde amnesia and the hippocampal complex. *Current Opinion in Neurobiology*, 7, 217–227.
- Nazzi, T., & Bertoncini, J. (2003). Before and after the vocabulary spurt : two modes of word acquisition ? *Developmental Science*, 6(2), 136–142.

- Neitz, J., Geist, T., & Jacobs, G. H. (1989). Color vision in the dog. *Visual Neuroscience*, 3, 119–125. <http://doi.org/10.1017/S0952523800004430>
- Nelson, C. A. (1995). The Ontogeny of Human Memory: A Cognitive Neuroscience Perspective. *Developmental Psychology*, 31(5), 723–738.
- Nelson, K., & Gruendel, J. (1981). Generalized event representations: Basic building blocks of cognitive development. In M. Lamb & A. Brown (Eds.), *Advances in developmental psychology* (p. Vol. 1 pp 131-158). Hillsdale, NJ: Erlbaum.
- Nelson, K., & Gruendel, J. (1986). Children's scripts. In K. Nelson (Ed.), *Event knowledge: Structure and function in development* (pp. 21–46). Hillsdale, NJ: Erlbaum.
- Newcombe, N., Huttenlocher, J., Drummey, A. B., & Wiley, J. G. (1998). THE DEVELOPMENT OF SPATIAL LOCATION CODING : PLACE LEARNING AND DEAD RECKONING IN THE. *Cognitive Development*, 13, 185–200.
- Newcombe, N. S. (2015). Commentary: Memory development: Halfway there? *International Journal of Behavioral Development*, 39(4), 304–305. <http://doi.org/10.1177/0165025415573647>
- O'Keefe, J., & Dostrovsky, J. (1971). Short Communications The hippocampus as a spatial map . Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34, 171–175.
- Overman, W., Bachevalier, J., Turner, M., Peuster, A., Roland, P., Turnbull, D., ... Evans, L. (1992). Object Recognition Versus Object Discrimination : Comparison Between Human Infants and Infant Monkeys. *Behavioral Neuroscience*, 106(1), 15–29.
- Oviatt, S. L. (1980). The emerging ability to comprehend language: an experimental approach. *Child Development*, 51(1), 97–106. <http://doi.org/10.1111/j.1467-8624.1980.tb02514.x>
- Pascalis, O., & Bachevalier, J. (1999). Neonatal Aspiration Lesions of the Hippocampal Formation Impair Visual Recognition Memory When Assessed by Paired-Comparison Task but Not by Delayed Nonmatching-to-Sample Task. *Hippocampus*, 9, 609–616.
- Pascalis, O., & De Haan, M. (2003). Recognition Memory and Novelty Preference: What Model? In H. Hayne & J. Fagen (Eds.), *Progress in Infancy Research* (Lawrence E, pp. 95–120). New Jersey.
- Pascalis, O., De Haan, M., Nelson, C. A., Matheny, L. S., Thomas, K. M., & Carver, L. J. (1998). Long-Term Recognition Memory for Faces Assessed by Visual Paired Comparison in 3- and 6-Month-Old Infants We thank parents and infants for participating in the study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(1), 249–260.
- Pascalis, O., & De Schonen, S. (1994). Recognition memory in 3- to 4-day-old human neonates. *NeuroReport*, 5, 1721–1724.
- Pascalis, O., Hunkin, N. M., Bachevalier, J., & Mayes, A. R. (2009). Change in background context disrupts performance on visual paired comparison following hippocampal damage. *Neuropsychologia*, 47, 2107–2113. <http://doi.org/10.1016/j.neuropsychologia.2009.04.001>
- Pascalis, O., Hunkin, N. M., Holdstock, J. S., Isaac, C. L., & Mayes, A. R. (2004). Visual paired comparison performance is impaired in a patient with selective hippocampal lesions and relatively intact item recognition. *Neuropsychologia*, 42, 1293–1300. <http://doi.org/10.1016/j.neuropsychologia.2004.03.005>

- Patterson, K., Nestor, P., & Rogers, T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8, 976–987.
- Patterson, T. L., & Tzeng, O. J. L. (1979). Long-term memory for abstract concepts in the lowland gorilla (*Gorilla g . Gorilla*). *Bulletin of the Psychonomic Society*, 13(5), 279–282.
- Pedreira, C., Mormann, F., Kraskov, A., Cerf, M., Fried, I., Koch, C., & Quiroga, R. Q. (2010). Responses of Human Medial Temporal Lobe Neurons Are Modulated by Stimulus Repetition. *Journal of Neurophysiology*, 103, 97–107. <http://doi.org/10.1152/jn.91323.2008>.
- Pena, M., Werker, J. F., & Dehaene-lambertz, G. (2012). Earlier Speech Exposure Does Not Accelerate Speech Acquisition. *The Journal of Neuroscience*, 32(33), 11159–11163. <http://doi.org/10.1523/JNEUROSCI.6516-11.2012>
- Pepperberg, I. M. (2002). In Search of King Solomon ' s Ring : Cognitive and Communicative Studies of Grey Parrots (*Psittacus erithacus*). *Brain, Behavior and Evolution*, 59(1–2), 54–67.
- Pepperberg, I. M., Koepke, A., Livingston, P., Girard, M., & Hartsfield, L. A. (2013). Reasoning by Inference : Further Studies on Exclusion in Grey Parrots (*Psittacus erithacus*). *Journal of Comparative Psychology*, 127(3), 272–281. <http://doi.org/10.1037/a0031641>
- Pilley, J. W. (2013). Border collie comprehends sentences containing a prepositional object , verb , and direct object. *Learning and Motivation*, 44(4), 229–240. <http://doi.org/10.1016/j.lmot.2013.02.003>
- Pilley, J. W. (2014). *Chaser: Unlocking the genius of the dog who knows a thousand words* (Mariner Bo). Boston, New York.
- Pilley, J. W., & Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, 86(2), 184–195. <http://doi.org/10.1016/j.beproc.2010.11.007>
- Pongrácz, P., Molnár, C., & Miklósi, Á. (2010). Barking in family dogs : An ethological approach. *The Veterinary Journal*, 183(2), 141–147. <http://doi.org/10.1016/j.tvjl.2008.12.010>
- Pongrácz, P., Molnar, C., Miklosi, A., & Csanyi, V. (2005). Human Listeners Are Able to Classify Dog (*Canis familiaris*) Barks Recorded in Different Situations. *Journal of Comparative Psychology*, 119(2), 136–144. <http://doi.org/10.1037/0735-7036.119.2.136>
- Portes, A., & Hao, L. (1998). *E pluribus Unum: bilingualism and language loss in the second generation*.
- Prichard, A., Cook, P. F., Spivak, M., Chhibber, R., & Berns, G. S. (2018). Awake fMRI Reveals Brain Regions for Novel Word Detection in Dogs. *Frontiers in Neuroscience*, 12(737). <http://doi.org/10.3389/fnins.2018.00737>
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT press.
- Quiroga, R. Q., Fried, I., & Koch, C. (2013). Brain Cells for Grandmother. *Scientific American*, 308(2), 30–35. <http://doi.org/10.1038/scientificamerican0213-30>
- Quiroga, R. Q., Kraskov, A., Koch, C., & Fried, I. (2009). Report Explicit Encoding of Multimodal Percepts by Single Neurons in the Human Brain. *Current Biology*, 19, 1308–1313. <http://doi.org/10.1016/j.cub.2009.06.060>
- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature Letters*, 435, 1102–1107. <http://doi.org/10.1038/nature03687>

- Rakic, P. (1995). Corticogenesis in human and nonhuman primates. In *The cognitive neurosciences* (pp. 127–145). Cambridge, MA: MIT Press.
- Ramos, D., & Ades, C. (2012). Two-Item Sentence Comprehension by a Dog (*Canis familiaris*). *PLoS ONE*, 7(2), 1–7. <http://doi.org/10.1371/journal.pone.0029689>
- Ramos, D., & Mills, D. S. (2009). Responsivness of domestic dogs (*Canis familiaris*) to verbal OBJECT and ACTION commands. In *XXXI International Ethological Conference*. Rennes.
- Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs, 339–347. <http://doi.org/10.1007/s10071-007-0123-2>
- Raslau, F. D., Mark, I. T., Klein, A. P., Ulmer, J. L., Mathews, V., & Mark, X. L. P. (2015). Memory Part 2 : The Role of the Medial Temporal Lobe. *Journal of Neuroradiology*, 36, 846–849.
- Rehn, T., & Keeling, L. J. (2011). The effect of time left alone at home on dog welfare. *Applied Animal Behaviour Science*, 129(2–4), 129–135. <http://doi.org/10.1016/j.applanim.2010.11.015>
- Reichmuth Kastak, C., & Schusterman, R. J. (2002). Long-term memory for concepts in a California sea lion (*Zalophus californianus*). *Animal Cognition*, 5, 225–232. <http://doi.org/10.1007/s10071-002-0153-8>
- Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, 11(2), 145–166. <http://doi.org/10.1017/S0142716400008742>
- Ribordy, F., Jabès, A., Banta Lavenex, P., & Lavenex, P. (2013). Development of allocentric spatial memory abilities in children from 18 months to 5 years of age. *Cognitive Psychology*, 66, 1–29. <http://doi.org/10.1016/j.cogpsych.2012.08.001>
- Richmond, J., & Nelson, C. A. (2007). Accounting for change in declarative memory: A cognitive neuroscience perspective. *Developmental Review*, 27(3), 349–373. <http://doi.org/10.1016/j.dr.2007.04.002>
- Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2005). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science*, 8(2), 162–172.
- Robinson, A. J., & Pascalis, O. (2004). Development of flexible visual recognition memory in human infants. *Developmental Science*, 7(5), 527–533.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*. Elsevier Ltd. <http://doi.org/10.1016/j.tics.2010.09.003>
- Rolls, E. T. (2010). A computational theory of episodic memory formation in the hippocampus. *Behavioural Brain Research*, 215(2), 180–196. <http://doi.org/10.1016/j.bbr.2010.03.027>
- Rolls, E. T. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus FUNCTION. *Frontiers in Systems Neuroscience*, 7, 1–21. <http://doi.org/10.3389/fnsys.2013.00074>
- Rose, D., & Blakemore, C. (1974). An analysis of orientation selectivity in the cat's visual cortex. *Experimental Brain Res*, 20(1), 1–17.
- Rosengren, A. (1969). *Experiments in color discrimination in dogs: Societas pro fauna et flora Fennica*.

- Rovee-collier, C. (1997). Dissociations in Infant Memory : Rethinking the Development of Implicit and Explicit Memory. *Psychological Review*, 104(3), 467–498.
- Rovee-collier, C., & Cuevas, K. (2009). Multiple Memory Systems Are Unnecessary to Account for Infant Memory Development : An Ecological Model. *Developmental Psychology*, 45(1), 160–174. <http://doi.org/10.1037/a0014538>
- Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103(4), 734–760. <http://doi.org/10.1037/0033-295X.103.4.734>
- Sakhon, S., Edwards, K., Luongo, A., Murphy, M., & Edgin, J. O. (2018). Equivalent fast-mapping and explicit encoding word retention in typical and atypical development. In *ICIS*.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103(3), 403–428.
- Savage-Rumbaugh, S., & Lewin, R. (1994). *Kanzi: The Ape at the Brink of the Human Mind*. (Wiley, Ed.).
- Schacter, D., & Moscovitch, M. (1984). Infants, amnesics, and dissociable memory systems. In M. Moscovitch (Ed.), *Advances in the study of communication and affect* (p. Vol 9, 173-216). New York.
- Schafer, G., & Plunkett, K. (1998a). Society for Research in Child Development Wiley Rapid Word Learning by Fifteen-Month-Olds under Tightly Controlled Conditions Rapid Word Learning by Fifteen-Month-Olds under Tightly Controlled Conditions. *Source: Child Development Child Development*, 69(2), 309–320. Retrieved from <http://www.jstor.org>
- Schafer, G., & Plunkett, K. (1998b). Society for Research in Child Development Wiley Rapid Word Learning by Fifteen-Month-Olds under Tightly Controlled Conditions Rapid Word Learning by Fifteen-Month-Olds under Tightly Controlled Conditions. *Source: Child Development Child Development*, 69(2), 309–320.
- Scheider, L., Grassmann, S., Kaminski, J., & Tomasello, M. (2011). Domestic Dogs Use Contextual Information and Tone of Voice when following a Human Pointing Gesture. *PLoS ONE*, 6(7). <http://doi.org/10.1371/journal.pone.0021676>
- Schmolck, H., Kensinger, E. A., Corkin, S., & Squire, L. R. (2002). Semantic Knowledge in Patient H.M. and Other Patients With Bilateral Medial and Lateral Temporal Lobe Lesions. *Hippocampus*, 12, 520–533. <http://doi.org/10.1002/hipo.10039>
- Schusterman, R. J., & Krieger, K. (1984). CALIFORNIA SEA LIONS ARE CAPABLE OF SEMANTIC COMPREHENSION. *The Psychological Record*, 34(1), 3–23.
- Schwab, J. ., & Lew-Williams, C. (2016). Repetition Across Successive Sentences Facilitates Young Children’s Word Learning. *Developmental Psychology*.
- Scoville, W. B., & Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL. *Journal of Neurosurgical Psychiatry*, 20, 11–21.
- Seress, L. (2001). Morphological changes of the human hippocampal formation from midgestation to early childhood. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 45–58). Cambridge, MA.
- Seress, L., & Abraham, H. (2008). Pre- and postnatal morphological development of the human hippocampal formation. In C. A. Nelson & M. Luciana (Eds.), *Developmental cognitive neuroscience. Handbook of developmental cognitive*

neuroscience (pp. 187–211). Cambridge, MA.

- Sharon, T., Moscovitch, M., & Gilboa, A. (2011). Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *PNAS*, *108*(3). <http://doi.org/10.1073/pnas.1005238108/-/DCSupplemental>.www.pnas.org/cgi/doi/10.1073/pnas.1005238108
- Shatz, C. J. (1996). Emergence of order in visual system development. *PNAS*, *93*, 602–608.
- Siegal, M., & Barlough, J. E. (1995). *The complete medical reference guide for dogs and puppies*. (H. Collins, Ed.) (UC Davis b). New York.
- Simcock, G., & Hayne, H. (2002). BREAKING THE BARRIER ? Children Fail to Translate Their Preverbal Memories Into Language. *Psychological Science*, *13*(3), 225–231.
- Smith, C. ., Urgolites, Z. ., Hopkins, R. ., & Squire, L. . (2014). Comparison of explicit and incidental learning strategies in memory-impaired patients. *PNAS*, *111*(1), 475–479.
- Smith, L. B., & Jones, S. S. (1988). The Importance of Shape in Early Lexical Learning. *Cognitive Development*, *3*(3), 299–321. [http://doi.org/10.1016/0885-2014\(88\)90014-7](http://doi.org/10.1016/0885-2014(88)90014-7)
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*, 1558–1568. <http://doi.org/10.1016/j.cognition.2007.06.010>
- Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology*. <http://doi.org/10.1016/j.jecp.2010.10.013>
- Squire, L. R. (1992). Declarative and Nondeclarative Memory: Multiple Brain Systems Supporting Learning and Memory. *Journal of Cognitive Neuroscience*, *4*(3), 232–243.
- Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. In *Neurobiology of Learning and Memory*. <http://doi.org/10.1016/j.nlm.2004.06.005>
- Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation : a neurobiological perspective. *Current Opinion in Neurobiology*, *5*, 169–177.
- Squire, L. R., Stark, C. E. L., & Clark, R. E. (2004). The Medial Temporal Lobe. *Annual Review of Neuroscience*, *27*, 279–306. <http://doi.org/10.1146/annurev.neuro.27.070203.144130>
- Squire, L. R., & Zola-morgan, S. (1991). The Medial Temporal Lobe Memory System. *Science*, *253*, 1380–1386.
- Tanaka, T., Watanabe, T., Eguchi, Y., & Yoshimoto, T. (2000). Color Discrimination. *Nihon Chikusan Gakkaiho*, *71*(3), 300–304.
- Thomas, K., Hunt, R., Sommer, T., Durston, S., Yang, Y., & Worden, M. (2004). Evidence of developmental differences in implicit sequence learning: an fMRI study of children and adults. *Journal of Cognitive Neuroscience*, *16*(8), 1339–51. <http://doi.org/10.1162/0898929042304688>
- Thorpe, S. (2011a). Grandmother cells, neocortical dark matter and very long term visual memories. *Journal of Vision*, *11*(11), 1243–1243.
- Thorpe, S. (2011b). Grandmother Cells and Distributed Representations. In N. Kriegeskorte & G. Kreiman (Eds.),

- Understanding visual population codes - Towards a common multivariate framework for cell recording and functional imaging.* Cambridge, MA: MIT press.
- Thunell, E., & Thorpe, S. J. (2019a). Memory for Repeated Images in Streams of Thousands of Images. *Psychological Science*, 1–12. <http://doi.org/10.1177/0956797619842251>
- Thunell, E., & Thorpe, S. J. (2019b). Regularity is not a key factor for encoding repetition in rapid image streams. *Scientific Reports*, 1–10. <http://doi.org/10.1038/s41598-019-39697-y>
- Tincoff, R., & Jusczyk, P. W. (1999). SOME BEGINNINGS OF WORD COMPREHENSION IN 6-MONTH-OLDS. *PSYCHOLOGICAL SCIENCE*, 10(2), 172–175.
- Tramo, M. J., Loftus, W., Thomas, C. E., Green, R. L., Mott, L. A., & Gazzaniga, M. S. (1995). Surface Area of Human Cerebral Cortex and Its Gross Morphological Subdivisions : In Vivo Measurements i n Monozygotic ' Wins Suggest Differential Hemisphere Effects of Genetic Factors. *Journal of Comparative Neuroscience*, 7(2), 292–301.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (Academic P, pp. 382–402). New York, NY.
- Tymofiyeva, O., Hess, C. P., Ziv, E., Lee, P. N., Glass, H. C., Ferriero, D. M., ... Xu, D. (2013). A DTI-Based Template-Free Cortical Connectome Study of Brain Maturation. *PLoS ONE*, 8(5), 1–10. <http://doi.org/10.1371/journal.pone.0063310>
- Van Der Zee, E., Zulch, H., & Mills, D. (2012). Word Generalization by a Dog (*Canis familiaris*): Is Shape Important ? *PLoS ONE*, 7(11). <http://doi.org/10.1371/journal.pone.0049382>
- Vargha-Khadem, F. (1997). Differential Effects of Early Hippocampal Pathology on Episodic and Semantic Memory. *Science*. <http://doi.org/10.1126/science.277.5324.376>
- Vilà, C., Savolainen, P., Maldonado, J. E., Amorim, I. R., John, E., Honeycutt, R. L., ... Honeycutt, R. L. (1997). Multiple and Ancient Origins of the Domestic Dog. *Science*, 276(5319), 1687–1689.
- Viswanathan, J., Rémy, F., Bacon-Macé, N., & Thorpe, S. J. (2016). Long term memory for noise: Evidence of robust encoding of very short temporal acoustic patterns. *Frontiers in Neuroscience*. <http://doi.org/10.3389/fnins.2016.00490>
- Vlach, H. A., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology*, 3(FEB), 1–8. <http://doi.org/10.3389/fpsyg.2012.00046>
- Wallis, L. J., Virányi, Z., Müller, C. A., Serisier, S., Huber, L., & Range, F. (2016). Aging effects on discrimination learning , logical reasoning and memory in pet dogs. *AGE*, 38(6). <http://doi.org/10.1007/s11357-015-9866-x>
- Warren, D. ., & Duff, M. . (2014). Not so fast: hippocampal amnesia slows word learning despite successful fast mapping. *Hippocampus*, 24(8), 920–933.
- Waxman, S. R., & Booth, A. E. (2000). Principles that are invoked in the acquisition of words, but not facts. *Cognition*, 77(2). [http://doi.org/10.1016/S0010-0277\(00\)00103-7](http://doi.org/10.1016/S0010-0277(00)00103-7)
- Wayne, R. K. (1986). Cranial morphology of domestic and wild canids: the influence of development on morphological change. *Evolution*, 40(2), 243–261.
- Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by

- 14-month-old infants. *Developmental Psychology*, 34(6), 1289–1309. <http://doi.org/10.1037/0012-1649.34.6.1289>
- White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59, 114–132.
- Wilhelm, I., Born, J., Friederici, A. D., & Friedrich, M. (2015). Generalization of word meanings during infant sleep. *Nature Communications*, 6, 1–9. <http://doi.org/10.1038/ncomms7004>
- Wojcik, E. H. (2013). Remembering new words: Integrating early memory development into word learning. *Frontiers in Psychology*, 4(APR), 1–12. <http://doi.org/10.3389/fpsyg.2013.00151>
- Wojcik, E. H. (2017). 2.5-Year-Olds' Retention and Generalization of Novel Words across Short and Long Delays. *Language Learning and Development*, 13(3), 300–316. <http://doi.org/10.1080/15475441.2016.1275644>
- Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994a). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*. <http://doi.org/10.1037/0012-1649.30.4.553>
- Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994b). Rapid word learning in 13-and 18-month olds. *Developmental Psychology*, 30(4), 553–566.
- Yakovlev, P. I., & Lecours, A. (1967). The myelogenetic cycles of regional maturation of the brain. In A. Minokowski (Ed.), *Regional development of the brain in early life* (pp. 3–70). Philadelphia: Davis.
- Zentall, T. R. (2006). Imitation: definitions, evidence, and mechanisms. *Animal Cognition*, 9, 335–353. <http://doi.org/10.1007/s10071-006-0039-2>
- Zola, S. M., & Squire, L. R. (2000). The medial temporal lobe and the hippocampus. In E. Tulving & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (Oxford Uni, pp. 485–500). New York.
- Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal Contrast: Competition Between Two Referents Improves Word Learning. *Applied Developmental Science*, 17(1), 20–28. <http://doi.org/10.1080/10888691.2013.748420>

Résumé substantiel

La vie est partout. Du plus petit organisme unicellulaire à la créature vivante la plus complexe. Mais la vie est aussi un enjeu. Pour vivre, les organismes doivent interagir et évoluer dans des environnements parfois étrangers voir périlleux. Le processus de sélection naturelle va favoriser l'émergence de certaines espèces mais aussi le déclin de d'autres. Pour survivre, être compétitif et se reproduire, une grande partie des Etres vivants sont pourvus d'étonnantes capacités d'adaptation. L'adaptation comportementale, qu'elle s'effectue de façon consciente ou réflexe, exigent des animaux qu'ils « *apprennent* » de leur environnement. Ils doivent notamment traiter une multitude d'informations à chaque instant, et si dotés de cette capacité, conserver ces informations pour une utilisation ultérieure. Pouvoir stocker des informations en mémoire confère aux Etres Vivants un extraordinaire avantage pour survivre dans la nature. Mais comment des espèces aussi diverses et génétiquement éloignées peuvent-elles apprendre et potentiellement mémoriser des informations ? Existe-il des principaux généraux de formation de la mémoire ? Y-a-t' il des mécanismes communs qui gouvernent les règles d'apprentissage ? Et comment pouvons-nous, humains, explorer ces règles et mécanismes chez un large panel d'espèces afin de mieux appréhender si le fonctionnement de l'apprentissage et de la mémoire a un caractère « universel »?

Voilà maintenant des décennies voir des siècles que ces questions passionnent les chercheurs. Cependant, comprendre comment une information est physiquement traitée puis stockée en mémoire n'est pas une mince affaire. Afin d'explorer ces questions, les chercheurs de placent à différentes échelles organisationnelles : moléculaire, cellulaire, tissulaire, comportementale, et depuis peu, computationnelle. Ils s'intéressent à des espèces très variées, allant d'organismes simples à des organismes parfois très complexes, avec à chaque fois la volonté de mieux comprendre si ces espèces partagent des traits communs ou bien si l'évolution a façonnée des modalités d'apprentissage et de mémorisation très distinctes les unes des autres. D'une part, certains chercheurs ont pour objectif de révéler des modèles animaux qui calquent le plus fidèlement possible le modèle humain, et ce, essentiellement pour des fins cliniques (par exemple, mettre au jour un modèle animal convaincant de la maladie d'Alzheimer afin de mieux étudier cette pathologie). D'autre part, d'autres chercheurs ont pour but d'étudier les mécanismes d'apprentissage et de formation de mémoires chez l'Homme, et de déterminer s'ils sont propres à l'Homme, ou si à l'inverse, ils sont largement répandus au sein du règne animal. D'un point de vue évolutionniste, ce type de recherche permet de mieux comprendre comment certains mécanismes auraient évolués et se seraient complexifiés au cours de l'évolution.

Au cours de cette thèse, nous avons cherché à savoir si certains des principes qui sous-tendent la formation de mémoires à long-terme chez l'Homme adulte, trouvent un écho chez des espèces animales relativement éloignées de l'Homme, ainsi qu'au cours du développement de l'Homme. Plus précisément, cette thèse vise à élucider si certains des principes énoncés par le projet M4 (Memory Mechanisms in Man and Machine) porté par S.Thorpe et mis en évidence chez l'Homme adulte, sont généralisables à d'autres espèces ainsi qu'aux premières années de vie de l'Homme. Les travaux de cette thèse ont été effectués à un niveau comportemental. Cependant, nous avons tenté de mettre nos résultats en relation avec les connaissances générales sur les corrélats neuraux aussi souvent que cela était possible. Cela permet de mieux comprendre les mécanismes cérébraux mis en jeu lors d'un comportement donné. C'est pourquoi cette thèse se situe à l'interface entre comportement, cognition et neurosciences.

Pour modèle d'étude, nous avons choisi le *chien domestique*, car bien que méconnu au sein de la communauté scientifique il y a seulement 20 ans de cela, il devient aujourd'hui un modèle de choix, notamment grâce à ses incroyables capacités de communications et de coopération avec l'Homme. De plus, un florilège d'études récentes ont mis en lumière d'étonnantes performances cognitives chez cette espèce, suggérant que malgré des divergences anatomiques, morphologiques et fonctionnelles notables avec l'Homme, le chien domestique a sû se parer d'atouts cognitifs largement insoupçonnés jusqu'à récemment. Il ne faut pas oublier que cette espèce est la première que l'Homme ait domestiquée, et que les millénaires de cohabitation avec l'Homme ainsi qu'une sélection génétique accrue effectuée par ce dernier, ont probablement été les principaux ingrédients permettant l'émergence de telles capacités cognitives, notamment de communication avec l'homme. C'est donc sur des critères de coopération et d'aisance de communication entre cette espèce et la nôtre, que nous avons choisi le chien comme modèle d'étude distant de l'Homme, pour étudier l'apprentissage et la mémoire de stimuli sensoriels complexes.

Par ailleurs, la façon dont les humains emmagasinent et stockent des informations précocement au cours de leur développement reste largement méconnue. Est-ce qu'un bébé ou un jeune enfant apprend et retiens de la même façon qu'un adulte ? A l'inverse, est-ce qu'apprentissage et mémoire sont des capacités cognitives qui s'acquièrent progressivement au cours du développement et qui dépendraient de corrélats neuraux mûres et fonctionnels ?

Pour ces deux modèles d'étude, nous nous sommes tout d'abord demandés si des informations sensorielles arbitraires pouvaient être conservées en mémoire à long terme sans qu'il n'y ait eu de réactivation en mémoire durant la période d'intervalle. Pour cela, nous nous sommes basés sur un concept abstrait complexe que les jeunes enfants mais aussi apparemment les chiens sont capables d'acquérir, celui que des mots peuvent désigner des objets. Nous avons notamment exploré

certaines paramètres susceptibles de faciliter l'encodage et le stockage en mémoire de ce type d'information. Parmi eux, nous avons cherché à savoir si le nombre de présentations de nouvelles associations mot-objets lors de l'apprentissage avait un effet sur la rétention du nom de ces objets après un délai. En effet, un des postulats du projet M4 propose que les poids synaptiques de neurones sélectifs (donc la trace mnésique) accroîtraient linéairement avec le nombre de répétitions de l'information lors de la phase d'encodage. En lien avec ce postulat, pour chacun de ces deux modèles d'étude, nous avons essayé de déterminer le nombre minimal de présentations d'une paire mot-objet nécessaire au stockage de cette nouvelle information en mémoire à long terme. Enfin, pour aller plus loin, l'effet de deux stratégies d'apprentissage sur la mémorisation de nouveaux mots ont particulièrement attiré notre curiosité : l'« *ostensive naming* » (à savoir le fait de nommer ostensiblement un objet nouveau ; dans ce cas l'apprentissage est dit *passif*) et le « *fast-mapping* » (c'est-à-dire le fait d'inférer un mot nouveau au seul objet inconnu d'une scène en utilisant une stratégie de déduction logique ; dans ce cas l'apprentissage est *actif*). Nous tenons à préciser que le but de cette thèse n'était pas de simplement comparer les performances des enfants à celles des chiens lors des différentes tâches expérimentales mises en œuvre, mais bien de comprendre si des règles communes à ces deux modèles d'études pouvaient être identifiées.

Mémoire à long terme d'associations mot-objets chez le jeune enfant

Cette première partie comprend six études, trois étant achevées (*études 1, 2 & 3*) et trois en cours (*études 4, 5 & 6*).

Des enfants de deux (et parfois trois) classes d'âge ont été recrutés, ainsi que des adultes naïfs quant à l'objectif de leur recrutement. Les âges étudiés étaient ciblées avec précision:

- 18 mois pour les plus jeunes (étude 1), c.à.d. des enfants ne produisant généralement pas plus d'une cinquantaine de mots (i.e. explosion lexicale non effectuée), dont le pic de densité synaptique du gyrus denté de l'hippocampe (i.e. un des composant nécessaire à la consolidation des informations) n'est pas atteint et dont le système mnésique est immature (amnésie infantile) ;
- 24 mois (études 1 à 5), c.à.d. des enfants ayant pour la majorité une production verbale nettement supérieure à 50 mots (i.e. explosion lexicale effectuée) et ayant atteint le pic de densité synaptique du gyrus denté, mais ne possédant pas un système mnésique mature (amnésie infantile) ;
- 4 ans (études 1 à 6), c.à.d. des enfants à la production verbale fluide, et possédant un système mnésique davantage fonctionnel mais pas pleinement mature (amnésie juvénile) ;

- *adultes de 20 à 60 ans* (études 1 à 6, sauf 4), c.à.d. des individus possédant un système mnésique mature et parfaitement fonctionnel (seuls des sujets sains ont été inclus).

Lors de ces différentes études, l'apprentissage a mis en jeu trois associations mot-objets arbitrairement choisies et présentées sur un écran tactile. La capacité des participants à se souvenir du nom de ces objets fut systématiquement testée via un test de choix forcé sur écran, faisant intervenir les trois associations apprises auparavant.

L'ensemble de ces études a permis de démontrer que la performance des enfants à retenir des noms d'objets était corrélée au nombre de fois que ces associations leur avaient été présentées lors de l'apprentissage (étude 1 et 3). Nous avons montré qu'un minimum de 5 présentations d'une nouvelle association mot-objet était nécessaire pour qu'un enfant de 18 mois se souvienne du nom de cet objet après un délai de 30min, et à condition que cet objet fussent nommé de façon ostensive et sans ambiguïté (« ostensive naming », *étude 1*). Dans ce même contexte d'apprentissage, nous avons pu mettre en évidence pour la toute première fois qu'une seule exposition à un objet nouveau était suffisante pour qu'un enfant de 2 ans retienne le nom de cet objet (*étude 1 et 3*). Cela suppose, qu'au contraire de tout ce qui avait été promulgué jusqu'alors, une mémoire déclarative peut s'établir rapidement et sans requérir de répétitions très tôt dans l'enfance. Nous proposons que la diminution significative du nombre de répétitions nécessaire à la mémorisation de nouveaux mots observée entre 18 et 24 mois serait due à de meilleures capacités attentionnelles (*étude 1*) et à une plus grande expertise langagière des enfants (la plupart des enfants de 2 ans ont déjà effectué leur explosion lexicale, *étude 1*). Nous suggérons aussi que l'« ostensive naming » est une stratégie d'apprentissage très efficace durant les 2 premières années de vie. En revanche, lorsque l'apprentissage de noms d'objets s'effectue de façon active en requérant de l'enfant qu'il utilise une stratégie de type déduction logique pour inférer un mot nouveau à son référent (« fast-mapping »), des enfants de 2 ans ne manifestent aucune trace de rétention de ces associations après un délai de 30min (*étude 3*). Dans cette thèse, nous avons apporté la preuve qu'un apprentissage passif et non ambigu de nouveaux mots (« ostensive naming ») est significativement plus efficace pour induire une trace mnésique à l'âge de 2 ans qu'un apprentissage actif de type déduction logique (étude 4).

En revanche, cela ne semble pas être aussi vrai plus tard au cours du développement. En effet, nos études ont également permis de mettre en évidence un shift développemental dans l'efficacité des stratégies d'apprentissage. Nous avons montré qu'à l'âge de 4ans, un apprentissage actif nécessitant un raisonnement logique est tout aussi efficace pour apprendre et retenir le nom d'un nouvel objet qu'un apprentissage passif (étude 4). C'est pourquoi, nous proposons que ces deux types d'apprentissage seraient médiés par des régions distinctes du cerveau : celle supportant un

apprentissage passif serait fonctionnelle plus précocement au cours du développement que celle supportant un apprentissage actif. L'hypothèse alternative qu'une seule et même région cérébrale serait responsable de la consolidation d'un nouveau mot, mais que seule une forme plus mature de cette région permettrait le stockage d'un mot appris via un processus d'apprentissage actif, n'est pas exclue.

Durant cette première partie de thèse, une étude en cours suggère également l'importance de laisser écouler du temps avant de solliciter la mémoire des jeunes enfants (étude 5). Nos résultats préliminaires supposent que tester des enfants de 2 ou 4ans immédiatement après un apprentissage (ce qui est pourtant classiquement effectué dans la littérature scientifique) aurait un effet délétère sur le processus de consolidation, ou du moins, ne reflèterait pas les capacités mnésiques réelles des jeunes enfants. En effet, il est communément admis que la mémoire à court terme (et la mémoire de travail), dépendante(s) des régions préfrontales, sont les plus tardives à s'établir au cours du développement de l'Homme. Chez le jeune enfant, il semblerait que lorsque la mémoire à court terme est sollicitée, elle amène à des erreurs qui pourraient vraisemblablement être évitées si un certain laps de temps (30min au moins) s'écoule après l'acquisition des nouvelles informations.

Enfin, nous avons pu re-tester la mémoire de certains enfants de 4 ans ainsi que de certains sujets adultes après un délai d'1 mois ou 6 mois afin d'évaluer le maintien dans la durée de la mémoire (*étude 6*). Précisons que le choix de nos stimuli, à savoir des objets inventés associés à des pseudo-mots, ont permis de nous affranchir de tout risque de réactivation de la mémoire chez ces participants pendant toute la durée de latence. Nos résultats actuels montrent que la mémoire reste relativement intacte 1 mois après l'apprentissage chez des enfants de 4 ans ainsi que chez des adultes (*étude 6*). En revanche, seule la mémoire des adultes semblent survivre à des latences plus longues.

Pour récapituler, il semblerait que dès très tôt au cours du développement, des enfants soient capables d'apprendre et mémoriser des informations de nature sémantique, telles que des nouveaux noms d'objets, et ce, malgré un très faible nombre d'expositions à ces informations. En revanche, la conservation en mémoire à très long terme de ces connaissances semblent compromise. Il se pourrait donc que l'amnésie infantile (et juvénile), majoritairement décrite dans le contexte de la mémoire épisodique, touche également la mémoire sémantique. Nous proposons, que tout comme chez l'adulte, une réactivation régulière des connaissances serait la clef pour prévenir d'une inaccessibilité à l'information durant l'enfance. Elle permettrait un renforcement des poids synaptiques des neurones mis en jeu et donc un maintien de la trace mnésique dans le temps. Dans cette thèse, nous avons également démontré qu'au-delà de la nature de l'information,

la stratégie mise en jeu pour acquérir cette information joue un rôle prépondérant dans la capacité de rétention de cette connaissance et que l'efficacité des stratégies d'apprentissage évolue au cours du développement.

Mémoire à long terme d'associations mot-objets chez le chien

Cette deuxième partie de thèse comporte deux études majeures, l'une impliquant un échantillon de 40 chiens de toutes races et âges confondus (*étude 1*), et la seconde, une étude de cas, impliquant le chien le plus expert de la cohorte. Cette étude de cas est elle-même constituée de 3 études exploratoires (*étude 2, 3 & 4*). L'ensemble des études se sont déroulées en partenariat avec l'Ecole Nationale Vétérinaire de Toulouse et une éducatrice canine spécialisée et ont eu lieu au sein de l'école vétérinaire de Toulouse.

A la différence des enfants, les chiens que nous avons recrutés étaient tous naïfs quant au fait que des mots peuvent désigner des objets. Il a donc fallu dans un premier temps leur enseigner le concept mot-objet. Cela a soulevé la question suivante : tous les chiens sont-ils capables d'acquérir le concept mot-objet ?

La littérature montre que les chiens sont capables d'associer une signification à un mot (notamment pour des actions, ex. *assis, coucher, debout, vient, stop, tourne à droite, à gauche, mords, prends, donne*, etc.) essentiellement par le biais de conditionnements classiques et opérants. Plus récemment, des études de cas ont montré que des chiens pouvaient également associer un mot à un item neutre, et que certains d'entre eux possédaient un « répertoire lexical » de plusieurs centaines voire un millier de noms d'objets. Cela amène donc tout d'abord à la question de l'extension de la compréhension de concepts humains complexes, tels que le concept mot-objet, chez cette espèce non dotée de production verbale.

Quarante chiens ont été recrutés sur la base du volontariat de leur maître. Nous n'imposons pas de critères de race, de sexe, ni d'âge, mais les animaux devaient avoir un niveau minimum d'obéissance et d'éducation (évaluation du comportement général des chiens effectuée lors de l'inclusion à l'étude).

Les chiens inclus ont ensuite été entraînés quotidiennement à l'apprentissage du concept mot-objet par leur maître, l'éducatrice canine et moi-même, en suivant scrupuleusement les protocoles mis au point par les études pionnières dans ce domaine. Des méthodes complémentaires ont également été développées par nos soins. Après six mois d'entraînement, la performance des chiens à identifier les objets associés à leur label a été évaluée en conditions contrôlées (i.e. salle expérimentale agréée, objets disposés aléatoirement par l'expérimentateur derrière un rideau

opaque, générateur aléatoire de l'objet cible à chaque essai, expérimentation entièrement filmée, etc.). Nos résultats ont montré que la cohorte de chiens affichait un niveau de performance significativement supérieur au hasard (niveau de hasard calculé via des simulations de Monte Carlo). Cependant, une forte variabilité interindividuelle était constatée, tant au niveau du nombre de noms d'objets appris en 6 mois (de 2.5 à 8) qu'au niveau de la performance en elle-même (de 20 à 80% de réussite). Nous avons démontré que l'âge était un facteur crucial dans l'acquisition d'une telle tâche cognitivement complexe. En effet, nous avons pu mettre en évidence que seul le groupe constitué de chiots (chiens <1 an) exécutait la tâche de façon significative. Les jeunes chiens avaient en moyenne appris un plus grand nombre de noms d'objets et étaient significativement plus performants au test que les chiens adultes.

En réponse à notre toute première question, il semblerait donc que les chiens ne soient pas tous en mesure d'acquérir le concept mot-objet (du moins, pas en 6 mois d'entraînement quotidien) et que l'acquisition de ce type de tâche est facilitée au cours du développement du chien, voire réservée aux chiots. Cela est concordant avec le fait que chez l'Homme, la plupart des apprentissages complexes sont effectués au cours des premières années de vie. On peut alors supposer que le chien, tout comme l'Homme, bénéficie d'une plasticité cérébrale accrue au cours du développement qui serait très certainement responsable de l'efficacité des apprentissages précoces.

Nous nous sommes ensuite demandés si les chiens pouvaient conserver durablement en mémoire les noms des objets appris lors de la phase d'entraînement et si la mémoire à très long terme était également âge-dépendante chez cette espèce. Nos résultats démontrent que la cohorte dans son ensemble parvenait de façon significative à se souvenir des noms des objets après un délai de 3 mois (aucune réactivation des informations n'a eu lieu durant ce délai). Pour la première fois, nous démontrons l'existence d'une mémoire à long terme pour des stimuli sensoriels bimodaux complexes chez le chien. Nous avons également mis en évidence que les chiots étaient significativement meilleurs lors du test de mémoire à long terme que les chiens adultes.

Dans un second temps, nous nous sommes demandés si l'apprentissage et la mémorisation de nouveaux mots chez des chiens familiers avec le concept mot-objet suivent des règles similaires à celles établies chez l'homme adulte ou chez l'enfant. Afin de pouvoir effectuer un quelconque parallèle avec les principes énoncés chez l'humain, il était indispensable de pouvoir contrôler précisément le nombre de présentations des nouvelles associations mot-objet à apprendre. Pour cela, seul un chien expert de la tâche pouvait faire l'objet des études que nous souhaitions réaliser. Malheureusement, à l'issue des 6 mois d'entraînement, un seul chien était suffisamment convaincant et constant dans sa capacité à identifier des objets par leur noms. C'est pourquoi, ces

expérimentations ont été uniquement menées avec ce chien le plus expert de la cohorte, Moka, et ne constituent que des études exploratoires et préliminaires. Ce chien, âgé de 6 mois au début de la phase d'entraînement (groupe chiot), avait environ 1 an et un « répertoire lexical » d'une dizaine de noms d'objets lorsque ces études exploratoires ont débutées.

Au cours de ces différentes études, nous avons pu montrer que les performances mnésiques de Moka à se remémorer le nom des objets étaient stables au cours du temps (étude 2), suggérant qu'une fois l'information acquise, elle pourrait être maintenue en « dormance » des mois, voire des années, tout comme cela a été démontré chez l'Homme dans d'autres contextes.

Nous avons également pu établir que la performance de Moka à apprendre des nouveaux noms d'objets était positivement corrélée au nombre de présentations de ces associations (étude 3). En revanche, à l'inverse de ce que nous avons pu montrer chez les jeunes enfants, ni une seule ni même plusieurs présentations des nouvelles paires mot-objet n'ont suffi à induire une rétention après un délai de 30min et dans un contexte relativement similaire d' « ostensive naming » (étude 3). Nous avons montré qu'un minimum de 25-30 répétitions des labels en présence de leurs objets correspondants, répartis sur trois sessions d'apprentissage distribuées dans le temps et couplés à des tests intermédiaires faisant intervenir des renforcements positifs, étaient nécessaires pour observer une rétention significativement supérieure au hasard (étude 3). Il n'est cependant pas possible de conclure ici qu'un chien, à l'inverse d'un humain (adulte ou enfant) ne puisse apprendre une information en un seul essai car la compréhension de la tâche ainsi que le « répertoire lexical » de Moka étaient largement inférieurs à ceux de n'importe lequel des enfants testés au cours de cette thèse. Une réplication de ce protocole avec un chien ayant un « répertoire lexical » plus riche, permettrait peut-être d'élucider si la formation et consolidation d'une mémoire sensorielle peut s'effectuer sans répétition chez cette espèce.

En revanche, nous pouvons affirmer que chez ce chien, l' « ostensive naming » n'était pas une méthode d'apprentissage efficace pour un maintien durable en mémoire du nom d'un objet (étude 3). Nous avons cependant révélé que l'apprentissage de nouveaux noms d'objets devaient faire intervenir des tests intermédiaires avec renforcement positif pour que Moka conservent le nom de ces objets après un délai d'un mois (étude 3).

Par ailleurs, les études de cas rapportées dans la littérature montrent que des chiens possédant un répertoire de plusieurs centaines voire un millier de noms d'objets étaient capables d'inférer le nom d'un nouvel objet par déduction logique (« fast-mapping ») et de retenir cette information après un délai d'un mois. Ici, nous montrons que Moka parvient également, mais fragilement, à utiliser une stratégie de type déduction logique pour identifier l'objet correspondant à un nouveau

label, malgré son « répertoire lexical » nettement inférieur à celui des cas rapportés dans la littérature (*étude 3*). En revanche, il ne semble pas capable de retenir le nom d'un objet appris de cette façon. Chez l'Homme, il est suggéré que l'apparition de ce type de raisonnement serait liée à la taille du répertoire lexical de l'enfant. Il se pourrait que cela soit également le cas chez le chien.

Enfin, nous avons cherché à mieux comprendre l'essence même du concept mot-objet chez le chien. S'agit-il d'un apprentissage purement associatif de « bas-niveau », qui ne reposerait que sur du conditionnement et qui serait dénué de toute signification sémantique ? Pour tenter d'y répondre, nous avons examiné les capacités de généralisation de Moka (*étude 4*). Nous avons démontré qu'il était parfaitement capable de reconnaître l'objet correspond à son label associé si son maître ou des playbacks de la voix de son maître le lui demandait. En revanche, il n'était pas capable de généraliser ses apprentissages à des voix inconnues, ni à des voix synthétisées par ordinateur (*étude 4*). Enfin, il affichait des capacités incertaines de généralisation à des objets similaires aux originaux. Il semble donc douteux que Moka ait acquis le concept mot-objet tel que défini chez l'Homme. Nous discutons de ce point dans cette thèse ainsi que des mécanismes sous-jacents potentiels.

Pour récapituler, ces études exploratoires nous ont permis de révéler des divergences notables dans la façon dont ce chien apprend des informations liées à un concept humain. Il ne semble pas capable d'apprendre via une simple dénomination ostensive, mais apprendrait essentiellement par essai-erreurs sur la base de renforçateurs. Cependant, nous avons montré que la performance de Moka à se remémorer du nom de ses objets ne déclinait pas au cours du temps. Il se pourrait que quelle que soit la méthode d'apprentissage, une fois l'information sensorielle acquise, elle serait stockée et conservée et façon inaltérée durant de longues périodes, tel que cela fut reporté chez l'Homme.

Dans cette thèse, il a été très frustrant de ne pas pouvoir explorer davantage les mécanismes mis en jeu chez le chien ainsi que les principes qui régissent l'apprentissage et la mémoire chez cette espèce dans le contexte précis d'un apprentissage associatif mot-objet. Il aurait été souhaitable que davantage de chiens aient parfaitement acquis le concept mot-objet à l'issue de la phase d'entraînement afin de pouvoir répliquer les protocoles sur davantage de sujets. Il aurait également été préférable de disposer d'un chien ayant une expertise de plusieurs années avec la tâche, ou une connaissance de plusieurs centaines de noms d'objets afin de pouvoir réellement faire la lumière sur d'éventuelles règles communes à ces deux modèles d'études.

En conclusion, nos résultats confortent le postulat que l'Homme n'est pas la seule espèce capable de former et conserver durablement en mémoire une information sensorielle faisant intervenir des

stimuli bimodaux et pour laquelle il n'y a pas eu de réactivation durant toute la durée d'intervalle. Le maintien à (très) long terme d'une information sensorielle chez le chien, dépendrait, tout comme chez l'Homme, du nombre d'expositions à cette information lors de l'apprentissage. En revanche, dans le cadre de ce concept précis lié au langage et spécifique à l'Homme, il se pourrait que seul l'Homme soit en mesure d'apprendre le nom d'un nouvel objet via une seule exposition et ce, dès très tôt au cours du développement. Le chien requerrait quant à lui, de multiples expositions et une méthode d'apprentissage différente de celle de l'Homme. Nos résultats montrent également que l'âge joue un rôle prépondérant, tant dans l'acquisition de nouvelles informations (montré ici chez le chien) que dans l'efficacité des stratégies employées (montré chez l'enfant) ou encore dans la conservation de ces informations (montré chez ces deux modèles d'étude).

Il serait intéressant d'examiner si l'ensemble de ces résultats est propre au concept mot-objet où généralisable à d'autres types d'informations sensorielles. Il serait également intéressant d'explorer si d'autres divergences peuvent être identifiées chez des espèces animales plus étrangères au mode de communication des humains et/ou plus éloignées génétiquement de l'Homme que le chien. Tester la mémoire après des latences plus longues pourrait également constituer une perspective à ce travail. Enfin, l'utilisation de techniques d'imagerie ou d'électroencéphalographie permettrait de faire la lumière sur d'éventuelles similitudes ou divergences quant aux mécanismes neuraux et neuronaux mis en jeu lors de ce type d'apprentissage.

Résumé court

Nous faisons tous l'expérience de nous souvenirs d'informations ou d'épisodes de vie vécus il y a fort longtemps. Mais qu'en est-il pour nos jeunes enfants et pour les espèces animales éloignées de nous ? Les enfants et les animaux forment-ils également des souvenirs qui peuvent demeurer intacts durant de longues périodes ? Et si tel est le cas, la formation de leurs souvenirs dépend-elle des mêmes règles que celles actuellement établies chez l'Homme adulte ? Cette thèse s'est intéressée à ces questions dans un contexte d'apprentissage associatif bimodal complexe: le concept mot-objet; chez le jeune enfant et le chien domestique. Pour ces deux modèles d'étude, nous avons exploré certains paramètres susceptibles de faciliter l'encodage et le stockage en mémoire de ce type d'information. Nous avons notamment cherché à savoir si le nombre de présentations de nouvelles associations mot-objets lors de l'apprentissage avait un effet sur la rétention du nom de ces objets après un délai, et avons tenté de déterminer le nombre minimal de présentations nécessaire pour induire une trace mnésique. Enfin, nous avons exploré l'efficacité de deux stratégies d'apprentissage sur la mémorisation de nouveaux noms d'objets et démontrons que l'efficacité des stratégies utilisées chez l'Homme évolue au cours du développement de l'enfant et ne semblent pas s'appliquer aux chiens. En bref, nos résultats apportent des éclaircissements quant à certains principes qui sous-tendent la formation de mémoires sensorielles chez un organisme en développement ainsi que chez une espèce animale non primate, et nous permettent d'émettre des hypothèses quant aux mécanismes cérébraux sous-jacents.

Short abstract

All of us occasionally remember information or personal events that occurred a long time ago. But what about young children and distantly related animal species? Do young children and animals also form memories that can last in time? And if so, does memory formation rely on the same principles than those established in human adults? This thesis intended to examine these questions in the context of an associative learning concept involving complex bimodal stimuli: the word-object concept; both in young children and domestic dogs. For these two models, we explored some parameters susceptible to facilitate the encoding and storage of this type of information in memory. In particular, we attempted to determine if the number of presentations of novel word-object associations during learning influenced the retention of the name of these objects after a delay. We also aimed to establish the minimal number of presentations of the pairs required to induce a memory trace. Finally, we examined the efficacy of two learning strategies on the ability to remember the names of novel objects and demonstrated that the efficacy of the strategies implemented in humans evolves during development and seems not to apply to dogs. In brief, our results enlightened some principles underlying the formation of sensory memories in an early-developing brain system as well as in a non-primate species, and allow us to make assumptions about the underlying brain mechanisms.