



HAL
open science

Apprentissage de dépendances non-adjacentes et traitement de grammaires supra-régulières chez le babouin et l'humain

Raphaëlle Malassis

► **To cite this version:**

Raphaëlle Malassis. Apprentissage de dépendances non-adjacentes et traitement de grammaires supra-régulières chez le babouin et l'humain. Sciences cognitives. Aix-marseille University, 2018. Français. NNT : 2018AIXM0134 . tel-04501374

HAL Id: tel-04501374

<https://hal.science/tel-04501374>

Submitted on 12 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Public Domain



Aix-Marseille Université

Ecole Doctorale Cognition, Langage, Education (ED 356)

Thèse pour l'obtention du grade de Docteur d'Aix-Marseille Université

Discipline : Psychologie

**Apprentissage de Dépendances Non-Adjacentes et
Traitement de Grammaires Supra-Régulières
chez le Babouin et l'Humain**

Présentée par

Raphaëlle Malassis

Sous la direction des Docteurs Joël Fagot & Arnaud Rey

Soutenue publiquement le 15 Juin 2018

Jury :

Judit Gervain	Université Paris Descartes	Rapporteur
André Didierjean	Université de Franche-Comté	Rapporteur
Alban Lemasson	Université de Rennes 1	Examineur
Patrick Lemaire	Université d'Aix-Marseille	Examineur
Joël Fagot	Université d'Aix-Marseille	Directeur
Arnaud Rey	Université d'Aix-Marseille	Co-directeur

« Trace l'inégal palindrome. Neige. Bagatelle, dira Hercule. Le brut repentir, cet écrit né Perec. L'arc lu pèse trop, lis à vice-versa. Perte. Cerise d'une vérité banale, le Malstrom, Alep, mort édulcoré, crêpe porté de ce désir brisé d'un iota. [...] A toi, nu désir brisé, décédé, trope percé, roc lu. Détrompe la. Morts : l'Ame, l'Élan abêti, revenu. Désire ce trépas rêvé : Ci va ! S'il porte, sépulcral, ce repentir, cet écrit ne perturbe le lucre : Haridelle, ta gabegie ne mord ni la plage ni l'écart. »

Georges Perec, Au Moulin d'Andé, 1969.

Résumé

Une hypothèse dominant actuellement les théories sur l'évolution des capacités syntaxiques au cours de la phylogénèse est celle d'une spécificité humaine pour le traitement des grammaires supra-régulières. Cette hypothèse tend à être confirmée par les données comparatives actuellement disponibles, les différentes espèces testées dans des paradigmes d'apprentissage de grammaires artificielles ne parvenant pas à traiter ce type de grammaire. Dans cette thèse, nous avons cherché à décomposer le traitement des grammaires supra-régulières en opérations cognitives élémentaires (apprentissage de dépendances non-adjacentes et extraction de la structure sous-jacente). Nous proposons que ces opérations posent des difficultés pour les animaux non-humains, difficultés pouvant expliquer les échecs des précédentes expériences. L'objectif de ma thèse était donc double : (1) tester l'hypothèse selon laquelle la détection de régularités entre éléments non-adjacents présente une difficulté particulière pour le primate non-humain, et (2) évaluer la capacité du primate non-humain à traiter des grammaires supra-régulières dans des conditions expérimentales prévenant l'exploitation de régularités plus locales. Pour cela, nous avons mené une série de quatre études chez le babouin (*Papio papio*).

Dans une première étude, nous avons examiné les dynamiques temporelles de l'apprentissage de dépendances non-adjacentes chez l'humain et le babouin, et montré que ces derniers requéraient une quantité d'exposition beaucoup plus importante. Dans une seconde étude, les babouins étaient testés sur l'apprentissage de patterns basés sur des répétitions adjacentes (ABB) ou non-adjacentes (ABA). Les babouins sont parvenus à généraliser les deux patterns à de nouveaux éléments, mais se sont montrés davantage sensibles aux répétitions adjacentes. Avec une étude corrélacionnelle nous avons ensuite examiné les performances individuelles des babouins en fonction du type de tâche et du type de dépendances (adjacentes *versus* non-adjacentes). Les analyses effectuées révèlent une influence majeure de ce dernier facteur : les babouins parvenant à apprendre les dépendances non-adjacentes dans les différentes tâches ne sont pas ceux obtenant les meilleures performances pour l'apprentissage de dépendances adjacentes, et *vice-versa*. Dans une dernière étude, nous avons contrasté la sensibilité des babouins à une organisation centré-emboîtée *versus* croisée des dépendances non-adjacentes. Les babouins se sont montrés sensibles à cette première structure, mais pas à la seconde. Ces résultats suggèrent que le babouin possède des capacités computationnelles supra-régulières, mais que des contraintes mnésiques importantes pèsent sur celles-ci. Ces résultats sont discutés au regard des modèles théoriques actuels sur l'évolution des capacités syntaxiques.

Remerciements

Je remercie les membres de mon jury de thèse, les docteurs Gervain, Didierjean, Lemasson et Lemaire, pour avoir accepté de lire et évaluer mon manuscrit et de discuter de mes travaux.

Je remercie mes deux directeurs de thèse, pour m'avoir aidée et soutenue tout au long de ces années. Joël, merci pour m'avoir appris tant de choses. Pour m'avoir laissé une grande liberté d'action tout étant toujours présent. Un grand merci également pour avoir construit cette plateforme permettant d'étudier les babouins dans un certain respect de l'animal ; en espérant que cette logique continuera de prévaloir et de porter ses fruits, au sein de la plateforme et ailleurs. Arnaud, merci pour m'avoir fait découvrir ce champ de la psychologie, et avoir su régulièrement prodiguer les encouragements nécessaires.

Je remercie également l'ensemble des membres présents et passés de l'équipe "Cognition Comparée" du Laboratoire de Psychologie Cognitive, Adrien, Anaïs, Carole, Charlotte, Florence, Marie, Nicolas, Julie, Laure, Sandra, Thierry, pour nos discussions riches, et l'esprit d'entraide que vous faites régner au sein de l'équipe. Carole, je t'adresse un remerciement particulier, car à tes côtés, un miracle s'est produit : je me suis sentie bien dans une ville nord-américaine ; merci pour ces ballades et longues conversations, qui finissaient bien trop tard pour les françaises assidues aux conférences et en plein décalage horaire que nous étions. Marie, je t'adresse également tous mes remerciements, pour n'avoir cessé de mener des discussions de science qui interrogent, et ne jamais hésiter à emmener les plus jeunes dans ton sillage. Un remerciement particulier aussi à Julie, pour avoir accompagné mes premiers pas dans le monde des Papios, et à Laure, pour mes premiers pas dans Marseille et la vie de doctorante en psychologie cognitive. Un grand merci enfin à Nicolas, pour n'avoir jamais été avare de son temps pour discuter et apporter son aide ultra-giga-précieuse.

Merci aux personnes travaillant à la station de primatologie, Romain, Pao, Marie, Slaveia, Laura, Corinne, Richard, Valérie, Chloé, Jean-Christophe, Jean-Noël, Jo, Sébastien, Fidji, Célia et Brigitte, pour leur aide indispensable à la conduite de nos recherches.

Mes remerciements vont également aux membres du Laboratoire de Psychologie Cognitive pour leur bienveillance ; et notamment à Fred et Colette pour leur accompagnement dans le dédale des démarches administratives.

Je remercie également Stanislas Dehaene pour son investissement dans la collaboration que nous avons menée.

Je remercie aussi l'ensemble des enseignants du master éthologie de l'université Paris XIII, pour tout ce qu'ils m'ont appris. Je remercie tout particulièrement Gilles Gheusi, pour m'avoir accordé une telle confiance, et m'avoir accompagnée jusqu'aux babouins. Pour m'avoir fait découvrir la métacognition. Ça a été une chance fabuleuse pour moi de vous avoir comme professeur. Je remercie également Stéphane Chameron, pour ses cours, mais également car lorsque j'ai découvert en L1 que la personne qui animait ainsi les assemblées générales contre la loi LRU serait l'un de mes enseignants, je me suis définitivement sentie à ma place dans ce cursus.

Chers parents, merci, vous qui m'avez donné comme seule injonction de faire ce que j'aimais, et m'y avez aidé. Je remercie également mes petits frères pour ce qu'ils sont, je suis si fière de vous (si Paolo, c'est un sentiment tout à fait valable et très précieux).

Mamili & Papilou, merci pour vos visites si joyeuses, votre patience lorsqu'il s'agissait de m'attendre tandis que je restais cloîtrée dans mon bureau à analyser mes derniers résultats, et pour avoir partagé mon enthousiasme. Pour votre aide si précieuse pour prendre soin de mes bêtes, et votre tendresse inconditionnelle.

Merci à toi Mamirèze, pour être venue à bout de tous mes articles, pour t'intéresser à mes recherches, et pour m'avoir attendu. Merci papi Claude, pour m'avoir dit au revoir. Merci à tous les deux de n'avoir jamais manqué de me dire combien vous étiez fiers.

Et puis merci à toute ma famille, et tout particulièrement à Xavier, pour m'avoir fourni une voiture pourvue de freins décents, puis régulièrement approvisionné en musiques, éléments sans lesquels rien n'aurait été pareil sur les routes de Provence. Pour avoir été un tonton exemplaire, toujours vache, et toujours là. Un remerciement particulier également à Cathy, pour m'avoir fourni en références biblio qui continuaient de me faire entrevoir un ailleurs en dehors de mes lectures spécialisées ; tes doctorants sont bien chanceux de t'avoir qui veille.

Tiphaine, merci. Sans toi, je serais devenue une pierre. L'expérience de la thèse prend tout son sens lorsqu'elle permet de telles rencontres, i.e. du troisième type. Plus prosaïquement : ces heures passées à gribouiller sur ce tableau pour faire naître le design parfait – mais si, il doit exister ! – sous tes yeux compatissants, ces séances de décorticage de concepts, tes idées claires, ont été une aide immense. François, je te remercie pour ta luminosité (sic !), ta sérénité et ton enthousiasme communicatifs, tes choix cinématographiques – presque – toujours de qualité ; et avoir été, pour toutes ces raisons, le colocataire & ami idéal. Seb, merci pour avoir réveillé l'animal politique qui s'était endormi en moi. Merci pour m'avoir emmené ailleurs, des Chronyctalopes aux trésors du vallon. Yannick, merci pour ton amitié, pour la folie Aixoise. Heureusement, tu étais là.

Merci à la joyeuse bande de copains, Guillaume, Pauline, Agathe, Pierre-Arth, Jacky, Thibalt, Vincent, Shiraz, Flo, Tiff, Marine, Damien, Camille, Derek, Laurane, Christophe, Keshia, Raph, p'tit Guillaume. On en a de la veine de s'avoir.

Merci Flora pour ces jolis moments à Marseille et ailleurs – partout et toujours, avec toi, les moments sont jolis.

Merci Alex, pour ta détermination et ta patience exemplaires, qui nous ont permis de traverser ces quatre années ensemble. Je ne peux imaginer une personne ne faisant pas de la recherche qui comprenne aussi bien les enjeux et contraintes de ce milieu. Connivence des passionnés : continue à vivre la tienne à tout prix.

Et enfin, merci à celles qui sont ce que le monde contient de plus génial, Ellie, Lise & Chloé. Merci pour notre amitié. Merci de vous être rendues disponibles chaque fois que je remontais à Paris, cha-que fois. Avec une mention spéciale pour Ellie, dont la seule tare aura été de bien mal choisir son ultime coup de théâtre.

Table des matières

Préambule.....	11
VOLET THEORIQUE.....	13
Chapitre I : Introduction à l’Hypothèse de la Supra-régularité.....	14
I. Théories et Définitions	14
I.1 Introduction à la théorie des langages formels	14
I.2 Hiérarchie des langages formels	15
I.3 Organisation des langues en syntagmes récursivement enchâssés	22
I.4 Théories générativistes sur l’évolution de la syntaxe	24
I.5 Confusions à éviter : supra-régularité, arborescence, récursivité	27
II. Données Comparatives	32
II.1 Production	32
II.2 Apprentissage de grammaires artificielles	40
III. Conclusion : Proposition d’une Nouvelle Approche pour Etudier le Traitement des Grammaires Supra-régulières.....	68
Chapitre II : Prérequis Cognitifs pour le Traitement de Grammaires Supra-régulières.....	71
IV. Apprentissage d’Associations Non-adjacentes.....	71
IV.1 Données chez l’humain.....	72
IV.2 Données comparatives	76
IV.3 Mécanismes	81
IV.4 Conclusions.....	86
V. Extraction de Propriétés Structurelles	90
V.1 Données chez l’humain.....	91
V.2 Données comparatives	92
V.3 Mécanismes.....	98
V.4 Conclusions.....	101

Chapitre III : Problématique.....	104
VOLET EXPERIMENTAL	106
I. Méthodologie Générale	107
I.1 Expériences chez le primate non-humain	107
I.2 Expériences chez l'humain	115
II. ETUDE I – Traitement des Dépendances Non-adjacentes chez le Primate et Humain et Non-humain : Dynamiques Temporelles et Généralisation.....	117
III. ETUDE II – Extraction de Régularités Structurelles chez le Babouin : Contraster Répétitions Adjacentes et Non-adjacentes.....	152
IV. ETUDE III – Etude Corrélacionnelle : Performances Individuelles des Babouins pour l'Extraction de Régularités Adjacentes et Non-adjacentes	180
V. ETUDE IV – Traitement de Grammaires Supra-Régulières chez le Babouin	202
DISCUSSION GENERALE	216
I. Résumé et Discussion des Travaux Expérimentaux	217
I.1 Effet de la distance.....	218
I.2 Traitement de grammaires supra-régulières.....	223
I.3 Conclusions sur les procédures expérimentales.....	227
II. Implications pour les Théories sur l'Evolution de la Syntaxe.....	228
III. Perspectives : Etudier l'Evolution du Langage à travers l'Etude Comparative des Fonctions Cognitives Générales.....	232
IV. Conclusions	271
Glossaire.....	273
Références	275
ANNEXES	302
I. Etude IV: Supplementary Information	303
II. Etude sur la Prévalence des Informations Ordinales et Relationnelles : Données Comparatives chez le Babouin et l'Humain	310

III. Extraction de Régularités Séquentielles : Données Comparatives chez le Babouin et l'Humain	324
---	-----

Préambule

Le langage se caractérise par sa remarquable productivité, permettant un usage infini de moyens finis (Humboldt, 1836) en combinant les mots en phrases organisées selon une certaine syntaxe. L'étude comparative des capacités syntaxiques vise à caractériser les mécanismes cognitifs impliqués dans le traitement de la syntaxe, et leur histoire évolutive. On peut distinguer deux approches complémentaires. La première consiste à étudier l'organisation syntaxique, ou proto-syntaxique, des productions naturelles des animaux non-humains (ci-après, NH). La seconde approche s'intéresse aux capacités d'apprentissage de ces animaux dans des contextes induits expérimentalement. Cette approche permet (1) d'étudier finement les mécanismes et représentations en jeu lors du traitement de régularités syntaxiques, et (2) d'examiner la capacité de différentes espèces à traiter des régularités plus complexes que celles observées dans leurs productions naturelles.

Certaines théories, notamment celles issues de la linguistique générative, font l'hypothèse d'une discontinuité évolutive des traitements syntaxiques (e.g. Hauser, Chomsky, & Fitch, 2002). Dans leurs versions les plus radicales, ces théories proposent qu'au cours de l'évolution récente de la lignée humaine seraient apparues des mécanismes spécialisés permettant l'acquisition et le traitement du type de syntaxe complexe caractérisant les langues naturelles (Chomsky, 2017). D'autres théories font l'hypothèse d'une continuité évolutive des mécanismes sous-tendant le traitement de ces régularités. Ces théories ne négligent pas l'existence de différences de capacités syntaxiques entre les espèces, mais proposent que celles-ci découlent de différences quantitatives entre l'humain et les autres espèces observées pour les mécanismes d'extraction de régularités non spécifiques au langage, et notamment des mécanismes sous-tendant l'extraction de régularités séquentielles (e.g. Christiansen & Chater, 2015). Cette perspective ouvre la voie à l'étude de l'histoire évolutive de prérequis cognitifs pour le traitement de la syntaxe.

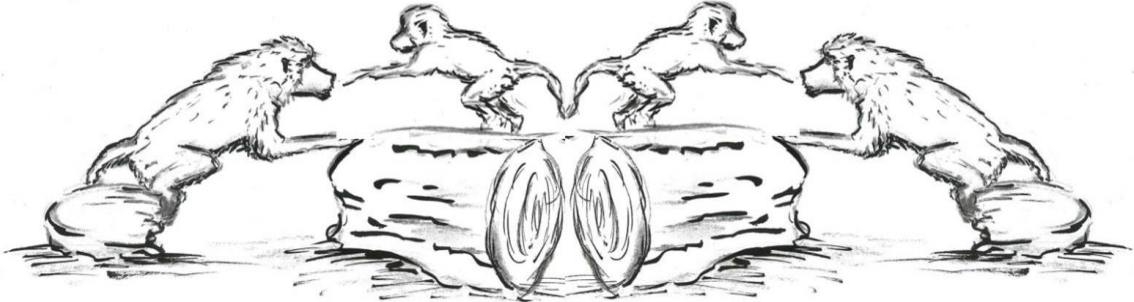
Dans ce cadre, cette thèse pose la question de l'existence d'une discontinuité entre l'humain et les autres animaux, ou bien d'une continuité évolutive, pour le traitement de régularités syntaxiques complexes. Pour cela, nous nous sommes intéressés aux capacités d'apprentissage de grammaires artificielles du babouin, un singe de l'ancien monde dont le dernier ancêtre commun avec l'humain remonterait à environ 35 millions d'années (Ciochon & Fleagle, 1987; Steiper, Young, & Sukarna, 2004). Nous nous sommes plus particulièrement

focalisés sur traitement de grammaires supra-régulières (Chomsky, 1956, 1957; Fitch & Friederici, 2012; Fitch & Hauser, 2004). Après une revue de la littérature sur le sujet, nous proposons une nouvelle approche consistant à examiner de manière comparative les prérequis cognitifs pour le traitement de ce type de grammaire.

Nous présentons ensuite une série de quatre études menées dans cette perspective, portant respectivement sur (1) l'apprentissage de dépendances non-adjacentes, (2) l'extraction de régularités structurelles, (3) l'analyse des performances individuelles dans différentes tâches d'extraction de régularités séquentielles, et enfin (4) le traitement de grammaires supra-régulières.

Dans un dernier volet, nous discutons les résultats obtenus au regard des théories sur l'évolution de la syntaxe, et proposons une ouverture sur l'étude de l'évolution du langage à travers l'étude comparative des fonctions cognitives dites *domain-general*.

VOLET THEORIQUE



Chapitre I : Introduction à l'Hypothèse de la Supra-régularité

I. Théories et Définitions

La linguistique générative est une approche initiée dans les années 50 par Noam Chomsky (1956, 1957), constituant aujourd'hui encore un courant majeur de cette discipline. Elle reprend les principes de la théorie des langages formels pour étudier l'organisation syntaxique des langues naturelles. Dans une perspective évolutive, elle se distingue des autres approches (e.g. linguistique fonctionnelle, Langacker, 1987, 1991 ; linguistique fondée sur l'usage, Bybee, 2010; Tomasello, 2003) par ses hypothèses fortes sur une discontinuité humain/non-humain. Nous verrons dans les sections suivantes que la mise à l'épreuve de ces hypothèses a prédominé dans les recherches comparatives menées ces dernières décennies. Avant d'examiner ces hypothèses, nous présentons donc ci-dessous quelques notions clés de la linguistique générative, après une brève introduction à la théorie des langages formels.

I.1 Introduction à la théorie des langages formels

La théorie des langages formels (en anglais, *formal language theory*, FLT) trouve ses sources dans les théories computationnelles élaborées en mathématiques (Turing, 1937) et en théorie de l'information (Shannon, 1948). Elle a pour objet la description formelle des systèmes régis par des règles (grammaires des langues naturelles, mais également de la musique, des langages informatiques, etc.).

L'étude des langues naturelles (e.g. le français, l'anglais) comme des systèmes formels a été initiée par Chomsky (1956), qui fait l'observation que le locuteur d'une langue sera généralement capable de juger si un énoncé est grammaticalement correct, même s'il est dépourvu de sens. Pour illustrer cette observation, il donne un exemple devenu célèbre, « *Colorless green ideas sleep furiously* ». Nous sommes capables de voir que cette phrase est

grammaticalement correcte, tandis que « *Furiously sleep ideas green colorless* » ne l'est pas. Pourtant, aucune de ces phrases ne fait sens, et un locuteur de l'anglais a peu de chances d'avoir déjà rencontré l'une des paires de mots les constituant l'une et l'autre, e.g. ce locuteur n'aura probablement jamais entendu ni « *sleep furiously* » ni « *furiously sleep* ». Par conséquent, selon Chomsky, ni la signification d'une phrase, ni ses propriétés distributionnelles, ne peuvent constituer la base des jugements de grammaticalité. Ces jugements s'effectueraient en vérifiant l'adéquation de l'exemplaire rencontré avec les règles grammaticales d'une langue donnée.

Dans la théorie des langages formels, un langage est défini comme un ensemble de séquences, ou d'expressions, générées à partir d'un ensemble fini de symboles constituant le vocabulaire (ou alphabet) de ce langage, et d'une grammaire donnée. Cette grammaire comprend l'ensemble des règles régissant l'organisation de ces symboles dans les séquences. La grammaire spécifie ainsi les opérations à partir desquelles peuvent être générées (ou dérivées) l'ensemble des expressions de ce langage, et seulement celles-ci. Ce dernier point est important : la formalisation d'une grammaire pour décrire un langage donné ne pourra être jugée adéquate que si elle respecte cette double contrainte. Ces opérations peuvent être spécifiées sous forme d'un ensemble de règles de substitution, ou de réécriture. Par exemple, l'expression "abab" peut être générée à partir d'un vocabulaire $V = \{a, b\}$, par l'application de l'ensemble des quatre règles suivantes : $\{ S \rightarrow aB ; B \rightarrow bA ; A \rightarrow aB ; B \rightarrow b \}$. Ces règles permettent de remplacer ou réécrire (\rightarrow) l'axiome S par une suite de symboles non-terminaux (en majuscule) devant être à leur tour remplacés par des terminaux (en minuscule ; correspondant aux symboles élémentaires, indécomposables, du langage).

I.2 Hiérarchie des langages formels

Cette hiérarchie, établie par Chomsky (1956, 1957), parfois simplement appelée « hiérarchie de Chomsky », est une typologie des différents types de grammaires et des automates correspondants, requis pour générer ou évaluer ces grammaires. On y distingue deux grands ensembles : les grammaires régulières et les grammaires supra-régulières.

1.2.1 Grammaires régulières

Les grammaires régulières, également appelées grammaires rationnelles, ou à états finis (en anglais, *finite-state grammars*), constituent la classe de grammaires nécessitant la puissance computationnelle la plus limitée. Une grammaire régulière comprend un état initial, un état final, et une succession d'états intermédiaires. On peut la représenter par un diagramme à états finis, comme exemplifié dans la Figure 1a, où chaque nœud du diagramme représente un état, et les flèches définissent les transitions légales d'un état à l'autre. Pour reconnaître qu'une expression correspond ou non à cette grammaire, l'automate dit, lui aussi, « à états finis », n'a qu'à faire « glisser » une « fenêtre » de vérification des transitions entre deux symboles, tout le long de la séquence (Figure 1b). La taille de cette mémoire peut éventuellement être plus importante (une chaîne de Markov d'ordre n pourra prendre en compte les n états précédents), mais elle est nécessairement limitée.

Encadré 1 Lexique

Automate : un modèle mathématique de computation ; également appelé « machine ».

Arbre syntaxique : mode de notation employé pour représenter la structure syntaxique d'une phrase (pour un exemple, voir la Figure 5).

Grammaire : ensemble des règles de dérivation à partir desquelles un langage donné peut-être généré.

Langage : ensemble de séquences ou d'expressions (e.g. phrases) générées à partir d'un vocabulaire et d'une grammaire donnée.

Merge : opération de « fusion », à partir de laquelle deux objets syntaxiques sont combinés, pour former un nouvel objet syntaxique situé à un niveau hiérarchique supérieur. L'objet syntaxique en résultant hérite de la nature de l'un de ses constituants (e.g. déterminant + nom = syntagme nominal ; adverbe + verbe = syntagme verbal).

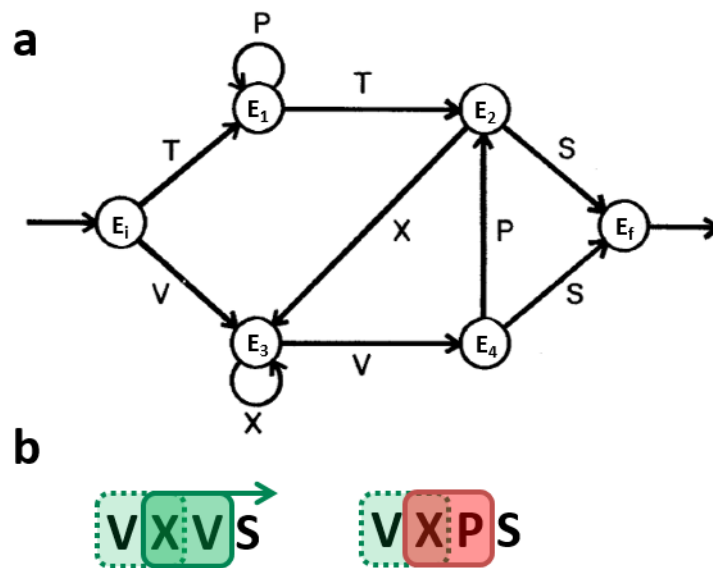


Figure 1 a. Exemple de diagramme représentant une grammaire régulière, d'après Reber (1967). E_i = état initial, E_{1-4} = états intermédiaires, E_f = état final. Le sens des flèches indique les transitions légales d'un état à l'autre. b. Représentation schématique de la vérification des transitions par un automate à états finis : la première séquence de lettres peut être générée à partir de la grammaire, tandis que la seconde comporte une transition illégale entre X et P. Dans cet exemple, l'automate ne prend en compte que le dernier élément rencontré.

Les grammaires régulières sont considérées comme non adéquates pour modéliser l'anglais, ou toute autre langue naturelle. Cette idée repose sur deux types d'arguments : une démonstration mathématique, et des raisons pratiques. La démonstration mathématique repose sur le fait qu'aucune grammaire régulière ne peut générer *l'ensemble* des phrases d'une langue. En effet, il existe des constructions linguistiques qui dépassent les capacités génératives de toute grammaire régulière, telles que les constructions centrées-emboîtées (e.g. en français, « L'homme qui a dit que (...), arrive aujourd'hui. » Dans cette phrase, une dépendance est entretenue entre « L'homme qui a dit que » et « arrive aujourd'hui », qui peuvent être séparés par un nombre arbitraire de propositions relatives emboîtées les unes dans les autres (e.g. « L'homme qui a dit que l'homme qui a dit que (...) tu avais tort, arrive aujourd'hui). Par conséquent, dans ces langues, il existe un nombre théoriquement infini d'expressions comportant ce type de construction (i.e. pour toute expression comprenant n occurrences de « L'homme qui a dit que », on peut former une nouvelle expression grammaticalement correcte comprenant $n+1$ occurrences). Or, aucune grammaire régulière ne peut générer un tel langage, de même qu'aucune ne pourra générer les trois langages artificiels suivants (pour une

démonstration, voir Chomsky, 1956; Miller & Chomsky, 1963 ; voir également Jäger & Rogers, 2012) :

- (i) $ab, aabb, aaabbb, \dots$, i.e. toutes les expressions comprenant n occurrences de "a" suivies par n occurrences de "b", i.e. $a^n b^n$, et seulement celles-ci ;
- (ii) $aa, bb, abba, baab, aaaa, bbbb, aabbaa, abbbba, \dots$, i.e. toutes les expressions comprenant une série X , suivie par l'image en miroir de X , et seulement celles-ci (noter que ce langage suit une organisation centrée-emboîtée comme notre exemple en français) ;
- (iii) $aa, bb, abab, baba, aaaa, bbbb, aabaab, abbabb, \dots$, i.e. toutes les expressions comprenant une série X , suivie par la répétition de cette même série X , et seulement celles-ci.

Notons bien que c'est la présence de dépendances à longue-distance dans ces différents langages, et la possibilité de générer un nombre théoriquement infini d'expressions comportant des dépendances à longue-distance, qui rendent les grammaires régulières inadéquates, et non pas le fait que l'anglais ou les langages (i-iii) comprennent un nombre potentiellement infini d'expressions. En effet, une grammaire régulière peut tout à fait générer un langage comprenant un ensemble non fini d'expressions. Un exemple en est le langage artificiel $(ab)^n$, i.e. un langage comprenant tout nombre n de paires d'éléments "ab" (cf. Figure 2).

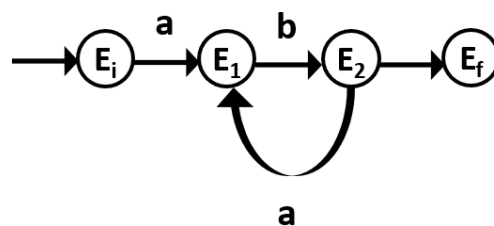


Figure 2 Exemple de grammaire régulière capable de générer un nombre infini d'expressions. Formellement, on note ce langage $(ab)^n$. E_i = état initial, E_f = état final.

Cette démonstration repose néanmoins sur un infini théorique. Or, comme le fait remarquer Chomsky lui-même, la non-adéquation des grammaires régulières pour décrire les langues naturelles pourrait être évitée en négligeant cet infini théorique : « *We might avoid this consequence by an arbitrary decree that there is a finite upper limit to sentence length in English.* » (Chomsky, 1956, p. 115). Or, les langues naturelles étant typiquement traitées par

un système cognitif limité, en termes de ressources attentionnelles, mnésiques, la définition d'une telle limite pourrait en effet être justifiée (Petersson, Folia, & Hagoort, 2012).

C'est ici qu'intervient le deuxième argument de Chomsky (1956, 1957) en faveur d'une caractérisation des langues naturelles par des grammaires supra-régulières : il serait fastidieux, et relativement inutile, de construire un diagramme à états finis à partir duquel on pourrait dériver l'ensemble des phrases grammaticales de l'anglais, ce qui reviendrait quasiment à établir une liste de ces phrases. Autrement dit, formaliser la langue anglaise par une grammaire régulière n'apporterait pas grand-chose à notre compréhension des principes la régissant, et ne permettrait pas de décrire de manière concise sa structure¹.

1.2.2 Grammaires supra-régulières

Chomsky (1956, 1957) formalise donc une hiérarchie de grammaires correspondant à des automates de puissance générative croissante, dont le niveau le plus élémentaire est occupé par les grammaires régulières. A l'extrême opposé de ce continuum, on trouve une machine de Turing, algorithme capable de générer l'ensemble des langages récursivement énumérables, c'est-à-dire l'ensemble des langages formels.

Les grammaires régulières sont trop limitées pour caractériser les langues naturelles, mais une machine de Turing l'est trop peu, d'où la formalisation de classes de grammaires supra-régulières de puissance expressive intermédiaire. Pour évaluer les langages qu'elles génèrent, ces grammaires requièrent un automate à états finis, augmenté d'une mémoire sur laquelle pèsent certaines contraintes (à la différence d'une machine de Turing). On distingue les grammaires *non contextuelles* et les grammaires *contextuelles*, chacune caractérisée par des contraintes différentes pesant sur la récupération de l'information en mémoire. Ces quatre types de grammaires sont organisées en classes hiérarchiques inclusives, comme illustré dans la Figure 3. C'est-à-dire que les langages pouvant être générés par les grammaires régulières constituent un sous-ensemble de ceux pouvant être générés par les grammaires non contextuelles, et ainsi de suite.

¹ Ceci est également valable pour la description d'autres systèmes hiérarchiques complexes, comme les langages informatiques ou la musique (e.g. Steedman, 1984).

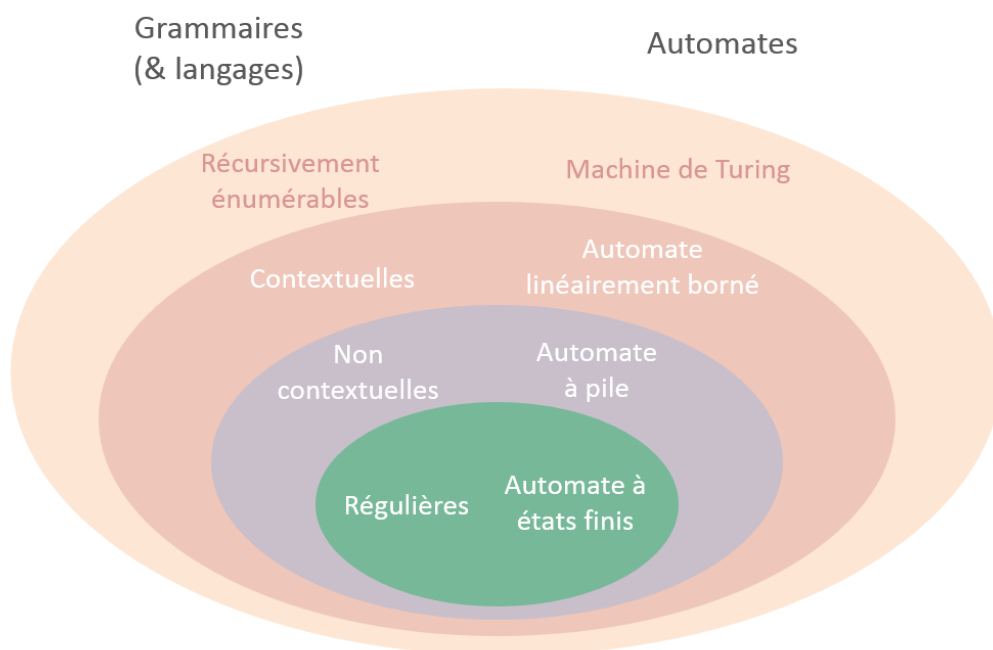


Figure 3 La hiérarchie des langages formels établie par Chomsky (1956, 1957), au sein de laquelle figurent les différents types de grammaires (et de langages) et les automates correspondants.

Les grammaires non contextuelles (*context-free grammars*) correspondent à un automate dit « à pile » (*push-down stack automaton*), qui est un automate à états finis doté d'une mémoire additive. L'automate peut seulement accéder aux items mémorisés en partant du haut de la « pile », i.e. en remontant à partir du dernier item stocké. Pour reprendre les exemples de langages artificiels donnés plus haut, une grammaire non contextuelle peut générer les langages (i) et (ii). Pour reconnaître si une expression appartient au langage (i), l'automate à pile pourra appliquer une stratégie consistant à compter puis comparer le nombre d'éléments "a" et "b". Nous nous référerons par la suite à la grammaire permettant de générer ce langage par « grammaire A^nB^n ». Pour une expression du langage (ii), il pourra vérifier que les *dépendances centrées-emboîtées* (en anglais, *center-embedded dependencies*) entre les différents éléments de la première et de la seconde moitié de l'expression sont respectées. Nous nous référerons à la grammaire correspondante dans la suite de ce manuscrit par la dénomination de « grammaire centrée-emboîtée ». Ces différentes stratégies sont illustrées dans la Figure 4.

Les grammaires contextuelles (*context-sensitive grammars*), correspondent à un automate dit « linéairement borné » (*linear-bounded automaton*), qui possède une mémoire sujette à des contraintes différentes : il peut par exemple accéder aux états passés à partir du

premier item mémorisé. La taille maximale de cette mémoire est toutefois linéairement fonction de la longueur de l'entrée sensorielle, et ne peut l'excéder. Ces grammaires peuvent générer les langages (i) et (ii), ainsi que le langage (iii). Pour reconnaître qu'une expression appartient à ce dernier, l'automate linéairement borné devra vérifier que les *dépendances croisées* (en anglais, *crossed dependencies*) entre les différents éléments de la première et de la seconde moitié de l'expression sont respectées (Figure 4c). La grammaire correspondant à ce langage sera ci-après nommée « grammaire croisée ».

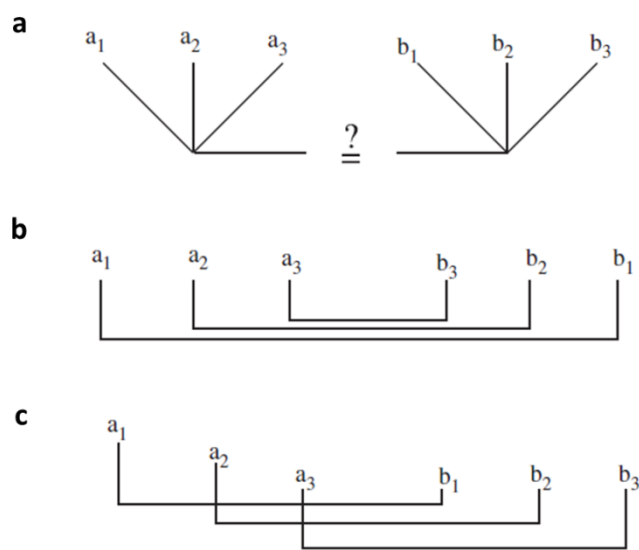


Figure 4 Illustration des différentes stratégies permettant de vérifier la grammaticalité d'expressions générées par trois grammaires supra-régulières. **a.** Pour une expression générée par la grammaire $A^n B^n$, l'automate devra compter et comparer le nombre d'éléments "a" et "b". **b.** Pour la grammaire centrée-emboîtée, il devra vérifier le caractère centré-emboîté des dépendances entre les différents éléments "a" et "b". **c.** Pour la grammaire croisée, il devra vérifier les dépendances croisées. D'après Fitch et Friederici (2012).

Où se situent les langues naturelles au sein de cette hiérarchie ? Les grammaires non contextuelles sont suffisantes pour générer des langues tels que l'anglais et le français (Chomsky, 1956). Pour d'autres langues, en revanche, des grammaires contextuelles seraient nécessaires (ou des grammaires de capacité générative intermédiaire, d'après certaines hiérarchies enrichies, telles que celle proposée par Jäger & Rogers, 2012). En suisse allemand par exemple, les dépendances entre les verbes et leurs objets sont croisées (pour un exemple, voir Jäger & Rogers, 2012). Des constructions similaires peuvent également être observées dans le néerlandais (Bresnan, Kaplan, Peters, & Zaenen, 1982) et le bambara (Culy, 1985). De

manière générale, il existe un consensus au sein de la grammaire générative, selon lequel l'ensemble des langues naturelles excéderaient les capacités génératives des grammaires régulières (e.g. Fitch & Friederici, 2012). A signaler toutefois que cette proposition n'est pas acceptée par l'ensemble de la communauté des linguistes. Il existe notamment des débats nourris autour d'une langue, le pirahã, qui serait susceptible de n'impliquer que des constructions tout à fait linéaires (Everett, 2005 ; pour un contre-argumentaire, voir e.g. Nevins, Pesetsky, & Rodrigues, 2009).

Résumé

- **L'anglais ou le français ne peuvent être générés à partir de grammaires régulières, car celles-ci sont trop restrictives pour pouvoir générer l'ensemble des phrases de ces langues.**
- **Ces langues (et, par extension, l'ensemble des langues naturelles) requerraient par conséquent des capacités génératives supra-régulières.**

I.3 Organisation des langues en syntagmes récursivement enchâssés

I.3.1 Une organisation hiérarchique

La grammaire générative décrit l'organisation des langues naturelles sous la forme de syntagmes hiérarchiquement enchâssés. En anglais, on parle de *constituent* ou *phrase structure grammars* (Chomsky, 1956, 1957)². Ces structures peuvent être formalisées par des arbres syntaxiques, comme l'exemple donné dans la Figure 5 où chaque nœud correspond à un objet syntaxique (syntagme nominal, syntagme verbal, etc.). Ces objets sont organisés en différents niveaux hiérarchiques (dans la Figure 5, le premier syntagme nominal est lui-même décomposable en un syntagme nominal et une proposition relative, etc.), dont les éléments terminaux sont les mots. Le locuteur formerait des représentations elles-mêmes hiérarchiques, et non linéaires, des phrases qu'il entend ou produit (e.g. Everaert, Huybregts, Chomsky, Berwick, & Bolhuis, 2015). Postuler que les représentations internes des expressions

² Un autre type de grammaires, dites grammaires transformationnelles (en anglais, *transformational grammars* ; Chomsky, 1956), a également été formalisé. Pour simplifier notre propos, nous ne nous intéressons pas ici à cette distinction.

linguistiques ne sont pas linéaires permet d'expliquer divers phénomènes linguistiques, notamment la présence des relations définies par la structure hiérarchique sur la proximité temporelle entre les éléments. Pour notre exemple, cela permet d'expliquer que l'on comprenne à la lecture de la phrase « Le babouin qui dort sur le rocher est brun. » (Figure 5), que c'est le babouin -et non le rocher- qui est brun, malgré que ce dernier nom soit plus proche.

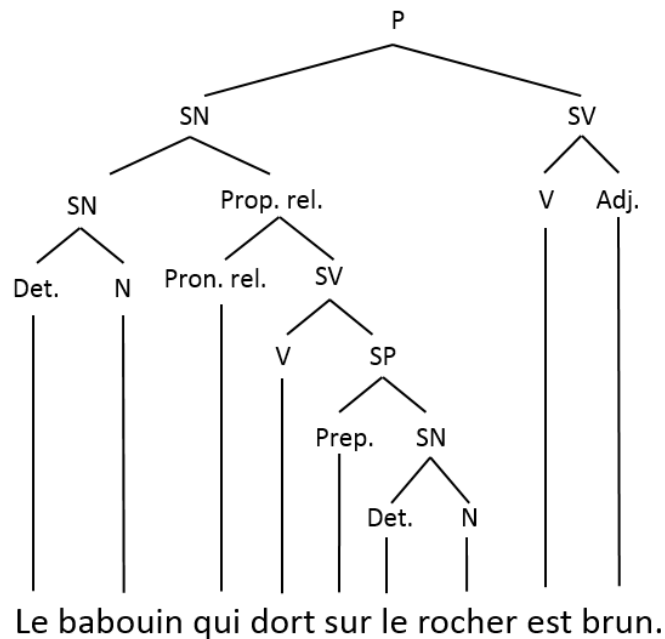


Figure 5 Arbre syntaxique représentant la structure d'une phrase comportant une proposition relative. P = phrase, SN = syntagme nominal, SV = syntagme verbal, Det. = déterminant, N = nom, Prop. rel. = proposition relative, Pron. rel. = pronom relatif, SP = syntagme prépositionnel, Prep. = préposition, V = verbe, Adj. = adjectif. Modifié d'après Hoch (2010).

1.3.2 Un système récursif

C'est l'enchâssement *récursif* de ces syntagmes qui permettrait la productivité de la faculté de langage (e.g. Hauser et al., 2002). Le sens donné à ce terme varie selon les différentes disciplines en faisant l'usage, comme les mathématiques, les sciences computationnelles ou la linguistique (Fitch, 2010), et parfois même au sein de chacune de ces disciplines (Tomalin, 2007). Nous nous limiterons ici aux deux définitions les plus pertinentes au regard de l'hypothèse de Hauser et al. (2002).

Chomsky (e.g., Everaert, Huybregts, Chomsky, Berwick, & Bolhuis, 2015; voir également Lobina, 2011, et Watumull, Hauser, Roberts, & Hornstein, 2014) se focalise sur la

notion de récursivité telle qu'elle est définie en mathématiques et logique, et qui renvoie à la possibilité de réappliquer une fonction à un résultat (*output*) issu d'une application antérieure de cette même fonction. Cette fonction peut également être dite définie par induction, ou encore « auto-référentielle ». Dans les développements les plus récents de la théorie générativiste (le « programme minimaliste » ; Chomsky, 1995), la construction d'objets syntaxiques a été réduite à l'application récursive d'une opération minimale, appelée « merge » (que l'on pourrait traduire en français par « combinaison », ou « fusion »). Cette opération combinatoire consiste, à partir de deux objets syntaxiques, X et Y (e.g. « le » et « babouin ») de construire un nouvel objet syntaxique, Z (ici un syntagme nominal, « le babouin »). Formellement, cette opération peut se noter : Merge (le, babouin) → {le, babouin}. L'application récursive de cette opération (en fusionnant par exemple Z, qui est ici un syntagme nominal, avec un syntagme verbal – résultant lui-même d'une opération de fusion– donnant par exemple la phrase « le babouin dort sur le rocher ») prend fin lorsque l'ensemble des objets syntaxiques ont été fusionnés en un seul objet (i.e. la phrase).

Certains auteurs ont préféré se référer à une définition plus stricte de la récursivité, entendue alors au sens d'auto-enchâssement (*self-embedding*) de constituants syntaxiques au sein de constituants syntaxiques *de même type* (e.g. Fitch, 2010; Kinsella, 2010; Martins, 2012), comme par exemple l'enchâssement d'un syntagme verbal au sein d'un autre syntagme verbal (noté : $SV_1 \rightarrow SV_2 + X$). Ce choix peut s'expliquer par le fait que c'est effectivement seulement dans le cas de l'auto-enchâssement qu'il n'est pas nécessaire de spécifier de nouvelles règles à chaque itération de l'opération (et que l'on peut donc obtenir des enchâssements de profondeur arbitraire ; e.g. « Je pense que ... », « Je pense que tu penses que ... », et ainsi de suite). On peut toutefois souligner que ces deux définitions ne comportent pas de réelle contradiction, car les opérations récursives telles que définies au sens strict constituent un sous-ensemble de celles définies au sens large (Hulst, 2010).

I.4 Théories générativistes sur l'évolution de la syntaxe

Hauser, Chomsky, et Fitch (2002) ont proposé de rassembler les différentes composantes du langage en deux facultés distinctes : la faculté de langage au sens large (*language faculty in the broad sense*, ou FLB) et la faculté de langage au sens étroit (*language faculty in the narrow sense*, ou FLN). Dans la FLB, on retrouverait au minimum trois systèmes :

(1) un système sensori-moteur, correspondant aux mécanismes de perception et de production de signaux, (2) un système conceptuel-intentionnel, pour les aspects sémantiques, et (3) un système computationnel récursif, ou syntaxique au sens strict. Comme l'illustre la Figure 6, ce dernier système est le seul que l'on retrouve dans la FLN. Il interagirait avec les autres systèmes constituant la FLB, mais serait indépendant de ceux-ci. La propriété fondamentale de la FLN, sa récursivité, permettrait sa productivité, c'est-à-dire sa capacité à générer une quantité infinie d'expressions (phrases), à partir d'un nombre fini d'unités discrètes (i.e. de faire un usage infini de moyens finis, Humboldt, 1836).

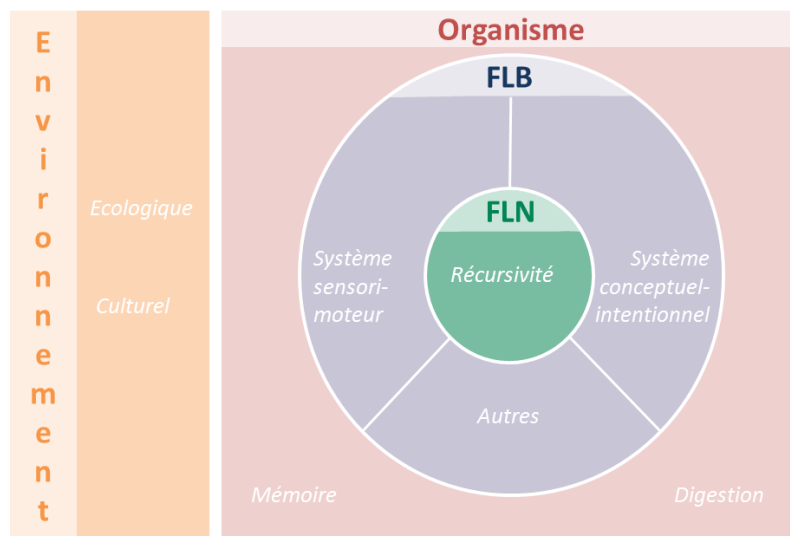


Figure 6 Représentation schématique des différents systèmes impliqués dans la faculté de langage au sens large (FLB) et au sens étroit (FLN), et exemples de facteurs internes et externes avec lesquels ces systèmes interagissent. Modifié d'après Hauser et al. (2002).

Du point de vue évolutif, ces auteurs proposent qu'aucune des capacités impliquées dans la FLB ne soit strictement spécifique à l'espèce humaine, et qu'elles puissent se retrouver sous diverses formes, potentiellement moins complexes, chez d'autres espèces, à l'exception de la FLN (voir également Fitch, Hauser, & Chomsky, 2005). Autrement dit, ils proposent une vision continuiste de la FLB, et font l'hypothèse qu'il existe une discontinuité qualitative concernant la FLN : « *On the basis of data reviewed below, we hypothesize that most, if not all, of FLB is based on mechanisms shared with nonhuman animals [...]. In contrast, we suggest that FLN - the computational mechanism of recursion- is recently evolved and unique to our species.* » (Hauser et al., 2002, p. 1573).

Cette proposition initiale sera par la suite enrichie. Fitch (2014) reprend une idée de Miller (1967), selon laquelle le système cognitif humain ne serait pas seulement capable de

former des représentations hiérarchiques, mais manifesterait également une *propension* à inférer ce type de structure à partir de données séquentielles. Fitch (2014) étend cette hypothèse à un cadre comparatif, en proposant qu'une telle préférence pour ce type de représentation puisse être exclusivement observée chez l'espèce humaine. Il l'appellera l'hypothèse de la « *dendrophilie* », en référence au caractère hiérarchique, ou arborescent, de ces représentations. Enfin, cette faculté générative récursive ne se retrouverait pas seulement dans le langage, mais aussi dans un vaste ensemble de domaines de la cognition humaine (tels que la musique, la numération, la géométrie, ou encore la théorie de l'esprit ; Fitch, 2014; Hauser & Watumull, 2017³. Cf Figure 7).

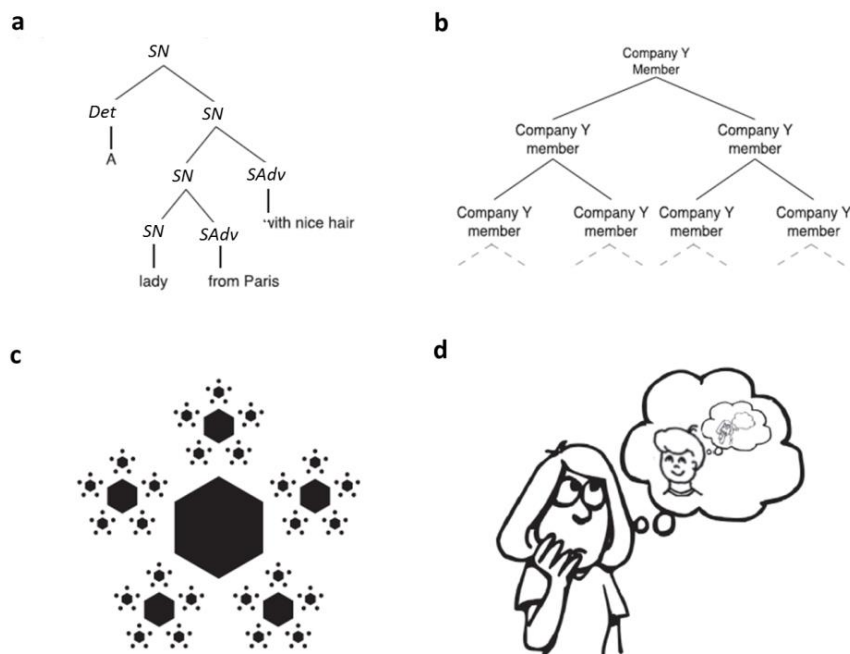


Figure 7 Exemples de domaines dans lesquelles des computations récursives pourraient être appliquées. Enchâssement récursif (a) d'objets syntaxiques, (b) de statuts hiérarchiques, (c) de figures géométriques, (d) d'états mentaux. *SN* = syntagme nominal, *Det* = déterminant, *SAdv* = syntagme adverbial. Modifié d'après Fischmeister, Martins, Beisteiner, & Fitch (2017).

Parallèlement, Fitch et Hauser (2004), sur la base de travaux antérieurs de Chomsky (1956, 1957 ; voir section I.2), soulignent que ce type de computation dépasse la puissance

³ Cette idée apparaissait déjà dans le papier de Hauser, Chomsky et Fitch (2002), qui proposaient que la spécificité de la cognition humaine ne réside peut-être pas dans sa capacité à effectuer des computations récursives (qui pourraient potentiellement exister chez d'autres espèces pour des fonctions autres que la communication, telle que la navigation spatiale), mais dans l'application de ces computations à divers domaines.

expressive des grammaires dites régulières. Ils relèvent que l'existence ou non de capacités génératives supra-régulières chez d'autres espèces nous est inconnue, et avancent une hypothèse complémentaire à la précédente : les capacités computationnelles des animaux non humains seraient restreintes au traitement des grammaires régulières.

Résumé

- **Trois hypothèses complémentaires ont été formulées pour caractériser la spécificité des traitements syntaxiques effectués par l'humain.**
- **Hypothèse de la récursivité⁴ : la faculté de langage implique un système computationnel récursif apparu récemment au cours de l'évolution de la lignée humaine, permettant l'enchâssement récursif de syntagmes dans des structures hiérarchiques de profondeur arbitraire (Hauser, Chomsky, & Fitch 2002).**
- **Hypothèse de la dendrophilie : l'humain manifeste une propension particulière à inférer une structure hiérarchique à partir de données séquentielles (Fitch, 2014).**
- **Hypothèse de la supra-régularité : la capacité à traiter des grammaires supra-régulières est restreinte à l'espèce humaine (Fitch & Hauser, 2004).**

I.5 Confusions à éviter : supra-régularité, arborescence, récursivité

La mise à l'épreuve des hypothèses présentées dans la section I.4 a jusqu'ici majoritairement été effectuée en examinant la capacité de diverses espèces à traiter différents langages artificiels : des langages générés par les trois grammaires supra-régulières vues plus haut d'une part, et le langage $(ab)^n$ d'autre part. Ces langages ont été choisis pour certaines de leurs propriétés (e.g. O'Donnell, Hauser, & Fitch, 2005). Mais du fait d'un manque de caractérisation précise des différentes notions que nous venons d'examiner (Fitch & Hauser, 2004; Hauser et al., 2002), des confusions ont été faites dans les interprétations de ces travaux expérimentaux (e.g. Gentner, Fenn, Margoliash, & Nusbaum, 2006; Rey, Perruchet, & Fagot,

⁴ A l'exception de l'hypothèse de la dendrophilie, ces dénominations nous sont propres, et ont pour seul but de permettre d'identifier ces hypothèses plus aisément dans ce manuscrit.

2012). Nous tenterons donc ici de clarifier ce que ces différents langages permettent et ne permettent pas de tester, avant de rapporter les résultats de ces expériences.

1.5.1 Supra-régularité et enchâssement récursif

L'évaluation de la capacité d'une espèce à traiter des grammaires supra-régulières peut s'effectuer en examinant sa capacité à traiter l'une des trois grammaires suivantes : A^nB^n , centrée-emboîtée, et croisée. Cette capacité peut éventuellement être contrastée avec celle de traiter le langage $(ab)^n$, qui constitue un exemple paradigmatique de langage régulier (cf. section I.2.1 ; voir également Fitch & Hauser, 2004; O'Donnell, Hauser, & Fitch, 2005). Il est attendu qu'une espèce dotée de capacités computationnelles limitées aux langages réguliers échoue à traiter une grammaire supra-régulière, tandis qu'il pourra traiter un langage généré à partir de la grammaire $(ab)^n$.

En revanche, la comparaison de ces langages ne permet pas de vérifier la capacité d'un système à former des représentations récursivement enchâssées, pour deux raisons. Premièrement, ces deux langages peuvent être générés par une procédure récursive, comme par une procédure non récursive (Fitch, 2010). Par exemple, les deux fonctions récursives suivantes peuvent générer les langages $a^n b^n$ et $(ab)^n$, respectivement :

$S \rightarrow a + S + b$

$S \rightarrow ab + S$

Deuxièmement, l'adéquation du langage $a^n b^n$ pour l'étude du traitement de structures hiérarchiques est sujet à discussion. Une grande majorité des expériences employant ce langage l'associe aux constructions syntaxiques hiérarchiques des langues naturelles (e.g. Bahlmann, Schubotz, & Friederici, 2008; de Vries, Monaghan, Knecht, & Zwitserlood, 2008; Fitch, 2014; Fitch & Hauser, 2004; Lai & Poletiek, 2011), et l'arbre syntaxique classiquement employé pour le représenter présente les paires les plus externes comme étant situées à un niveau hiérarchique supérieur aux paires les plus internes (cf. Figure 8a). Le langage $a^n b^n$ est souvent employé dans ces études comme un équivalent artificiel d'une phrase comportant une ou plusieurs propositions relatives organisées de manière centrée-emboîtée, telle que « Le babouin qui dort sur le rocher est brun. » qui, selon les théories génératives, sera représentée par un locuteur du français sous la forme d'une proposition principale dans laquelle une proposition relative aura été enchâssée (cf. Figure 5, p. 23).

Mais, comme le soulignent Ojima et Okanoya (2014), si de telles relations de dominance/subordination ne sont pas rendues explicites lors de la familiarisation des sujets à un langage artificiel généré à partir des grammaires A^nB^n , centrée-emboîtée ou croisée, ceux-ci ne pourront pas se représenter les relations entre les paires comme telles. Par conséquent, ces auteurs proposent que les sujets de ces expériences soient davantage susceptibles de former des représentations linéaires des expressions de ce langage, dans lesquelles seules les relations de dépendance entre les éléments sont encodées (cf. Figure 8b). Parler de structure « centrée-emboîtée » pour ce langage a pu être à l'origine de cette confusion, mais cette dénomination ne fait que référer au fait qu'il corresponde à une insertion de paires "AB" au milieu d'autres paires "A_B". Cette procédure n'implique pas nécessairement que ces paires soient enchâssées les unes au sein des autres, et donnent lieu à une structure hiérarchique (Fitch, 2010).

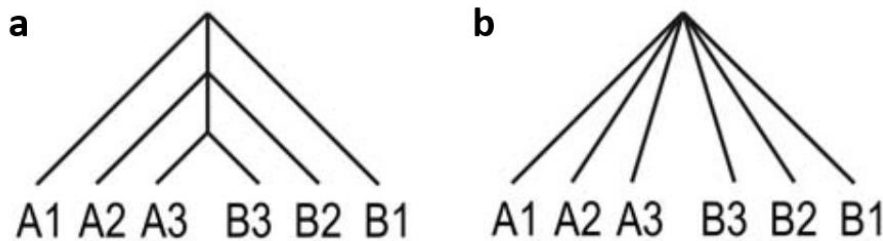


Figure 8 a. Arbre syntaxique classiquement utilisé pour représenter une grammaire $a^n b^n$ (ici $n = 3$). Selon ce diagramme, la représentation de la séquence est organisée hiérarchiquement. b. Arbre syntaxique pouvant également représenter une grammaire $a^n b^n$. Noter que cette représentation alternative n'implique pas d'organisation hiérarchique de la séquence, mais seulement des dépendances entre les différents éléments (les éléments partageant un même index (1-3) étant appariés). D'après Ojima & Okanoya (2014).

De plus, le contraste entre les langages $a^n b^n$ et $(ab)^n$ n'est pas nécessairement pertinent (Ojima & Okanoya, 2014). En effet, ces deux langages peuvent tout autant donner lieu à des représentations linéaires ou hiérarchiques, en fonction des stimuli employés pour les implémenter. Par exemple, des expressions de la forme $(ab)^n$ peuvent être trouvées dans les langues naturelles, sous la forme de grammaires branchées à droite (e.g. « C'est ce qu'a dit la mère de la tante du cousin de Marie (...). »).

Mais il existe également des arguments en faveur de l'existence de mécanismes communs pour le traitement des grammaires artificielles $A^n B^n$, centrée-emboîtée ou croisée, et de structures linguistiques. Le traitement de phrases organisées en syntagmes implique le maintien temporaire en mémoire de résultats intermédiaires du traitement (e.g. encodage des

différents mots composant le syntagme nominal, stockage temporaire du résultat de ce traitement, encodage du syntagme verbal, et intégration de l'ensemble ; voir Fitch, 2014). On peut envisager que le même type de mémoire soit impliqué lors du traitement d'un langage supra-régulier construite à partir de stimuli artificiels. Par exemple, traiter une expression générée à partir de la grammaire $A^n B^n$ nécessite un stockage en mémoire du nombre de "a" ayant été comptés, afin de pouvoir le comparer ultérieurement avec le nombre de "b". Le traitement du langage $(ab)^n$, à l'inverse, qui ne présente que des régularités locales, ne nécessiterait pas, ou dans une moindre mesure, le maintien d'informations en mémoire. Nous verrons dans la section II.2.4 que les résultats de plusieurs expériences de neuro-imagerie suggèrent que le traitement de grammaires artificielles supra-régulières, et le traitement d'expressions linguistiques complexes, présentent effectivement certaines similarités.

1.5.2 Récurtivité et organisation hiérarchique

L'enchâssement récursif est à distinguer de la décomposition d'un ensemble hiérarchiquement organisé en composants élémentaires. Par exemple, Pinker et Jackendoff (2005), dans une réponse à l'article de Hauser et al. (2002), avaient donné un exemple de « récursivité » dans le domaine de la perception visuelle, illustré par la Figure 9. Cette figure complexe peut être décomposée en constituants élémentaires : quatre matrices rectangulaires, chacune étant elle-même décomposable en quatre matrices, etc., jusqu'à arriver aux éléments minimaux "x".

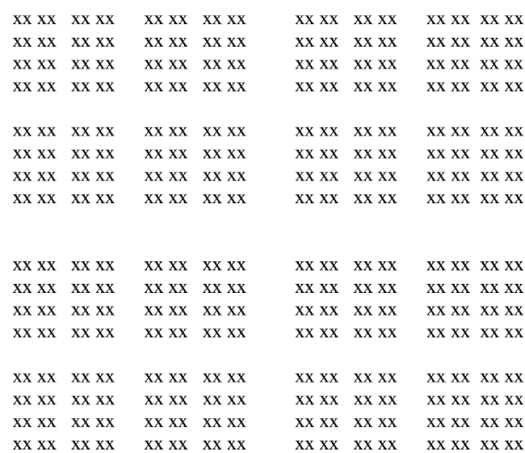


Figure 9 Exemple de structure hiérarchique visuelle, d'après Pinker & Jackendoff (2005).

Traiter une structure organisée de manière hiérarchique telle que celle-ci requière bien la capacité de se représenter des relations de subordination entre un certain nombre d'entités, occupant un certain nombre de niveaux. Toutefois, seuls les niveaux hiérarchiques « connus » peuvent être représentés : aucune inférence ne pourra être faite quant aux règles à appliquer pour générer de nouveaux niveaux. Car ce n'est pas une expression linguistique (ou une structure) qui est récursive, mais la procédure générative qui en est à l'origine. Du fait de sa récursivité, une telle procédure permet de générer ou représenter de nouveaux niveaux hiérarchiques, sans que de nouvelles règles aient à être spécifiées.

Un test valable de la capacité d'un système à effectuer ce type de computations devra donc vérifier que ce système est capable de générer (ou de reconnaître) de nouveaux niveaux hiérarchiques, au-delà de ceux disponibles (Martins, 2012). L'emploi d'un matériel expérimental tout à fait différent des langages artificiels que nous avons vu jusqu'à présent, comme des figures géométriques fractales (cf. Figure 7c et Figure 10), pourrait être plus adéquat pour évaluer la capacité d'un système cognitif à effectuer des computations récursives (e.g. Fischmeister et al., 2017).

Résumé

- **Les langages $a^n b^n$ sont impropres pour tester la capacité d'un système à former des représentations récursivement enchâssées.**
- **Ils constituent en revanche un matériel idéal pour tester la capacité d'un système à traiter des grammaires supra-régulières, en le contrastant par exemple avec le langage $(ab)^n$, qui est régulier.**
- **La capacité d'un système cognitif à former des représentations hiérarchiques de manière récursivement enchâssées peut s'effectuer en employant d'autres types de matériel expérimental (Figure 10). Un test valable de la capacité d'un système cognitif à former ce type de représentations devra examiner si ce système peut générer de nouveaux niveaux hiérarchiques.**



Figure 10 La récurtivité et la supra-régularité sont deux propriétés caractérisant la faculté de langage (au centre). L'évaluation de la capacité d'un système à effectuer des computations récurtives, et celle de sa puissance générative peuvent toutefois requérir d'employer des matériels expérimentaux tout à fait différents.

II. Données Comparatives

On distingue classiquement deux approches pour l'étude des capacités syntaxiques des animaux NH : l'étude des productions naturelles, et l'étude des capacités d'apprentissage en laboratoire (e.g. Petkov & Wilson, 2012).

II.1 Production

II.1.1 Communication

Nous ne traitons ici que de travaux portant sur des aspects strictement syntaxiques des productions vocales, i.e. sur la description des règles gouvernant l'organisation des éléments dans des séquences de vocalisations. Nous ne rapportons pas les études et discussions autour des questions de sémantique combinatoire (pour des travaux récents sur ces questions, voir e.g. Schlenker, Chemla, & Zuberbühler, 2016; Suzuki, Wheatcroft, & Griesser, 2016). Nous n'aborderons pas non l'étude des productions induites expérimentalement dans le cadre des recherches sur l'apprentissage de langues chez l'animal (*Animal Language Research*)⁵.

⁵ Ces travaux consistent à apprendre à des animaux la langue des signes (e.g. Gardner & Gardner, 1969) ou à utiliser un tableau de symboles visuels (« lexigramme » ; e.g. Savage-Rumbaugh & Lewin, 1996).

II.1.1.1 Chez les espèces non-primates

Les chants de certaines espèces d'oiseaux chanteurs peuvent être considérés comme étant organisés de manière hiérarchique, au sens où ils sont composés d'unités discrètes (ten Cate, Lachlan, & Zuidema, 2013), pouvant être regroupées entre elles en unités plus larges (e.g. une série de répétitions d'une même syllabe) appelées « motif », ou *phrase* (ten Cate & Okanoya, 2012), et organisées selon une certaine structure. Cette structure peut être décrite par des grammaires régulières comprenant un nombre fini d'états, et des probabilités transitionnelles entre ces états (i.e. les règles de passage d'un état à un autre ne sont pas déterministes, mais probabilistes ; Berwick, Okanoya, Beckers, & Bolhuis, 2011). Les analyses formelles appliquées à ces chants ont toutefois identifié des variations de complexité entre différentes espèces.

Des chaînes de Markov de premier ordre (ne prenant en compte que le dernier élément de la séquence pour prédire le suivant ; i.e. ne vérifiant que des bigrammes) semblent par exemple suffire pour modéliser la plupart des chants de l'étourneau sansonnet (*Sturnus vulgaris*, Gentner & Hulse, 1998) et de différentes espèces de grives américaines (Dobson & Lemon, 1979). Les chants du moineau du japon (*Lonchura striata* var. *domestica*) en revanche, sont typiquement organisés en fragments (en anglais, *chunks*), chacun comprenant une à cinq notes (Suge & Okanoya, 2010). Les grammaires régulières décrivant la structure de ces chants comportent des boucles, qui ont pour effet qu'un nombre arbitraire de répétitions d'un même fragment puisse apparaître entre deux autres fragments (Hosino & Okanoya, 2000). Pour décrire ces chants, il est donc nécessaire d'utiliser des chaînes de Markov « *n*-gram » ayant une fenêtre de vérification de taille supérieure (Kakishita, Sasahara, Nishino, Takahasi, & Okanoya, 2009)⁶. De la même manière, les chants du rossignol philomèle (*Luscinia megarhynchos*) présentent des boucles permettant l'insertion d'un nombre arbitraire de fragments entre d'autres fragments (Todt & Hultsch, 1998). Les chants de ces espèces ne sont toutefois pas organisés sous la forme d'enchâssements de *profondeur* arbitraire, i.e. organisés sur plusieurs niveaux hiérarchiques (Berwick et al., 2011).

L'étude de l'organisation syntaxique des expressions produites par ces individus suggère qu'elles sont d'un niveau de complexité limité (Terrace, Petitto, Sanders, & Bever, 1979; Truswell, 2017; Yang, 2013).

⁶ Bien que l'immense majorité des travaux de formalisation de ces productions vocales aient employé des chaînes de Markov de diverses sortes, il est à noter que d'autres types de modèles pourraient s'avérer plus appropriés (e.g. voir Kershenbaum et al., 2014).

Des chants complexes composés d'unités élémentaires ont également été rapportés chez des mammifères marins comme la baleine à bosses (*Megaptera novaeangliae* (Green, Mercado, Pack, & Herman, 2011; Suzuki, Buck, & Tyack, 2006). La formalisation de ces chants requerrait également des chaînes de Markov d'ordre supérieur, car ils impliquent des dépendances à longue distance ; mais ici encore des grammaires régulières suffisent (pour une discussion, voir Kinsella, 2010). Enfin, bien que la grande majorité des travaux sur la question se soient concentrés sur les oiseaux chanteurs et les mammifères marins, on peut toutefois noter que d'autres espèces présentent des séquences de vocalisations comportant des unités élémentaires regroupées en fragments plus larges (e.g. chez la tourterelle, *Streptopelia decaocto*, ten Cate & Ballintijn, 1996 ; ou la souris, *Mus domesticus*, Holy & Guo, 2005).

II.1.1.2 Chez les primates non-humains

L'étude de l'organisation séquentielle des signaux de communication n'est apparue que très récemment dans le champ de la primatologie. Mais ces travaux, fruits de collaborations entre primatologues et linguistes, offrent une approche rigoureuse pour l'étude des propriétés formelles des systèmes de communication des primates NH⁷ (Schlenker, Chemla, & Zuberbühler, 2016).

Chez certaines espèces de singes, différentes vocalisations peuvent être émises au cours d'une même séquence de production, dans un ordre non aléatoire. Chez le mone de Campbell, la vocalisation *boom* est typiquement émise par paire et en début de séquence (*Cercopithecus campbelli* ; Zuberbühler, 2002). Chez le colobe, deux vocalisations, *snort* et *roar*, sont toujours émises dans le même ordre lorsqu'elles sont présentes dans une même séquence, un seul *snort* étant suivi par plusieurs *roar* (*Colobus polykomos* et *Colobus guereza* ; Schel, Tranquilli, & Zuberbühler, 2009). Chez une espèce de cercopithèque (putty-nosed monkeys, *Cercopithecus nictitans*), les mâles émettent des séquences consistant en une concaténation de plusieurs *pyows* suivis de plusieurs *hacks*. Des régularités sont également rapportées dans les séquences de vocalisations du titi, un singe du nouveau monde (*Callicebus nigrifrons* ; Cäsar, Zuberbühler, Young, & Byrne, 2013).

L'interprétation selon laquelle l'ordre d'émission de ces vocalisations serait gouverné par des règles n'en est toutefois qu'une parmi d'autres possibles, et plus ou moins pertinente

⁷ A notre connaissance, ces travaux n'ont jusqu'à présent porté que sur les singes non-anthropoïdes ; à l'exception du gibbon (voir Clarke, Reichard, & Zuberbühler, 2006).

selon les cas (pour une discussion complète, voir Schlenker, Chemla, Schel, et al., 2016a, 2016b). Par exemple, les singes *putty-nosed* pourraient se représenter les séquences constituées d'une série de *pyows* suivie d'une série de *hacks* sous la forme d'une règle (e.g. *pyow+ hack+*). Mais ils pourraient également avoir mémorisé une série d'exemplaires, chacun correspondant à une longueur possible de séquence (e.g. *pyow hack*, *pyow pyow hack*, etc). Le nombre relativement important de séquences différentes de ce type ayant été enregistrées par les chercheurs constitue un argument en faveur du premier type d'encodage, mais des études supplémentaires seraient nécessaires pour répondre à cette question. Par ailleurs, le rôle joué par différents types de contraintes sur cette organisation reste à déterminer. Dans le cas du mone de Campbell, par exemple, les caractéristiques articulatoires du *boom*, qui requièrent que les sacs d'air de l'individu émetteur soient pleins, pourraient expliquer que cette vocalisation n'apparaisse qu'en tout début de séquence.

Deux séries de travaux suggèrent de leur côté la présence d'affixes dans les systèmes de communication de deux espèces de singes, ce qui pourrait constituer des cas de régularités morphologiques (Schlenker, Chemla, Schel, et al., 2016a, 2016b). Chez le mone de Campbell toujours, un suffixe *-oo* peut être observé à la suite de deux vocalisations différentes, *krack* et *hok* (Ouattara, Lemasson, & Zuberbühler, 2009; Schlenker et al., 2014). Chez le cercopithèque diane, une vocalisation particulière peut se trouver à la suite de différentes autres vocalisations (*Cercopithecus diana*, Candiotti, Zuberbühler, & Lemasson, 2012; Coye, Zuberbühler, & Lemasson, 2016).

Cet ensemble de données suggère que l'organisation séquentielle des vocalisations n'atteint que des niveaux de complexité assez limités chez les primates non-humains. Mais le nombre d'espèces étudiées reste encore très restreint, et la conduite d'études ne se focalisant plus seulement sur les vocalisations, mais également sur les gestes de communication offre une piste complémentaire prometteuse (Byrne et al., 2017).

II.1.2 Fonctions non-communicatives

En dehors des systèmes de communication, l'existence de représentations hiérarchiques récursivement enchâssées, et nécessitant des capacités computationnelles supra-régulières, a été postulée dans plusieurs autres domaines de la cognition humaine : dans la cognition spatiale, la cognition sociale, la numération et la programmation de séquences d'actions complexes (e.g. Hauser & Watumull, 2017). Pour illustrer l'état d'avancement des recherches comparatives

dans ces domaines, nous développerons deux aspects : la cognition sociale et les séquences d'actions complexes.

II.1.2.1 Cognition sociale

Les travaux des primatologues D. Cheney et R. Seyfarth sur le babouin hamadryas (*Papio hamadryas ursinus*) indiquent que la structure sociale de cette espèce respecte une hiérarchie de dominance organisée en deux niveaux : un premier niveau correspondant à une hiérarchie entre les familles (avec une transmission matrilineaire du rang social), et un second niveau de hiérarchie entre les individus au sein de ces familles (cf. Figure 11; Cheney & Seyfarth, 2007). Ces auteurs ont effectué des expériences de play-back afin de déterminer si les babouins se représentent ces deux niveaux de hiérarchie (Bergman, Beehner, Cheney, & Seyfarth, 2003). Pour réaliser ces expériences, ils ont utilisé deux types de vocalisations : une vocalisation typiquement émise par les femelles hamadryas pour menacer un individu subordonné (*threat-grunt*), et une vocalisation typiquement émise par un individu menacé par une femelle de rang supérieur (*scream*). Une procédure de violation des attentes était employée : des séquences de vocalisations violant les rangs hiérarchiques étaient diffusées (e.g. *threat-grunt* de l'individu de rang 2 suivi d'un *scream* de l'individu de rang 1, cf Figure 11), et les réponses d'orientation des babouins en direction du haut-parleur mesurées.

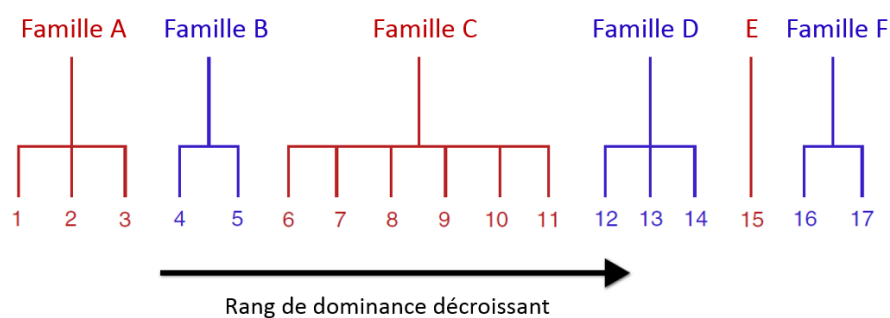


Figure 11 Organisation hiérarchique typique d'un groupe de babouins hamadryas. Les familles (lettres), et les individus (nombres) entretiennent des relations de dominance linéaire. Modifié d'après Seyfarth & Cheney (2014).

Les babouins ont davantage réagi aux séquences présentant une violation de rang entre deux familles, qu'à celles avec une violation de rang au sein d'une même famille, et ce indépendamment de l'écart de rang entre les individus. Ainsi, dans l'exemple de la Figure 11,

les babouins auront plus tendance à réagir si dans la séquence diffusée l'individu 4 menace l'individu 3 (inter-famille), que si l'individu 3 menace l'individu 2 (intra-famille).

Ces résultats suggèrent que les babouins hamadryas se représentent les rangs de dominance au sein de leur groupe sous la forme d'une hiérarchie à deux niveaux. Toutefois, ils ne nous permettent pas de tirer de conclusions définitives quant au caractère récursif de ce système de représentation. Par exemple, l'encodage des rangs familiaux pourrait être représenté de trois manières différentes : (1) par une série de représentations des liens de subordinations entre chaque famille deux-à-deux (Famille A > Famille B ; Famille B > Famille C ; etc), (2) sous la forme d'une hiérarchie de dominance enchâssée (Famille A > Famille B > Famille C > ...), ou (3) une hiérarchie de dominance *récursivement* enchâssée (ce dernier type de représentation aurait alors pour propriété de pouvoir inclure de nouvelles familles comme autant de nouveaux niveaux dans le système). Pour ce qui est de la distinction entre les deux premières hypothèses, il n'est pas rare que les groupes de babouins hamadryas comportent au moins dix familles (Cheney & Seyfarth, 2007), ce qui rend le nombre de comparaisons deux-à-deux à mémoriser assez important (e.g. pour un cas où la troupe comporte dix familles, il serait nécessaire de mémoriser 45 combinaisons). L'adoption d'une représentation transitive du rang hiérarchique des différentes familles (hypothèse 2) pourraient donc constituer un mode d'encodage moins coûteux en termes de ressources mnésiques. Mais cette hypothèse n'a pas été vérifiée expérimentalement. La distinction entre les hypothèses (2) et (3) est plus difficile encore, les familles constituant un groupe, et la hiérarchie entre ces familles étant extrêmement stable chez cette espèce (Cheney & Seyfarth, 2007).

II.1.2.2 *Actions complexes*

Byrne et Russon (1998 ; voir également Byrne, Corp, & Byrne, 2001) ont étudié les séquences d'actions impliquées dans les comportements alimentaires du gorille des montagnes (*Gorilla g. beringei*), et dans l'imitation d'actions complexes effectuées par des humains, par des orangs-outans (*Pongo pygmaeus*). Les auteurs proposent que ces phénomènes relèvent d'une organisation hiérarchique, plutôt que linéaire. Leur argumentation repose sur l'observation que ces séquences sont organisées en buts (e.g. chez le gorille, retirer les parties non-comestibles de la plante) pouvant être décomposés en sous-buts (e.g. détacher le pétiole, arracher les épines). Les auteurs proposent ainsi que ces séquences comportementales puissent être décrites en utilisant le même type de règles de réécriture employées en linguistique pour

décomposer une expression linguistique en ses différents composants, et le même type de diagramme (Figure 12).

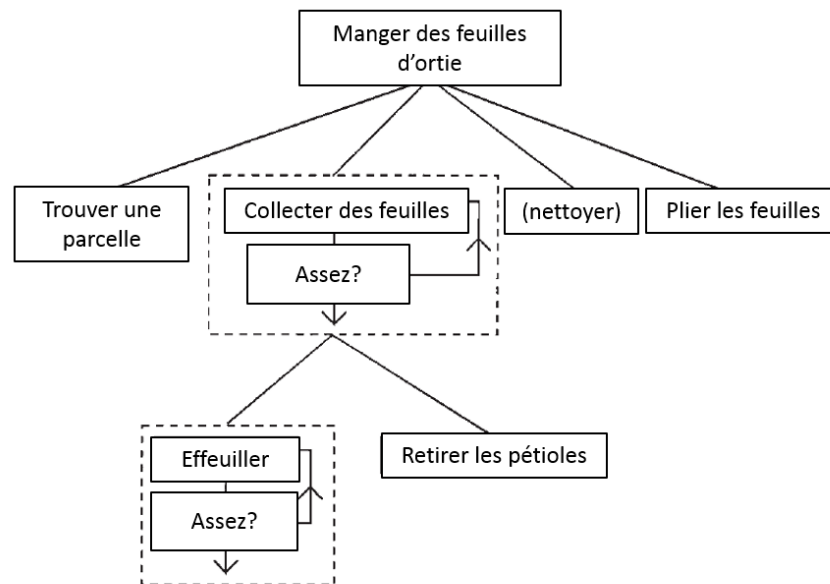


Figure 12 Diagramme représentant la hiérarchie des buts et sous-buts impliqués dans les séquences comportementales visant à la consommation de feuilles d'orties chez le gorille. Modifié d'après Byrne & Russon (1998).

Ces formalisations permettent de rendre compte de l'ordre dans lequel se déroulent les différentes opérations. Elles permettent également de rendre compte du fait que certaines opérations sont optionnelles (nettoyer), et que d'autres peuvent être répétées de manière itérative (e.g. effeuiller). Corp et Byrne (2002) décrivent également chez le chimpanzé (*Pan troglodytes*) des séquences comportementales flexibles pour la préparation de deux espèces de plantes dont les feuilles comportent des parties difficiles à ingérer. Ces séquences sont ici aussi décrites comme étant composées de différents modules correspondant à différents sous-buts, certains de ces modules pouvant être optionnels ou répétés. De plus, les auteurs rapportent dans ce cas l'existence de différents modules alternatifs pour l'atteinte d'un même sous-but (e.g. plusieurs techniques sont observées pour le détachement des pétioles, chez chacun des individus).

Ces séquences de préparation d'aliments avant leur ingestion peuvent donc être décrites de manière hiérarchique. L'existence d'une certaine flexibilité dans ces séquences comportementales rend plausible l'interprétation des auteurs selon laquelle les représentations

mentales de ces séquences par les grands singes seraient elles-mêmes hiérarchiques, plutôt que de consister en un stockage de ces différents programmes sous la forme de séquences strictement linéaires et fixes. Mais là encore, il est difficile de tirer une conclusion définitive.

Par ailleurs, ces observations ne permettent pas de conclure quant aux propriétés de récursivité ou de supra-régularité de ces systèmes. Nous avons vu dans la section I.5.2 que les procédures récursives ont pour particularité de permettre la génération de nouveaux niveaux hiérarchiques, au-delà de ceux donnés. Or, si le comportement alimentaire des gorilles décrit dans la Figure 12 implique l'application itérative d'une opération, cette itération se déroule au sein d'un même niveau hiérarchique, et n'accroît pas la profondeur de la hiérarchie. Cette limitation est toutefois inhérente à la relative simplicité de la tâche. Elle ne met donc pas en évidence une limitation cognitive chez cette espèce, mais rend inadéquate l'observation de ce comportement pour mettre à l'épreuve l'hypothèse d'une spécificité humaine pour la capacité de récursivité.

De plus, comme le soulignent les auteurs, à la différence du traitement d'une phrase, par exemple, qui requière un stockage en mémoire de travail des différents sous-éléments la composant afin de les intégrer et d'en former une représentation hiérarchique complète, ici aucun maintien en mémoire des actions passées et/ou de l'état d'avancement de la tâche n'est nécessaire. Les végétaux sur lesquels sont appliquées ces opérations sont en effet continuellement sous les yeux de l'animal. Cet aspect de la situation la rend par conséquent également impropre pour étudier la capacité d'un organisme à traiter des grammaires supra-régulières (qui se caractérisent, rappelons-le, par la possession a minima d'une mémoire à pile permettant de stocker les différentes étapes intermédiaires du traitement de la séquence).

II.1.3 Conclusions

L'importance de ces études réside donc essentiellement dans le fait qu'elles documentent une certaine variété de comportements naturels (i.e. non induits expérimentalement) pour lesquels des capacités syntaxiques et d'organisation hiérarchique, et éventuellement de récursivité et supra-régularité, pourraient avoir été sélectionnées au cours de l'évolution. Mais ces observations sont, pour une majorité d'entre elles, inadéquates pour déterminer précisément la nature des représentations sous-jacentes à ces différents comportements. Etudier ce dernier aspect nécessite la conduite d'expériences contrôlées.

Résumé

- **Des systèmes syntaxiques complexes, formalisables par des grammaires régulières, sont mis en évidence chez plusieurs espèces.**
- **Des représentations hiérarchiques sont susceptibles de sous-tendre ces systèmes de communication, ainsi que la représentation des relations sociales et la production de séquences d'actions complexes.**
- **Il n'y a cependant aucune indication claire sur l'existence d'un système supra-régulier, ou d'enchâssement récursif, chez les espèces ayant été étudiées.**
- **Il est difficile de déterminer la nature exacte des représentations sous-jacentes sans conduire des expériences contrôlées.**

II.2 Apprentissage de grammaires artificielles

Les études visant à situer les capacités computationnelles de différentes espèces animales au sein de la hiérarchie de Chomsky (1956, 1957) ont très largement dominé au cours de la dernière décennie. De fait, contrairement aux déclarations apparaissant dans certaines de ces études (voir section I.5), il n'existe pas à ce jour d'étude testant la capacité d'une espèce non-humaine à effectuer des computations récursives (Fitch & Friederici, 2012; Watumull et al., 2014). Après une introduction aux principes généraux de ces expériences, nous présentons rapidement les principales études comparatives réalisées sur l'apprentissage de grammaires régulières. Nous offrons ensuite une revue détaillée des expériences visant à tester l'hypothèse d'une spécificité humaine pour le traitement de grammaires supra-régulières.

II.2.1 Procédures expérimentales : principes de l'AGL

L'apprentissage de grammaires artificielles (en anglais, *Artificial Grammar Learning*, ci-après AGL) est un paradigme expérimental utilisé pour caractériser les processus à l'œuvre dans le traitement de la syntaxe chez l'adulte humain (e.g. Petersson, Folia, & Hagoort, 2012). Il est également largement employé pour étudier l'acquisition du langage (pour une revue, Gómez & Gerken, 2000), et l'évolution des capacités syntaxiques au cours de la phylogenèse (e.g. ten Cate, 2017; ten Cate & Okanoya, 2012; Fitch & Hauser, 2004). Il part du postulat que

l'on peut dissocier l'étude des traitements sémantiques et syntaxiques, en utilisant pour cette dernière des stimuli dépourvus de sens, tels que des syllabes (ne correspondant pas à des mots utilisés dans la langue des participants), des sons (notes musicales, sons, vocalisations), des formes géométriques, etc. Le principe est simple : ces stimuli sont présentés dans des séquences, organisées selon une certaine grammaire plus ou moins complexe, le but étant d'évaluer l'apprentissage de cette grammaire par les participants. Ces expériences sont divisées en deux phases. Dans une première phase d'exposition (également appelée phase de familiarisation, ou d'habituation), le participant est exposé à un ensemble de séquences générées par la grammaire d'intérêt. Dans une deuxième phase de test, de nouvelles séquences sont présentées. Classiquement, deux types de nouvelles séquences sont utilisées : la moitié est congruente avec la grammaire initiale, tandis que la deuxième moitié est non-congruente (i.e. les séquences comportent des violations de cette grammaire). Une réponse différentielle aux deux types de séquences test permet en principe d'inférer un apprentissage réussi de la grammaire sous-jacente⁸. Ces tâches peuvent être exclusivement comportementales, ou combinées avec des techniques d'imagerie cérébrales.

Les expériences d'AGL peuvent être implémentées dans divers types de tâches, qui peuvent différer selon que l'on teste des participants humains ou non-humains. Chez l'adulte humain, un premier type de tâche consiste à présenter aux participants une tâche de classification, postérieurement à la phase d'exposition passive. Les consignes fournies lors de cette tâche de classification sont plus ou moins directives. Dans la plupart des expériences, les participants doivent émettre des jugements de grammaticalité sur les séquences test, après avoir été informés que les séquences de la phase d'exposition correspondent à une certaine structure ou grammaire sous-jacente. Dans d'autres types de tests, il est simplement demandé aux participants de juger si les séquences test sont « similaires » ou « différentes » des séquences d'exposition, ou de donner un jugement de préférence, sans que l'existence d'une grammaire sous-jacente soit mentionnée (Forkstam, Elwér, Ingvar, & Petersson, 2008; Manza & Bornstein, 1995). Un deuxième type de tâche implémente le principe de l'AGL dans une tâche de pointage sériel (en anglais *Serial Response Time task*, ou *SRT* ; Nissen & Bullemer, 1987). Dans ces tâches, une cible se déplace sur un écran selon certaines régularités prédéfinies, et le participant doit toucher cette cible le plus vite possible à chacun de ses déplacements. Les temps de réponse

⁸ De nombreux facteurs confondus sont toutefois susceptibles d'exister dans ces expériences, et nous allons voir dans notre revue qu'une difficulté majeure des expériences AGL est de parvenir à prévenir l'utilisation par les participants de stratégies permettant de distinguer les séquences congruentes et non-congruentes sans traiter la grammaire sous-jacente.

reflètent les prédictions du participant quant à l'emplacement de la prochaine cible. Enfin, des tâches d'habituation-déshabituaton peuvent être employées chez les enfants pré-linguistiques : les réactions des participants aux différents types de séquences test sont alors mesurées par ses réponses d'orientation (e.g. nombre, durée, ou latence des réponses d'orientation vers le haut-parleur).

Pour l'expérimentation chez l'animal NH, on distingue trois grands types de protocoles. Comme pour les enfants pré-linguistiques, des protocoles d'habituation-déshabituaton peuvent être employés. Classiquement, les participants sont exposés à des séquences diffusées via un haut-parleur pendant une certaine durée, généralement assez courte (< 1h ; éventuellement augmentée d'une courte période de ré-habituaton précédant immédiatement le test). Un deuxième type de protocole implique un entraînement des participants, selon les principes du conditionnement opérant, à discriminer entre deux ensembles de séquences générées par deux grammaires différentes. Ces protocoles peuvent être implémentés dans des tâches à choix forcé (*forced choice tasks*, où le participant doit sélectionner le stimulus positif parmi deux alternatives, ou davantage) ou des tâches *Go/No-go* (où le sujet doit répondre, par exemple en appuyant sur un levier, aux stimuli positifs, et négliger de répondre aux stimuli négatifs). Les participants sont typiquement renforcés sur les essais de base (*baseline*) qui constituent une part importante des essais lors des tests, afin de maintenir la motivation des animaux, mais ne sont pas renforcés, ou le sont aléatoirement, dans les essais tests. Enfin, les animaux peuvent également être testés à l'aide d'une tâche de pointage sériel. Dans ce dernier cas, ils sont renforcés non différenciellement, i.e. qu'ils extraient ou non les régularités sous-jacentes au déplacement de la cible.

II.2.2 Etudes comportementales : grammaires régulières

Saffran et al. (2008) ont testé des enfants humains de 12 mois et des tamarins (*Saguinus oedipus*) avec une procédure d'habituation-déshabituaton. Des séquences auditives étaient diffusées aux participants, séquences générées à partir d'un vocabulaire comprenant cinq mots dépourvus de sens (A = *biff*, C = *cav*, D = *klor*, F = *dupp*, G = *jux*), et qui suivaient la grammaire régulière illustrée par la Figure 13.

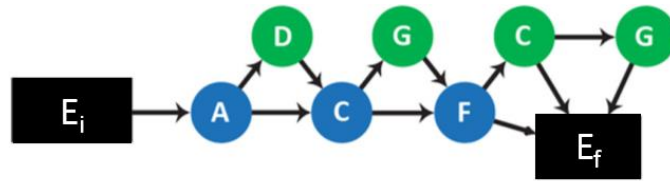


Figure 13 Diagramme représentant la grammaire régulière utilisée dans les différentes expériences présentées dans cette section. Les nœuds correspondent à différents éléments, obligatoires (en bleu) ou optionnels (en vert). Les séquences congruentes peuvent être générées en suivant n'importe quel chemin spécifié par les flèches, de l'état initial (E_i) à l'état final (E_f). Les séquences non-congruentes sont composées des mêmes éléments, mais leur ordre de présentation ne suit pas les flèches. Modifié d'après Wilson et al. (2013).

Testés ensuite à l'aide de nouvelles séquences composées de ces mêmes mots, les enfants (Expérience 1) comme les tamarins (Expérience 3) se sont davantage orientés vers le haut-parleur lors de la diffusion de séquences comportant des violations (e.g. ADF), par rapport à celle de séquences grammaticalement correctes (e.g. ACF). Toutefois, ils observent des différences entre les deux espèces : dans une autre expérience, les participants étaient exposés à un langage où chaque classe ne comportait plus seulement un mot, mais plusieurs (e.g. $A = \{biff, hep, mib, rud\}$, etc.). Les tamarins n'ont pas réagi différemment aux nouvelles séquences congruentes et non-congruentes (Expérience 4). Les bébés humains au contraire, testés avec le même matériel, ont réagi aux violations (Expérience 2). Selon les auteurs, ces résultats suggèrent donc que les tamarins partagent avec l'humain la capacité à reconnaître une grammaire régulière sous-tendant un jeu de séquences auditives. Mais elles indiquent également que dans des conditions expérimentales comparables, les enfants mais non les tamarins sont parvenus à extraire la grammaire d'un langage artificiel plus proche d'une langue naturelle (celle-ci comprenant typiquement plusieurs mots par catégorie syntaxique, plutôt qu'un seul).

Des répliques partielles de cette étude, effectuées chez le pigeon (*Columba livia* ; Herbranson & Shimp, 2003, 2008), le moineau du japon (*Lonchura striata* var. *domestica* ; Abe & Watanabe, 2011), le macaque rhésus (*Macaca Mulatta*) et le marmouset commun (*Callithrix jacchus* ; Wilson et al., 2013) ont également donné des résultats positifs. Dans cette dernière étude, les auteurs ont toutefois mis au jour des différences interspécifiques après une analyse plus fine de leurs résultats. Cette analyse détaillée révèle que les marmousets réagissent essentiellement à des violations touchant le premier élément de la séquence (i.e. réagissent lorsque la séquence ne commence pas par l'élément A). Les macaques rhésus, eux, réagissent correctement aux violations survenant à d'autres positions de la séquence.

Dans une revue critique de l'ensemble de ces études⁹, Beckers, Berwick, Okanoya, et Bolhuis (2016) ont souligné qu'elles comportaient systématiquement un facteur confondu avec la grammaticalité des séquences : leur similarité acoustique. En effet, une analyse détaillée de leurs matériels expérimentaux révèle que les nouvelles séquences grammaticales présentées lors du test comportent systématiquement des fragments de séquences présentées lors de la période d'habituation, tandis que les séquences agrammaticales en comportent moins, ou aucun. Ces résultats peuvent donc également être expliqués par la détection des violations de ces fragments par les participants, plutôt que par la détection de violations de la grammaire *per se*. Cette critique ne touche cependant pas spécifiquement les études menées chez l'animal NH. Perruchet et Pacteau (1990) ont par exemple montré que les performances de leurs participants dans une tâche classique de jugement de grammaticalité ne différaient pas, qu'ils aient été préalablement exposés à des séquences suivant cette grammaire ou exposés à une simple liste des bigrammes composant ces séquences. Distinguer les effets de la similarité perceptive entre les séquences et les effets de grammaticalité nécessite donc un contrôle minutieux de la fréquence d'occurrence de fragments de séquences familières, au sein des séquences test congruentes et non-congruentes (e.g. Meulemans & Van der Linden, 1997)¹⁰. A notre connaissance, aucun contrôle expérimental de ce type n'a été réalisé chez l'animal NH.

II.2.3 Etudes comportementales : grammaires supra-régulières

II.2.3.1 Expérience princeps

La première étude testant l'hypothèse selon laquelle seul l'humain serait capable de traiter des grammaires supra-régulières a été effectuée par Fitch et Hauser (2004). La capacité d'une espèce de primate non-humain, le tamarin à crête blanche (*Saguinus oedipus*), à discriminer des langages générés par les grammaires A^nB^n et $(AB)^n$ a été testée par une tâche d'habituation-déshabituation (Figure 14). Comme nous l'avons vu dans la section I.2, la première de ces grammaires constitue un exemple paradigmatique de grammaire supra-

⁹ A l'exception des expériences d'Herbranson et Shimp (2003,2008), dans lesquelles le même type d'analyse avait toutefois été réalisé, menant à des conclusions similaires à celles de Beckers et al. (2016).

¹⁰ L'effet de la fréquence d'occurrence de bigrammes et trigrammes familiers au sein des nouvelles séquences grammaticale et non-grammaticales était contrôlé dans l'expérience de Saffran et al. (2008) par l'ajout d'une condition contrôle présentant un langage *non prédictif* contenant ces n-grammes. L'analyse réalisée par Beckers et al. (2016) montre toutefois que la procédure employée est insuffisante pour permettre d'écarter définitivement cette explication alternative.

régulière, et la deuxième est un exemple de grammaire régulière (voir également O'Donnell, Hauser, & Fitch, 2005). Des participants humains étaient également testés.

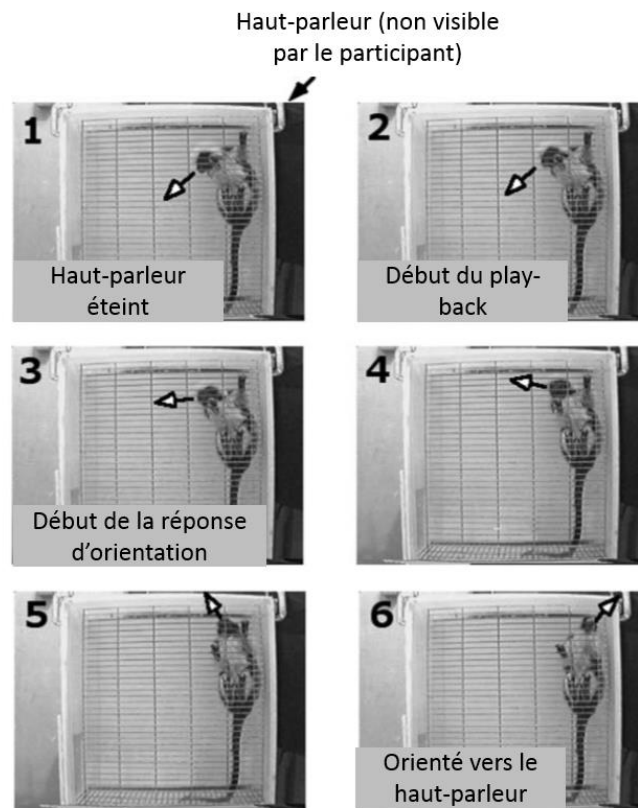


Figure 14 Illustration de la procédure d'habituation-déshabituatoin employée par Fitch et Hauser (2004). Modifié d'après Hauser, Newport, et Aslin (2001).

Les participants étaient d'abord familiarisés avec une série de séquences auditives générées par l'une ou l'autre de ces grammaires. Le vocabulaire était constitué de syllabes énoncées par des voix humaines, dont la classe était marquée perceptivement : les syllabes implémentant la classe A et celles implémentant la classe B étaient respectivement énoncées par une voix féminine et une voix masculine. Ces séquences de familiarisation comportaient quatre ou six éléments, donnant par exemple, « *di yo gu do* » et « *la no yo mo bi gu* » pour la grammaire A^nB^n , et « *di mo yo pa* » et « *di pa wu bi tu nu* » pour la grammaire $(AB)^n$ (les syllabes en gras correspondant à la voix féminine). De nouvelles séquences de même longueur (i.e. $n = \{2, 3\}$) et générées à partir du même vocabulaire étaient ensuite diffusées lors du test. La moitié de ces séquences test correspondait à la grammaire A^nB^n , et l'autre à $(AB)^n$.

Au cours de cette phase de test, le groupe de tamarins familiarisés à la grammaire $(AB)^n$ a exprimé davantage de réponses d'orientation en direction du haut-parleur lors de la diffusion

des séquences non-congruentes ; tandis que les tamarins familiarisés à la grammaire A^nB^n n'ont pas exprimé de réponse différentielle aux séquences congruentes et non-congruentes. Les humains, eux, ont davantage classifié les séquences non-congruentes comme étant différentes des séquences familières, quel qu'ait été le langage avec lequel ils avaient été familiarisés. Selon les auteurs, ces résultats suggèrent qu'au moins une espèce de primates NH, le tamarin, est capable de traiter et de reconnaître une grammaire régulière, mais pas une grammaire supra-régulière. Ces données supportent donc l'hypothèse selon laquelle seule l'espèce humaine aurait les capacités computationnelles nécessaires pour traiter ce dernier type de grammaire.

II.2.3.2 Revue critique des données chez l'humain

Un certain nombre de publications offrent une analyse critique de cette étude princeps (M. H. de Vries et al., 2008; Hochmann, Azadpour, & Mehler, 2008; Perruchet & Rey, 2005). Toutes mettent en avant le principe suivant : l'échec des tamarins à traiter la grammaire supra-régulière dans cette tâche n'est pertinent, au regard des conclusions tirées par les auteurs, que si les participants humains ont effectivement effectué la tâche en traitant cette grammaire. Avant d'examiner deux des points pertinents soulevés par ces analyses, il nous faut signaler que l'une des critiques adressées est que les humains pourraient avoir « simplement compté et comparé le nombre d'éléments A et B » (e.g. Corballis, 2007; Perruchet & Rey, 2005). Cette critique ne sera pas retenue ici, car l'emploi d'une telle stratégie satisfait pleinement à la caractérisation de supra-régularité, selon la théorie des langages formels (voir section I.2). En effet, comparer le nombre d'éléments A et B nécessite le stockage temporaire en mémoire du résultat de l'opération de comptage des éléments A. Simplement, elle ne requière pas de traiter de dépendances non-adjacentes centrées-emboîtées.

Une première critique émise à l'encontre de l'étude de Fitch et Hauser (2004) porte sur l'utilisation de voix humaines féminines et masculines pour implémenter les éléments A et B (Perruchet & Rey, 2005). L'argument de Perruchet et Rey (2005) est que le choix de catégories perceptives différant de manière aussi saillante induirait des différences majeures de caractéristiques acoustiques entre les deux types de séquences : les séquences générées par la grammaire A^nB^n ne comportent qu'une seule transition entre les deux types de voix, tandis que celles générées par la grammaire $(AB)^n$ en comportent trois à cinq. Par conséquent, ces dernières sont plus variables (notamment en termes de hauteur). Perruchet et Rey (2005 ; voir également Hochmann et al., 2008) proposent que la détection de ces différences acoustiques ait pu être utilisée par les participants humains dans l'expérience de Fitch et Hauser (2004) pour

discriminer les deux types de séquences. Toutefois, cette stratégie étant relativement simple, les tamarins auraient *a priori* également pu y recourir, ce que les résultats ne suggèrent pas. Perruchet et Rey formulent donc une hypothèse complémentaire pour tenter d'expliquer le comportement des tamarins : ils proposent qu'une séquence de forme A^nB^n pourrait davantage être perçue comme une « conversation naturelle », et donc comme indicatrice de la présence d'humains, qu'une séquence de type $(AB)^n$ (en partant de l'idée que l'alternance de monosyllabes constituerait un mode de conversation rare). La réaction d'orientation des tamarins lors de la diffusion de séquences A^nB^n après une exposition à des grammaires $(AB)^n$ pourrait ainsi découler d'un intérêt accru pour la « conversation », tandis que la présentation de séquences $(AB)^n$ suite à une familiarisation avec des séquences A^nB^n ne susciterait pas le même intérêt (et donc, moins de réponses d'orientation).

Une deuxième série de critiques de l'expérience de Fitch et Hauser (2004) souligne l'absence de tests de généralisation et de violation appropriés pour déterminer ce qui a été appris par les participants. Deux types de tests sont requis (Fitch & Friederici, 2012; Hochmann et al., 2008; O'Donnell et al., 2005). Le premier type de test à appliquer est un test d'extension, qui doit vérifier que la grammaire apprise n'est pas trop restrictive, e.g. qu'elle n'est pas limitée à des séquences où $n = \{2,3\}$ comme dans les séquences de familiarisation, mais généralisable à tout nouveau n ¹¹. Le second type de test doit vérifier que la grammaire apprise n'est pas trop inclusive. La présentation de séquences comportant un nombre inégal d'éléments A et B (i.e. A^nB^m , où $n \neq m$) est requise pour contrôler que les sujets ne vérifient pas simplement qu'une succession d'éléments A est suivie d'une succession d'éléments B (grammaire formellement notées $A*B^*$; Hochmann et al., 2008; Perruchet & Rey, 2005). Hauser et Fitch (2004) rapportent dans leurs *Informations Supplémentaires* que dans un test additionnel, les participants humains ont correctement généralisé à des séquences où $n = 4$, suggérant que leur reconnaissance de la grammaire n'est pas limitée au n familier. Ils ne rapportent toutefois pas de test impliquant des séquences A^nB^m . Hochmann et collaborateurs (2008) ont toutefois répliqué l'expérience de Fitch et Hauser (2004), en y ajoutant un test impliquant des séquences non-congruentes de ce type (AABBB et AAABB ; Expérience 2). Leurs participants n'ont pas classifié ces séquences comme étant grammaticalement incorrectes, ce qui suggère qu'ils se sont basés sur une stratégie plus simple que celle consistant à compter et comparer le nombre

¹¹ Dans la limite des capacités perceptives et mnésiques de l'espèce testée (e.g. Fitch & Friederici, 2012). Ainsi, bien que la grammaire A^nB^n puisse générer un nombre de séquences théoriquement non fini, les capacités de mémoire de travail imposent des restrictions sur la longueur des séquences qui pourront être traitées et donc reconnues ; restrictions pouvant varier d'une espèce à l'autre.

d'éléments de chaque classe (e.g. ils ont pu simplement vérifier que les séquences étaient constituées d'une série de syllabes énoncées par des voix féminines, suivies de syllabes énoncées par une voix masculine).

Un dernier point à mentionner, qui rejoint les critiques mentionnées à l'encontre des expériences d'apprentissage de grammaires régulières, est la présence de caractéristiques de surface en commun dans les séquences familières et les séquences test congruentes. En effet, les nouvelles séquences présentées lors du test sont générées à partir du même vocabulaire que les séquences utilisées au cours de la période d'habituation. Ainsi, les syllabes « mo » et « yo » par exemple, qui n'auront jamais été entendues à la suite par les participants familiarisés à la grammaire A^nB^n , pourront l'être dans une séquence non-congruente, mais pas dans une nouvelle séquence congruente.

Il existe donc un certain nombre de stratégies alternatives ayant pu être employées par les participants de l'expérience de Fitch et Hauser (2004), plutôt que de traiter les grammaires sous-jacentes. La plupart de ces interprétations s'appliquent également à une étude menée à la même période, dans laquelle les corrélats neuronaux du traitement des grammaires A^nB^n et $(AB)^n$ étaient examinés chez l'humain (Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006). Une liste des stratégies alternatives pouvant être utilisées pour discriminer des séquences générées par les grammaires A^nB^n et $(AB)^n$ est présentée dans la Table 1.

Table 1 **Inventaire des stratégies alternatives pour discriminer des séquences générées par les grammaires A^nB^n et $(AB)^n$**

Stratégies	Conditions d'acceptation	
	A^nB^n	$(AB)^n$
Régularités de surface		
Variabilité perceptive	faible	forte
Mémorisation des exemplaires	contient une séquence/fragment familier	
Primauté	commence par "aa"	commence par "ab"
Récence	fini par "bb"	fini par "ab"
Bigramme	contient "aa"/"bb"	contient "ba"
Transition inter-catégories (a-b)	en contient une	en contient plusieurs
Régularités grammaticales		
Grammaire trop inclusive	contient toute alternance d'éléments "a" et "b"	contient tout nombre d'éléments "a" suivis d'éléments "b" (i.e. A^*B^*)
Grammaire trop restrictive	contient le même nombre de paires "ab" que les séquences familières	contient le même nombre d'éléments "a" suivis d'éléments "b" que les séquences familières

Note : Modifié d'après Ravignani, Westphal-Fitch, Aust, Schlumpp, & Fitch (2015)

La majorité des expériences conduites chez l'humain suite à ces expériences princeps ont été menées avec des grammaires A^nB^n indexées (Ojima & Okanoya, 2014)¹². Dans ces grammaires, les différents éléments A et B sont appariés, et ces dépendances sont organisées de manière centrée-emboîtée (e.g. $A_1A_2B_2B_1$, les index indiquant les appariements : A_1 est apparié avec B_1 , mais pas avec B_2 , etc.) ou croisée (e.g. $A_1A_2B_1B_2$). Pour effectuer la tâche, le participant ne doit pas compter et comparer le nombre d'éléments A et B, mais doit par exemple vérifier l'organisation centrée-emboîtée des dépendances, afin de repérer des violations de cette organisation (e.g. $A_1A_2A_3B_3B_1B_2$). Ces grammaires préviennent l'utilisation d'un certain nombre des stratégies listées dans la Table 1. Par exemple, la stratégie consistant à se baser sur la variabilité perceptive des séquences, comme celle consistant à vérifier qu'une série d'éléments A est suivie d'une série d'éléments B (que leur nombre soit égal ou non), ne sont plus utilisables avec ce type de grammaire (Corballis, 2007; M. H. de Vries et al., 2008; Perruchet & Rey, 2005)). Les premières études recourant à ce type de grammaire rapportent

¹² Ou avec une combinaison des deux procédés (i.e. avec marquage perceptif des classes A et B, et un appariement entre les éléments, e.g. Bahlmann, Schubotz, & Friederici, 2008).

des résultats négatifs, les participants n'ayant pas été capables de détecter des violations de la grammaire (e.g. (M. H. de Vries et al., 2008; Perruchet & Rey, 2005). Ces expériences montrent également que les participants humains testés exploitent des indices plus simples lorsque ceux-ci sont disponibles. Par exemple, de Vries et al. (2008, Expérience 2) observent que leurs participants basent leurs jugements de grammaticalité sur la détection de la présence/absence d'éléments répétés dans les séquences test.

Par la suite, un certain nombre d'études ont rapporté des résultats positifs chez l'humain, sous certaines conditions. Il faut notamment utiliser des tâches impliquant un entraînement explicite (Bahlmann et al., 2008), administrer les tests après une période d'exposition de longueur conséquente (constituée de plusieurs sessions d'écoute réparties sur plusieurs journées, Uddén, Ingvar, Hagoort, & Petersson, 2012), ou exposer au préalable les sujets aux éléments appariés présentés de manière adjacente (Lai & Poletiek, 2011). Avec des aménagements similaires de procédures, des résultats positifs ont également été obtenus avec des stimuli non linguistiques (e.g. avec des figures visuelles, Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009).

Pris dans leur ensemble, ces résultats indiquent donc que, sous certaines conditions, les humains peuvent traiter des grammaires supra-régulières A^nB^n dans des paradigmes d'AGL. En parallèle de ces expérimentations menées chez l'humain, la question d'une spécificité humaine pour ce type de traitement a été étudiée par une série d'études menées chez d'autres espèces.

II.2.3.3 Revue critique des données chez l'animal NH

Grammaires non-indexées

On peut recenser trois études menées à la suite de l'étude de Fitch et Hauser (2004) visant à examiner la capacité de différentes espèces à traiter des grammaires supra-régulières, et utilisant des grammaires non-indexées. Toutes ont été menées chez des oiseaux.

Gentner, Fenn, Margoliash, et Nusbaum (2006) ont été les premiers à annoncer l'obtention de résultats positifs chez une espèce NH, l'étourneau sansonnet (*Sturnus vulgaris*). Ils ont entraîné leurs onze participants à discriminer huit séquences auditives de la forme A^2B^2 et huit séquences de la forme $(AB)^2$ dans une tâche Go/No-go. Les éléments A et B étaient instanciés par huit « cliquetis » et huit « gazouillis », respectivement, extraits du répertoire

vocal d'un individu de cette espèce. Ils ont ensuite soumis les quatre individus ayant rapidement atteint le critère d'apprentissage à une série de tests, dans lesquels leurs étaient présentés les types de séquences suivants : (1) nouvelles séquences grammaticales de même longueur, (2) séquences grammaticales de longueur étendue (A^3B^3 , A^4B^4), (3) séquences agrammaticales ne correspondant ni aux deux grammaires examinées, ni à une grammaire A^*B^* (AAAA, BBBB, ABBA, BAAB), et (4) séquences agrammaticales ne correspondant pas aux deux grammaires examinées, mais correspondant à une grammaire A^*B^* (A^1B^3 , A^3B^1 , A^2B^3 et A^3B^2).

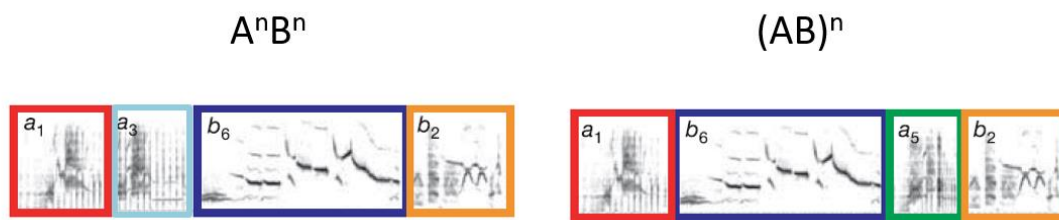


Figure 15 Exemples de stimuli diffusés aux étourneaux dans l'expérience de Gentner et al. (2006). Les cadres de couleur dénotent des syllabes identiques.

Les résultats indiquent que les étourneaux ont discriminé les séquences agrammaticales AAAA, BBBB, ABBA et BAAB des nouvelles séquences grammaticales. Selon les auteurs, ces résultats permettent d'éliminer trois stratégies alternatives : « primauté », « récence », et « bigramme » (cf. Table 1). Pour l'élimination de la stratégie « primauté » par exemple, leur argumentaire est le suivant : un oiseau qui aurait appris à discriminer entre les deux ensembles de séquences sur la base de leurs deux premiers items seulement (i.e. à discriminer entre AA-- et AB--), devrait classer AAAA comme A^2B^2 , et ABBA comme $(AB)^2$. Or, les étourneaux discriminent mieux les nouvelles séquences A^2B^2 et $(AB)^2$ que les séquences AAAA et ABBA (voir Figure 16). Des arguments similaires sont développés pour éliminer les stratégies « récence » (i.e. discrimination entre --BB et --AB) et « bigramme » (présence/absence de transition B-A).

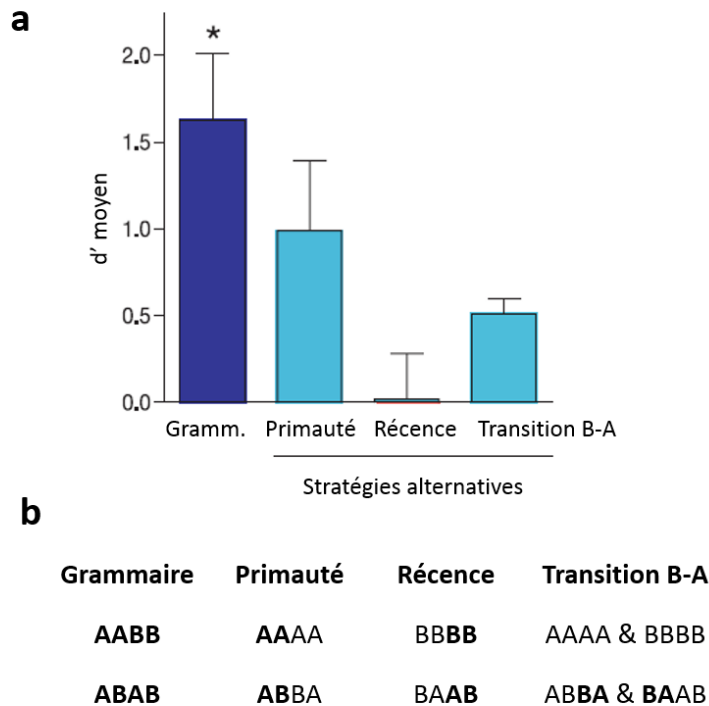


Figure 16 **a.** Performances de discrimination des étourneaux pour les nouvelles séquences congruentes (« Grammaire ») et pour chacune des stratégies alternatives. L’astérisque indique que la valeur d' obtenue pour la stratégie « grammaire » diffère significativement ($p < .05$) des valeurs d' calculées pour les autres stratégies. **b.** Contrastes examinés par Gentner et al. (2006) afin de calculer les valeurs de d' pour chacune des stratégies potentielles. Les régularités discriminatrices sont en gras. Modifié d’après Gentner et al. (2006).

Par ailleurs, les étourneaux ont passé avec succès le test d’extension (e.g. AAABBB *versus* ABABAB), et ont correctement discriminé ces séquences grammaticales de longueur étendue des séquences où le nombre d’éléments A et B était inégal¹³. Les auteurs concluent de cet ensemble de résultats que les étourneaux testés ont été capables de traiter une grammaire supra-régulière A^nB^n , sur la base d’une stratégie consistant à compter et comparer le nombre d’éléments A et B constituant chaque séquence.

Van Heijningen, Visser, Zuidema, & ten Cate (2009) ont soulevé deux problèmes posés par l’étude de Gentner et al. (2006). Le premier concerne la présence de fragments de séquences

¹³ Dans les *Informations Supplémentaires* du papier de Gentner et al. (2006), une information étonnante apparaît dans la légende de la Table S1 : « Stimuli marked with (*) comprised speech syllables instead of starling song motifs. ». Or, les seuls stimuli accompagnés d’un astérisque sont ceux des séquences agrammaticales $A*B*$ (faisant référence aux séquences où le nombre d’éléments A et B est inégal). Il pourrait s’agir d’une erreur. Mais si les séquences $A*B*$ ont effectivement été instanciées par des syllabes, cela invaliderait les résultats de ce dernier test crucial : les étourneaux pourraient avoir correctement rejeté les séquences $A*B*$ non pas en raison de l’inégalité du nombre d’éléments A et B, mais du fait de la nouveauté des stimuli utilisés.

familiales au sein des séquences congruentes, point sur lequel nous reviendrons plus bas. Le second concerne la méthode d'analyse statistique employée. Ces auteurs remarquent en effet que si l'indice d' est plus important pour la discrimination entre les nouvelles séquences A^2B^2 et $(AB)^2$ que pour les autres contrastes effectués (cf. Figure 16), la valeur d' obtenue pour la stratégie primauté est néanmoins significativement différente de zéro¹⁴. Ce dernier résultat indique que les étourneaux discriminent AAAA de ABBA, alors qu'aucun de ces patterns ne correspond à l'une ou l'autre des grammaires étudiées. Les oiseaux se basent donc probablement, dans certains essais au moins, sur les deux premiers items de la séquence pour effectuer la tâche. Gentner et al. (2006) tentent d'expliquer ce phénomène en proposant que la stratégie primauté soit une stratégie « par défaut », adoptée par les oiseaux lorsque les séquences présentées ne correspondent ni à un stimulus positif, ni à un stimulus négatif. Cette stratégie serait couplée à une tendance à porter davantage attention aux régularités présentes en début de séquence qu'à celles situées en milieu ou fin de séquence. Van Heijningen et al. (2009) offrent une explication alternative, selon laquelle à aucun moment les oiseaux n'ont traité la grammaire sous-jacente, et proposent que les résultats s'expliquent par des différences de stratégies individuelles. En effet, si une partie des individus se base sur les premiers items de la séquence pour effectuer la tâche, et une autre partie d'entre eux sur la présence/absence de transition B-A, alors les valeurs d' moyennées sur le groupe pour l'une ou l'autre de ces stratégies seront nécessairement plus faibles que pour la discrimination entre A^2B^2 et $(AB)^2$, où les deux indices sont présents. Gentner et al. (2006) n'ont malheureusement pas rapporté les valeurs d' pour chacun des individus testés (y compris dans leur réponse à van Heijningen et al., 2006 ; cf. Gentner, Fenn, Margoliash, & Nusbaum, 2010), ce qui ne permet pas d'invalidier cette hypothèse.

Van Heijningen et al. (2009) ont donc partiellement répliqué l'expérience de Gentner et al. (2006) en reprenant l'essentiel de leur méthodologie, mais en s'intéressant aux patterns de réponse individuels. Cette expérience a été réalisée chez une autre espèce d'oiseaux chanteurs, le diamant mandarin. Après un entraînement à la discrimination de séquences $(AB)^2$, et A^2B^2 (les éléments A et B étant des syllabes issues du répertoire vocal de cette espèce), ils ont exposé les oiseaux à un sous-ensemble des séquences test utilisées par Gentner et al (2006). En guise d'illustration, le taux de réponse moyen obtenu pour ces différentes séquences chez les deux oiseaux entraînés avec A^2B^2 comme S+, sont les suivants :

¹⁴ D'après les valeurs d' rapportées dans Gentner et al. (2006), c'est également le cas pour la stratégie bigramme. Van Heijningen, Visser, Zuidema, & Cate (2009) ne mentionnent toutefois pas ce cas-ci.

	ABAB	AABB	AAAB	AAAA	BBBB	BAAB	ABBA	AAABBB	AAAABBBB
Oiseau 1	1.5	28.0	22.5	22.5	0.0	8.5	1.5	17.5	14.0
Oiseau 2	1.0	24.5	8.0	8.0	23.0	9.5	1.5	20.0	21.0

On peut remarquer que le premier individu présente un taux de réponse important chaque fois que le pattern commence par deux éléments A, et beaucoup plus faible voire nul dans le cas contraire. Le deuxième individu présente le plus fort taux de réponses pour les patterns terminant par deux éléments B. En soumettant ces données à une analyse par sélection de modèle basée sur l'estimation du maximum de vraisemblance (en anglais, *maximum likelihood*), les auteurs confirment que les stratégies le plus probablement employées par ces deux oiseaux sont, respectivement, une stratégie « primauté » et une stratégie « récence ». Les deux oiseaux entraînés avec $(AB)^2$ comme S+ se seraient eux basés sur une stratégie « bigramme » consistant à négliger de répondre à toute séquence contenant "aa". Plutôt que d'effectuer une série de comparaisons deux-à-deux, à l'instar de Gentner et al. (2006), cette méthode d'analyse permet d'examiner avec quelle(s) stratégie(s) les réponses de chaque individu sont les plus compatibles.

Dans leur réponse à l'article de van Heijningen et al. (2009), Gentner, Fenn, Margoliash, et Nusbaum (2010) pointent une faiblesse de leur argumentation : il n'est nulle part fait mention du test de Gentner et al. (2006) impliquant des séquences agrammaticales compatibles avec une grammaire $A*B^*$. Or, ce test crucial avait été passé avec succès par les étourneaux. Toutefois, ici encore, la même logique peut être appliquée : en l'absence des données individuelles, il reste envisageable que les données moyennées sur le groupe ne reflètent pas les comportements individuels sur ce test également. Et bien que l'expérience de van Heijningen et al. (2009) n'offre pas une réplication avec la même espèce, et dans des conditions d'apprentissage strictement similaires (voir Gentner et al., 2010), cette hypothèse est corroborée par une deuxième étude.

(Ravignani et al., 2015)¹⁵, ont employé une méthode d'analyse similaire à celle de van Heijningen et al. (2009), pour mettre au jour les stratégies employées par deux autres espèces

¹⁵ Dans cette étude, les grammaires étaient en réalité indexées (voir Stobbe, Westphal-Fitch, Aust, & Fitch, 2012). Les tests dispensés étant toutefois limités à des tests typiques des expériences utilisant des

d'oiseaux : des kéas (*Nestor notabilis*) et des pigeons (*Columba livia*). Dans une précédente expérience (Stobbe, Westphal-Fitch, Aust, & Fitch, 2012), ces individus avaient été entraînés à discriminer des séquences de la forme A^nB^n et $(AB)^n$, dans une tâche à choix forcé. A la différence des expériences précédentes, celles-ci étaient implémentées dans des chaînes de stimuli visuels présentés simultanément sur un écran tactile ($n = \{2,3\}$), les éléments A et B correspondant à des stimuli composites répartis en deux catégories aisément discriminables, différant l'une de l'autre par leur gamme de couleurs et de formes, voir la Figure 17). Les oiseaux avaient ensuite été testés sur des séquences de longueur étendue, et des séquences de la forme A^nB^m (où $n \neq m$). Les participants des deux espèces avaient correctement généralisé leurs réponses à ce premier type de séquence, mais n'ont pas rejeté les séquences où le nombre de A et de B était déséquilibré.

A partir de ces données, enrichies de tests additionnels, Ravignani et al. (2015) ont examiné plus en détail les stratégies alternatives employées par ces oiseaux. En utilisant une analyse par sélection de modèle basée sur l'estimation du maximum de vraisemblance, ils ont mis au jour une certaine diversité de stratégies, entre espèces et entre individus. Pour une majorité de Kéas entraînés avec A^nB^n comme séquence stimulus positif (S+), la stratégie la plus probablement utilisée est la stratégie « récence » (cf. Table 1). Pour ceux entraînés avec $(AB)^n$ comme séquence S+, la stratégie utilisée était la détection d'un bigramme "BA", bigramme qui n'apparaissait que dans ces séquences. Chez les pigeons, une plus grande diversité de stratégies est observée : le groupe entraîné avec A^nB^n comme S+ a eu majoritairement recours à des stratégies locales (« récence » et « bigramme »). Quatre des cinq pigeons entraînés à sélectionner les séquences $(AB)^n$, auraient recouru à la fois à une stratégie basée sur la variabilité visuelle globale des séquences, et à une stratégie consistant à sélectionner la séquence ressemblant le plus aux séquences d'entraînement (i.e. incluant les séquences familières, ou comportant le moins de substitutions d'éléments par rapport à ces séquences). Enfin, un dernier individu aurait utilisé, en plus de la stratégie basée sur la variabilité, une stratégie consistant à traiter la grammaire $(AB)^n$.

grammaires non-indexées (i.e. absence de test de violation des dépendances), nous avons préféré l'inclure dans la section « Grammaires non indexées ».



Figure 17 Exemples de stimuli employés par Stobbe et al. (2012) et Ravignani et al. (2015).

Pris dans leur ensemble, les résultats de van Heijningen et al. (2009) et Ravignani et al. (2015) suggèrent donc que le traitement des grammaires n'est pas la stratégie privilégiée par les individus testés. Les participants des différentes espèces d'oiseaux testés montrent une forte propension à se baser sur des indices locaux (i.e. ne concernant que des stimuli adjacents : « primauté », « récence », « bigramme »), lorsque des grammaires non indexées sont employées. On observe toutefois une diversité dans les stratégies adoptées : les résultats de l'étude de Ravignani et al. (2015), qui offre une comparaison inter-espèce dans des conditions expérimentales similaires, suggèrent que cette diversité reflète des différences interspécifiques. Le mode de présentation des séquences semble toutefois également avoir un impact sur la stratégie privilégiée : dans cette expérience où les oiseaux sont testés avec des chaînes de stimuli visuels présentés simultanément, la stratégie « récence » est globalement privilégiée par rapport à la stratégie primauté. Chez les oiseaux de Gentner et al. (2006), en revanche, c'est la stratégie « primauté » qui semble prévaloir sur celle de « récence ».

Une deuxième critique émise par van Heijningen et al. (2009) à l'encontre de l'expérience de Gentner et al. (2006) concerne l'absence de test de généralisation des régularités apprises à de nouveaux éléments. Van Heijningen et al. (2009) ont donc conduit un test additionnel, dans lequel des séquences congruentes et non-congruentes étaient construites à partir de nouveaux stimuli (d'autres types de syllabes issues du répertoire vocal de cette espèce, notées C et D). Les oiseaux échouent à discriminer ces séquences, indiquant qu'ils ont appris à détecter des répétitions immédiates d'éléments particuliers (AA ou BB, selon les individus), et non à détecter la présence d'éléments immédiatement répétés, quels qu'ils soient. Ces résultats suggèrent que les animaux testés exploitent préférentiellement des informations locales de surface, plutôt que (1) les grammaires sous-jacentes, mais également que (2) des régularités structurelles plus simples (ici, la présence de répétitions dans les séquences générées par la grammaire $A^n B^n$, absentes de celles générées par la grammaire $(AB)^n$).

Grammaires indexées

L'emploi de grammaires indexées constitue une bonne alternative à l'utilisation de grammaires non-indexées, car ces grammaires empêchent l'utilisation de certaines stratégies locales que nous venons d'identifier. En effet, lorsqu'une grammaire non indexée est employée, les classes d'éléments A et B doivent nécessairement être marquées perceptivement, ce qui rend par exemple possible l'utilisation de stratégies consistant à repérer la présence/absence de répétitions immédiates d'éléments appartenant à une même classe perceptive. Les grammaires indexées reposent sur des appariements entre éléments arbitraires, sans que le marquage perceptif des classes soit nécessaire. Ce type de grammaire prévient donc l'utilisation de telles stratégies.

On recense à ce jour deux expériences menées avec une grammaire indexée centré-emboîtée, l'une conduite chez une espèce d'oiseau, et l'autre chez une espèce de primate NH. A notre connaissance, aucune étude examinant le traitement d'une grammaire indexée croisée n'a été menée à ce jour.

Abe et Watanabe (2011) ont utilisé une procédure d'habituation-déshabitude chez le moineau du japon (*Lonchura striata* var. *domestica*). Les oiseaux ont tout d'abord été exposés durant une heure à des séquences acoustiques de la forme A^nCB^n ($n = \{1,2\}$). Chaque élément correspondait à une « syllabe » provenant du répertoire vocal d'individus de leur espèce, arbitrairement attribué à l'une des trois classes A, B ou C (i.e., les éléments composant les classes A B et C n'appartenaient pas à des classes de vocalisations différentes), avec quatre syllabes par classe (notées A_{1-4}). Lors de la construction du matériel expérimental, chaque élément A a été arbitrairement associé à un élément B, ce qui produit des dépendances centré-emboîtées entre ces éléments, organisées autour d'un élément central C (e.g. $A_1C_1B_1$, $A_2A_1C_1B_1A_2$, les indices signalant les éléments associés). Il y avait 52 séquences de familiarisation au total. Les oiseaux étaient ensuite exposés à une série de tests, au cours desquels le nombre de vocalisations émis par l'individu testé était mesuré.

Les séquences test grammaticales étaient de deux types : nouvelles séquences n'excédant pas la longueur des séquences familières ($n = 2$), et nouvelles séquences de longueur étendue ($n = 3$). Les oiseaux n'ont pas présenté d'augmentation significative de leurs réponses à l'écoute de ces nouvelles séquences, par rapport à la période contrôle. Des séquences agrammaticales étaient également présentées, pouvant être réparties en trois types, en fonction des violations qu'elles comportaient : violation de position d'une classe (e.g. $A_1A_2B_2C_1B_1$, les

caractères en gras signalant la violation ; correspond aux tests *ACF*, *Random* et *AES2*), violation de dépendance (e.g. $A_1A_2C_1B_2B_4$, test *AFP*), et violation de l'emboîtement (i.e. séquences dans lesquelles les bonnes syllabes étaient présentées, mais pas dans le bon ordre, e.g. $A_1A_2A_3C_2B_1B_3B_2$, test *AES*). Pour chacun de ces tests, les auteurs observent une augmentation significative des vocalisations émises, par rapport à une période contrôle. Les auteurs en concluent que ces oiseaux ont été capables de discriminer les séquences sur la base de la grammaire supra-régulière sous-jacente.

Cette étude a fait l'objet d'une revue critique par Beckers, Bolhuis, Okanoya, et Berwick (2012). Ces auteurs effectuent une analyse détaillée du matériel expérimental employé par Abe et Watanabe, et observent que si l'on néglige une syllabe (la syllabe C), toutes les nouvelles séquences grammaticales contiennent une séquence familière. En effet, les nouvelles séquences ne sont pas construites à partir de nouvelles combinaisons de paires, mais en présentant des combinaisons de paires déjà présentes dans les séquences familières, autour d'un nouvel élément C. Par exemple, les séquences $A_1A_2C_2B_2B_1$, $A_1A_2C_3B_2B_1$ et $A_1A_2C_4B_2B_1$ étaient présentées lors de la période de familiarisation, et les séquences $A_1A_2C_1B_2B_1$ et $A_3A_1A_2C_1B_2B_1B_3$ étaient présentées dans le test de généralisation (nous soulignons les fragments partagés). Les séquences non congruentes, elles, ne comportaient pas, ou comportaient moins de fragments en commun avec les séquences familières (e.g. $A_1A_2C_1B_2B_4$). Par conséquent, la présence de longs fragments de séquences acoustiques familières dans les nouvelles séquences grammaticales, présents dans une moindre mesure dans les séquences agrammaticales, pourrait expliquer le pattern de réponse des oiseaux, sans qu'ils aient traité la grammaire sous-jacente¹⁶.

Une dernière étude, employant une procédure radicalement différente, a été menée dans notre laboratoire (Rey, Perruchet, & Fagot, 2012). Cette étude ne visait pas à démontrer la capacité d'une espèce NH à traiter des grammaires supra-régulières. Elle était davantage destinée à étudier les mécanismes cognitifs impliqués dans la production de structures centré-emboîtées chez une espèce de primates NH. Par conséquent, elle ne comporte pas les tests de violations et généralisation classiquement présents dans cette littérature. Les participants, onze babouins de guinée (*Papio papio*), étaient d'abord entraînés à associer six paires de stimuli visuels, dans une tâche d'appariement réalisée sur un écran tactile, à travers une succession

¹⁶ Dans une publication ultérieure, Beckers, Berwick, Okanoya, et Bolhuis (2016) offrent un autre type d'analyse de ces similarités acoustiques entre les différents types de séquences. Ces nouvelles analyses laissent inchangées leurs conclusions.

d'étapes de difficulté croissante. Lors de la phase test¹⁷, deux paires étaient présentées au cours d'un même essai (e.g. A_1B_1 et A_2B_2). Chaque essai de cette phase était divisé en trois affichages successifs, chacun requérant la sélection d'un stimulus par le babouin (cf. Figure 18a) : (1) le premier élément d'une paire (e.g. A_1) et un élément non apparié (distracteur), (2) A_2 et distracteur, et enfin, (3) B_1 , B_2 et B_3 . Lors du dernier affichage, les babouins devaient sélectionner deux éléments sur les trois disponibles, *sans contrainte d'ordre*.

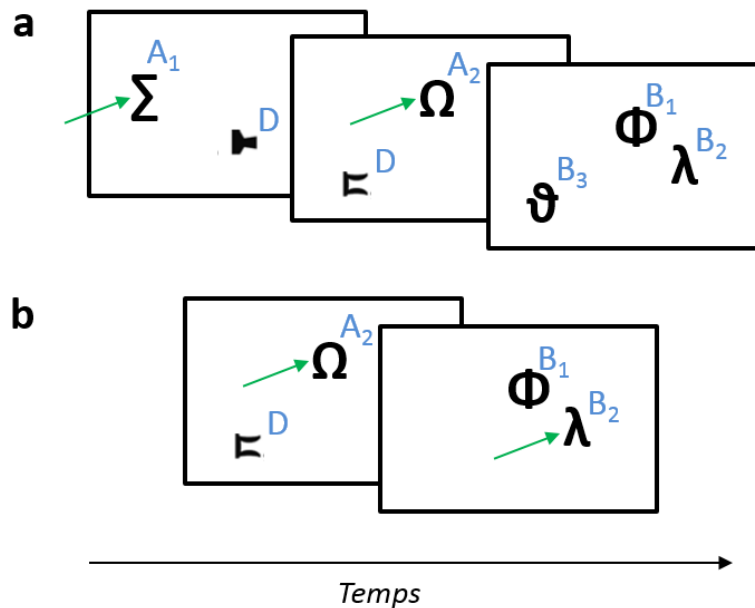


Figure 18 Illustration de la procédure expérimentale employée par Rey et al. (2012). **a.** Exemple d'un essai de la phase test. **b.** Exemple d'un essai de la dernière phase d'entraînement précédent le test. Les caractères en bleu indiquent le type de stimuli (A_{1-6} : premiers éléments des paires, B_{1-6} : deuxièmes éléments des paires, D : distracteurs). Les flèches vertes indiquent la réponse correcte pour chaque affichage. Dans le dernier affichage de la phase test, le participant devait successivement sélectionner deux stimuli sur les trois disponibles, sans contrainte d'ordre. Noter que les types de stimuli et les flèches sont représentés ici pour illustrer le fonctionnement de la tâche, mais n'apparaissent pas lors de l'expérience. L'emplacement des stimuli sur l'écran variait à chaque essai.

Le pattern de réponse des singes, en termes de proportion de séquences produites, est le suivant : $A_1A_2B_2B_1 > A_1A_2B_2B_3 > A_1A_2B_1B_2 > \text{autres}$. Les babouins auraient donc sélectionné en priorité l'élément apparié au dernier élément vu (B_2), et préféré ensuite sélectionner l'élément apparié au premier élément de la séquence (B_1), plutôt qu'un élément n'étant apparié à aucun des éléments de la séquence (B_3), produisant ainsi des séquences centré-emboîtées plus souvent que tout autre type de séquences. Les auteurs tirent deux conclusions de ces résultats.

¹⁷ Nous ne présentons ici que la procédure du test le plus complet, le Test 2.

Premièrement, les babouins sont capables de maintenir en mémoire de travail le premier élément d'une paire et de traiter une autre paire, avant de compléter la première. En effet, la proportion plus importante de sélection de B_1 que de B_3 (après une sélection de B_2) indique que les babouins parviennent à maintenir en mémoire l'élément A_1 , au cours du traitement de la paire intermédiaire A_2B_2 . Deuxièmement, dans cette tâche les babouins auraient spontanément produit des séquences à dépendances centré-emboîtées. Cette dernière conclusion semble toutefois être sujette à caution. En effet, la dernière étape de l'entraînement précédant le test consistait à sélectionner, dans une série de deux affichages sur un écran tactile : (1), le premier élément d'une paire (e.g. A_2) en évitant le distracteur, puis (2) le deuxième élément de la paire (B_2) en évitant le deuxième élément d'une autre paire (e.g. B_1 ; cf. Figure 18b). Ainsi, la sélection préférentielle de l'élément de la paire venant d'être vue pourrait avoir été fortement induite par les contingences de renforcement appliquées dans cette dernière phase d'entraînement.¹⁸ La conduite d'un test supplémentaire impliquant de nouvelles paires d'éléments et de nouveaux distracteurs, non soumis au même entraînement, permettrait d'écarter cette hypothèse.

II.2.3.4 Conclusions

Aucune des expériences menées à ce jour ne semble donc pouvoir de résultats non ambigus sur la capacité d'une espèce NH à traiter une grammaire supra-régulière. Nous avons vu que l'emploi de certains tests (extension, A^nB^m , violations du caractère centré-emboîté des dépendances) est indispensable pour répondre à cette question, mais la conduite de tests supplémentaires en cas d'échec est également informative. Couplés avec une analyse des stratégies individuelles, ces tests apportent ainsi des éléments cruciaux pour une meilleure compréhension du comportement adopté par les animaux testés dans ces expériences d'AGL (ainsi que par des participants humains, e.g. Zimmerer, Cowell, & Varley, 2011). Ils suggèrent notamment que si le traitement de la grammaire A^nB^n ne semble pas constituer une stratégie adoptée par les différentes espèces NH testées, le traitement de la grammaire $(AB)^n$ ne correspond pas non plus à une stratégie privilégiée. Ces résultats poussent à revoir les conclusions de l'expérience fondatrice de Hauser et Fitch (2004) concernant les résultats positifs obtenus chez le tamarin avec cette dernière grammaire. Ces tests montrent également que les oiseaux testés ont été particulièrement enclins dans ces paradigmes AGL à se baser sur

¹⁸ Pour une critique complémentaire, voir Poletiek, Fitz, et Bocanegra (2016).

un apprentissage de régularités locales (e.g. bigrammes), et de surface (i.e. non généralisables à de nouveaux éléments), plutôt que de traiter les régularités structurelles sous-jacentes.

Mais comme le soulignent van Heijningen et al. (2009), on peut considérer qu'il est effectivement pertinent d'utiliser une règle de décision simple lorsqu'elle suffit. Or, dans l'ensemble des tâches de conditionnement opérant passées en revue ici, l'utilisation de ces stratégies permettait effectivement de réussir la tâche dans une grande majorité des essais : la plupart des animaux testés montrent des niveaux de performance très élevés dans les essais d'entraînement (qui sont les seuls essais renforcés, voir la section II.2.1). Par conséquent, l'absence de résultats clairement positifs dans ces expériences ne nous permet ni de conclure à la capacité des espèces animales NH de traiter des grammaires supra-régulières, ni de conclure à une incapacité, du fait de la présence de régularités plus simples dans les matériels expérimentaux utilisés.

Résumé

- **Sous certaines conditions, les humains parviennent à traiter des grammaires supra-régulières dans des expériences d'AGL.**
- **Aucune donnée ne permet cette conclusion chez l'animal non-humain.**
- **En raison de nombreux problèmes méthodologiques, il est toutefois également difficile de conclure sur la base de ces travaux que ces animaux en sont incapables.**
- **Participants humains et non-humains ont tendance à exploiter des régularités plus simples (e.g. détection de répétitions) lorsqu'ils en ont la possibilité, plutôt que de traiter la grammaire sous-jacente.**

II.2.4 Corrélats neuronaux

II.2.4.1 Chez l'humain

Un ensemble convergent de données indique que le traitement de la syntaxe chez l'humain implique un réseau temporo-frontal, ainsi que des régions pariétales et sous-corticales (pour une revue, voir Friederici, 2011). Ce réseau est principalement latéralisé à gauche, mais certaines de ces régions sont également impliquées dans l'hémisphère droit. Les deux régions clés de ce réseau sont l'aire de Broca, dans le lobe préfrontal, et les régions supérieures du lobe temporal (notamment le sillon temporal supérieur, ci-après STS ; voir Encadré 2 et Figure 19).

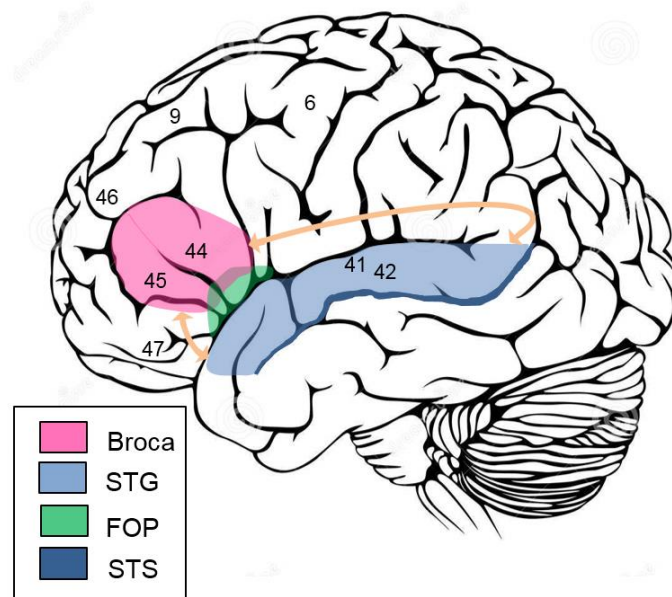


Figure 19 Localisation des principales régions du cerveau humain impliquées dans les traitements de la syntaxe mentionnées dans cette section. Vue de l'hémisphère gauche. Les flèches illustrent les deux voies principales, ventrales et dorsales, connectant les régions frontales et temporales impliquées dans le réseau. Les numéros indiquent les aires de Brodmann. STG = gyrus temporal supérieur, FOP = opercule frontale, STS = sillon temporal supérieur. Voir également Encadré 2.

Une première méthode employée pour étudier les corrélats neuronaux du traitement de la syntaxe est la manipulation orthogonale de l'information sémantique (e.g. mots *vs* pseudo-mots) et syntaxique (e.g. phrases *versus* listes). En prenant soin de manipuler séparément congruence sémantique (i.e. la possibilité de dériver un sens) et structure syntaxique, Humphries, Binder, Medler, & Liebenthal (2006) ont par exemple montré qu'une région du lobe temporal gauche (la partie antérieure du sulcus temporal supérieur, aSTS) s'active davantage en réponse au traitement d'une phrase grammaticalement correcte, par rapport au traitement d'une liste de mots. En faisant varier systématiquement la taille des syntagmes d'une séquence de longueur fixe (une phrase de 12 mots, deux phrases de 6 mots, etc., jusqu'à une liste de 12 mots isolés), Pallier, Devauchelle, et Dehaene (2011) montrent que la partie postérieure du STS, ainsi que le gyrus frontal inférieur (IFG) sont activés proportionnellement à la taille des syntagmes traités (i.e. le nombre de mots qu'ils contiennent). Les résultats de cette expérience indiquent également que pour ces deux régions les activations sont similaires. Cette

similarité s'observe que les participants aient à traiter des phrases naturelles, ou des séquences construites en « Jabberwocky »¹⁹, ce qui suggère que ces deux régions sont impliquées dans le traitement syntaxique plutôt que sémantique.

Une seconde étude de ce type montre que cette augmentation paramétrique du niveau d'activation au cours du traitement successif des mots composant le syntagme s'accompagne d'une soudaine baisse de l'activité lors de sa complétion (e.g. à la frontière entre le syntagme nominal et le syntagme verbal ; Nelson et al., 2017). On observe également un effet de la complexité syntaxique sur le degré d'activation de ce réseau. L'aire de Broca présente par exemple des activations plus importantes lors du traitement d'une phrase présentée dans un ordre non canonique (e.g. Friederici, Fiebach, Schleewsky, Bornkessel, & von Cramon, 2006; Santi & Grodzinsky, 2007), ou comportant une proposition relative (Makuuchi, Bahlmann, Anwander, & Friederici, 2009), que lors du traitement d'une phrase simple.

Des patterns d'activation similaires ont été observés dans des paradigmes d'AGL, lors du traitement d'une grammaire régulière complexe (Forkstam, Hagoort, Fernandez, Ingvar, & Petersson, 2006; Petersson, Forkstam, & Ingvar, 2004; Petersson, Folia, & Hagoort, 2012). En utilisant la stimulation magnétique transcrânienne (TMS), Uddén et al. (2008) ont confirmé le rôle causal joué par l'aire de Broca dans ces traitements. Ces recherches indiquent également que l'aire de Broca est impliquée dans le traitement de grammaires supra-régulières, impliquant de compter et comparer un certain nombre de syllabes (e.g. Friederici, Bahlmann, et al., 2006), ou de vérifier l'organisation centrée-emboîtée de dépendances non-adjacentes (e.g. Bahlmann

Encadré 2 Principales régions cérébrales impliquées dans le traitement de la syntaxe

Aires auditives primaires :
dans le lobe temporal,
comprend BA 41/42

Aire de Broca : Correspond aux
BA 44/45

DLPFC : Cortex préfrontal
dorso-latéral, comprend BA
9/46.

FOP : opercule frontale

vFOC : cortex frontaux et
operculaires ventraux

IFG : Gyrus frontal inférieur,
comprend l'aire de Broca, et
d'autres régions plus
antérieures (BA 47)

STG : Gyrus temporal supérieur

STS : Sillon temporal supérieur

VLPFC : Cortex préfrontal
ventro-latéral, comprend BA
47/45/44

BA = Aire de Brodmann.
Par souci de clarté, les
acronymes anglais sont utilisés.
Voir également la Figure 19.

¹⁹ Procédé expérimental consistant à remplacer dans une phrase les mots contenu (e.g. noms, verbes) par des pseudo-mots, tout en conservant les mots fonction (e.g. déterminants) et, généralement, les marqueurs morphologiques des différentes catégories syntaxiques (e.g. terminaisons et accords).

et al., 2008). De plus, ce réseau, et notamment l'aire de Broca, ne sous-tendrait pas uniquement l'extraction de régularités linguistiques. Il est également impliqué dans des expériences AGL présentant (1) des séquences de stimuli visuels (e.g. Bahlmann, Schubotz, Mueller, Koester, & Friederici, 2009), (2) des séquences musicales (e.g. Maess, Koelsch, Gunter, & Friederici, 2001), ou encore (3) lors de la production de séquences d'actions hiérarchiquement organisées (Greenfield, 1991; Koechlin & Jubault, 2006; Roy & Arbib, 2005). L'ensemble de ces données suggère donc que l'aire de Broca soit impliquée dans le traitement de toute séquence d'événements hiérarchiquement organisés (Fiebach & Schubotz, 2006).

De nombreux auteurs s'accordent sur cette idée, mais les propositions divergent quant au rôle précis que jouerait l'aire de Broca dans ces traitements. Pour Fitch (2014), cette région supporterait la fonction de mémoire à « pile » (cf. section I.2.2). Cette aire serait donc le lieu de *stockage* temporaire des représentations intermédiaires des sous-unités ou syntagmes composant une séquence, en vue de leur intégration. D'autres auteurs proposent que le rôle joué par l'aire de Broca serait plutôt de supporter les mécanismes permettant *l'intégration* de ces sous-unités en une représentation unifiée de la séquence (e.g. Wilson, Marslen-Wilson, & Petkov, 2017). Un dernier ensemble de modèles théoriques avance que les différentes sous-régions de cette aire supporteraient différentes fonctions. Par exemple, Zaccarella et Friederici (2016) ont proposé qu'une région spécifique de l'aire de Brodmann 44 soit le siège de l'opération *merge* (cf. section I.3.2), leurs expériences ayant rapporté une augmentation de l'activité de cette sous-région lors de l'intégration de syntagmes minimaux composés de deux mots.

II.2.4.2 Données anatomiques comparées

La question de l'existence d'un équivalent de l'aire de Broca chez le primate NH est sujette à controverse. Selon certains auteurs, certains primates NH auraient des régions comparables d'un point de vue architectonique aux aires de Brodmann 44 et 45 (e.g. chez le chimpanzé, Schenker et al., 2010 ; chez le macaque, Petrides & Pandya, 2002; Petrides, Cadoret, & Mackey, 2005).

Comme chez l'humain, ces régions présentent chez le chimpanzé une asymétrie, leur taille ayant tendance à être plus importante dans l'hémisphère gauche que dans l'hémisphère droit (Cantalupo & Hopkins, 2001). Cette asymétrie est corrélée avec les préférences manuelles de ces chimpanzés (mesurée pour des gestes de communication, Meguerditchian, Gardner,

Schapiro, & Hopkins, 2012 ; ainsi que pour des manipulations complexes d'outils impliquant une planification motrice, Hopkins et al., 2017; Hopkins, Russell, & Cantalupo, 2007).

Toutefois, cette région aurait connu une expansion particulièrement importante au cours de l'évolution de la lignée humaine. Des données comparatives suggèrent que sa taille serait environ 32 fois plus importante chez l'humain que chez le macaque, tandis que la taille de la plupart des autres régions du cortex cérébral ne seraient que deux à seize fois plus importante (e.g. Chaplin, Yu, Soares, Gattass, & Rosa, 2013). Des différences importantes entre primates humains et non-humains sont également observées en termes de connectivité structurelle. La voie dorsale connectant les régions temporales et préfrontales impliquées dans le traitement de la syntaxe (voir Figure 19) consiste en un réseau de fibres beaucoup plus important chez l'humain que chez le chimpanzé (Rilling et al., 2008 ; mais voir également Petrides & Pandya, 2009; Romanski & Averbeck, 2009). Des différences de connectivité fonctionnelle entre ces régions sont également observées avec le macaque (Neubert, Mars, Thomas, Sallet, & Rushworth, 2014).

II.2.4.3 Données fonctionnelles comparées

Wilson et collaborateurs (2015) ont effectué une expérience recourant à l'imagerie fonctionnelle (fMRI) chez des macaques rhésus et des humains pour observer les corrélats neuronaux du traitement d'une grammaire régulière relativement complexe impliquant des dépendances entre éléments adjacents (Figure 13). Après une exposition prolongée à des séquences auditives générées à partir de cette grammaire, l'expérimentateur a diffusé des séquences comportant des violations de cette grammaire. Ces violations ont suscité une activité accrue dans des régions similaires chez les participants humains et macaques, dans les régions les plus ventrales des cortex frontaux et operculaires (FOP, voir Figure 19). Les données sont moins claires concernant l'aire de Broca. Chez la majorité des participants humains, aucune augmentation significative de l'activité de la région en lien avec ces violations n'était observée. De manière intéressante, une augmentation de cette activité était cependant observée chez les macaques (mais celle-ci restait plus faible que l'augmentation enregistrée pour les régions ventrales). Les auteurs en concluent que les régions les plus ventrales du cortex frontal jouent un rôle clé dans le traitement des dépendances adjacentes, fonction qui serait évolutivement conservée, tandis que l'implication de l'aire de Broca dans ces traitements est moins claire.

Ces données sont congruentes avec celles obtenues chez l'humain dans des expériences d'AGL examinant les corrélats neuronaux du traitement de la grammaire (AB)ⁿ, qui comporte elle aussi exclusivement des dépendances adjacentes. Une augmentation de l'activité du FOP est observée lors de la diffusion de séquences présentant des violations de cette grammaire (Friederici, Fiebach, et al., 2006). Ces différents résultats suggèrent donc que cette région joue un rôle central dans le traitement de régularités locales (Friederici, 2011). Plus largement, cette région serait également impliquée chez le primate NH dans le traitement de propriétés saillantes des séquences, telles que l'encodage des éléments situés à leurs extrémités (Amiez & Petrides, 2007), ou encore l'encodage et le maintien en mémoire pour une courte période de l'identité et de l'ordre d'apparition de stimuli présentés séquentiellement (Inoue & Mikami, 2006).

L'ensemble de ces données a conduit à l'hypothèse d'un gradient d'activation ventro-dorsal. Le traitement des régularités séquentielles les plus élémentaires est sous-tendu par les régions les plus ventrales du cortex préfrontal ventro-latéral (VLPFC), tandis que le traitement de régularités plus complexes (dépendances non-adjacentes, organisation hiérarchique) serait supporté par les régions plus dorsales du VLPFC, dont l'aire de Broca (voir Figure 20 ; Friederici, 2011; Wilson et al., 2017). Les différents modèles de ce type postulent que le rôle joué par ces zones les plus ventrales du VLPFC dans les traitements séquentiels aurait été déjà présent *a minima* chez notre dernier ancêtre commun avec les macaques, et conservé au cours de l'évolution des primates. Mais leurs propositions divergent concernant l'histoire évolutive du rôle joué par les régions plus dorsales. Friederici et collaborateurs (e.g. Friederici, Bahlmann, et al., 2006) avancent que ces dernières régions sont impliquées dans le traitement de régularités plus complexes, mais seulement chez l'humain. De leur côté, Wilson et collaborateurs (2017) proposent que cette fonction soit également apparue précocement au cours de l'évolution des primates, et ait été conservée au cours de la phylogenèse. Leur modèle prédit donc que des activations plus dorsales devraient également être observées chez le primate NH lors du traitement de séquences comprenant des dépendances non-adjacentes et/ou organisées de manière hiérarchique. Ils font cependant l'hypothèse que cette homologie fonctionnelle pourrait être accompagnée de différences importantes entre les cerveaux des primates humains et NH, telles qu'un degré de latéralisation plus important ou une implication particulière des aires de Brodmann 44 et 45 pour les premiers, ou encore des différences de connectivité entre ces régions et le reste du réseau impliqué dans le traitement de régularités séquentielles.

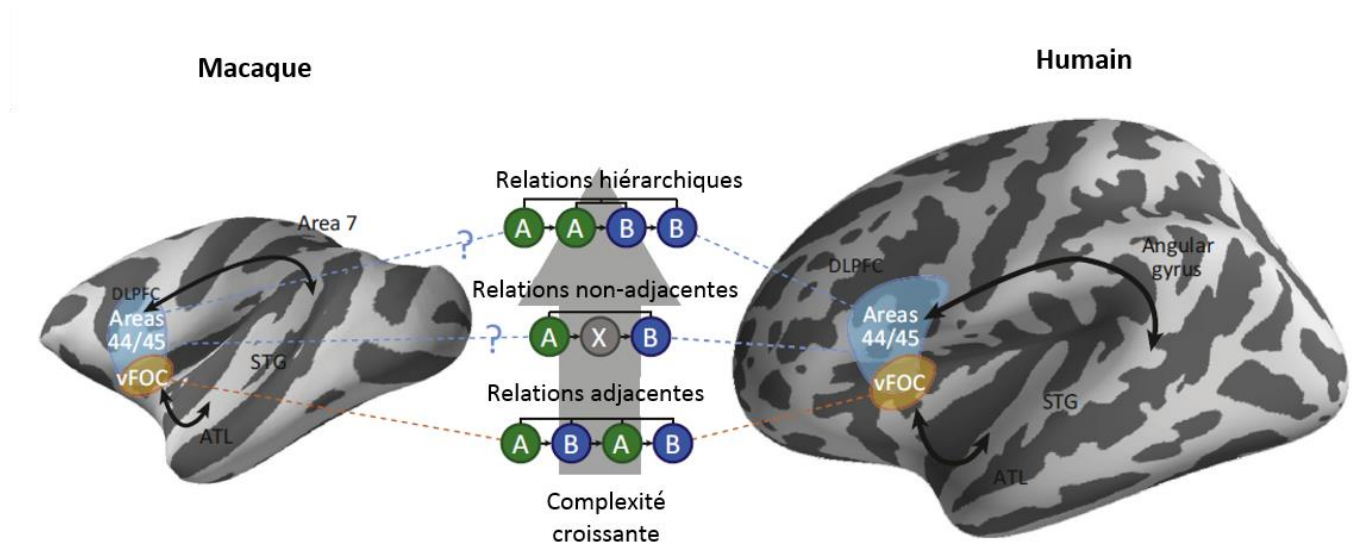


Figure 20 Illustration du modèle d'un gradient ventro-dorsal pour le traitement de régularités séquentielles de complexité croissante chez les primates. vFOC = cortex frontaux et operculaires ventraux. Modifié d'après Wilson et al. (2017).

Aucune donnée expérimentale ne nous permet cependant de discuter ces deux hypothèses à l'heure actuelle. A notre connaissance, le traitement de dépendances non-adjacentes n'a fait l'objet d'aucun travail comparatif en imagerie cérébrale (cependant, voir la section V.2.2 pour une étude d'électroencéphalographie). Et comme nous l'avons vu dans la section II.2.3.3, les paradigmes expérimentaux élaborés jusqu'à présent pour étudier le traitement de grammaires supra-régulières chez l'animal NH ne sont pas satisfaisants. Répondre à ces questions nécessiterait donc comme préalable la conception de tâches plus adaptées.

Résumé

- Une homologie fonctionnelle chez les primates humains et NH est observée pour les régions ventrales du cortex préfrontal ventrolatéral, qui supporteraient le traitement de régularités séquentielles simples (e.g. dépendances adjacentes).
- Chez l'humain, l'aire de Broca joue un rôle clé dans l'extraction de régularités syntaxiques plus complexes.
- A ce jour, l'implication ou non d'une région équivalente dans ce type de traitement chez le primate NH n'est pas documentée.

III. Conclusion : Proposition d'une Nouvelle Approche pour Etudier le Traitement des Grammaires Supra-régulières

Un examen de la complexité syntaxique des langues naturelles, d'une part, et des systèmes de communication des autres espèces, d'autre part, révèle des différences fondamentales. Toutefois, les données passées en revue dans la section précédente mettent également en lumière un certain nombre de similarités entre les capacités syntaxiques des humains et des autres espèces étudiées :

(1) Concernant les aspects de production, humains et non-humains présentent une certaine diversité de comportements qui semblent impliquer des capacités syntaxiques relativement complexes, y compris en dehors du domaine de la communication, tel que pour la production de séquences complexes d'actions.

(2) Dans les paradigmes d'AGL, participants humains et non-humains manifestent une tendance à se baser sur des régularités plus élémentaires que le traitement d'une grammaire A^nB^n lorsque le matériel expérimental employé le leur permet (détection de fragments d'exemplaires familiers, de répétitions d'éléments, etc.).

(3) On observe une homologie fonctionnelle chez les primates humains et non-humains pour certaines régions cérébrales impliquées dans le traitement de régularités séquentielles simples.

Ces similarités nous font suggérer qu'il existe, malgré les différences observées, un important recouvrement entre les systèmes de traitement des régularités séquentielles des humains et des NH.

De plus, nous avons vu que la principale différence qui émerge des données issues des expériences d'AGL, à savoir l'absence de données clairement positives sur le traitement de régularités supra-régulières chez une espèce NH, nous semble pouvoir être imputée à des défauts des procédures expérimentales employées. Enfin, dans ces mêmes expériences, le traitement de la grammaire régulière $(AB)^n$ ne constitue pas non plus une stratégie privilégiée par les différentes espèces NH testées. Ces données suggèrent que la distinction entre grammaires régulières et supra-régulières ne rend pas compte suffisamment finement des capacités syntaxiques des différentes espèces et de leurs limites (Pettersson et al., 2012; Petkov

& Wilson, 2012; Rogers & Pullum, 2011). Le fait qu'une somme relativement importante de travaux comparatifs aient été menés sur la base de cette distinction sans qu'aucune conclusion claire n'en émerge suggère que cette méthodologie n'est pas adaptée.

Dans la présente thèse, nous avons par conséquent poursuivi une stratégie alternative, consistant à décomposer le traitement des grammaires supra-régulières en opérations élémentaires. Selon cette perspective, nous faisons l'hypothèse que les échecs obtenus jusqu'à présent dans les expériences cherchant à tester la capacité des animaux NH à traiter des grammaires supra-régulières pourraient davantage découler de difficultés à effectuer certaines des opérations cognitives impliquées dans le traitement de ce type de grammaires, plutôt que d'une incapacité à accéder à ce type de grammaires *per se*. La maîtrise d'une grammaire supra-régulière implique de traiter des dépendances non-adjacentes, entre deux ensemble d'éléments (dans le cas de la grammaire A^nB^n), ou entre des items multiples (dans le cas des grammaires centrée-emboîtée et croisée). Par ailleurs, cela implique d'être sensible à la structure sous-jacente des séquences, et d'être capable de généraliser cette structure à de nouveaux éléments.

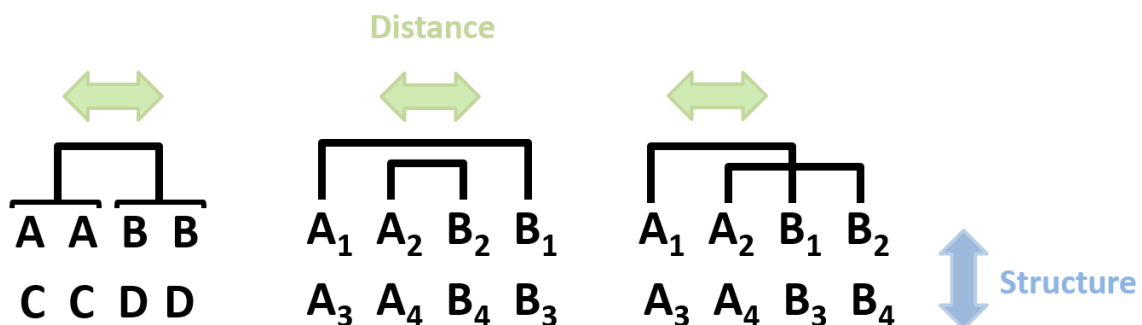


Figure 21 Illustration des deux prérequis cognitifs pour l'extraction de grammaires supra-régulières (A^nB^n , centrée-emboîtée et croisée) : la capacité à apprendre des relations non locales, et la sensibilité à la structure sous-jacente.

Nous identifions donc deux prérequis cognitifs pour le traitement de grammaires supra-régulières : (1) la capacité à apprendre des dépendances non-adjacentes, et (2) la capacité à extraire les propriétés structurelles des séquences (Figure 21). Il existe une littérature comparative sur ces deux aspects, que nous présentons dans le chapitre suivant. Nous verrons qu'un examen complet de cette littérature révèle des manques importants dans la caractérisation de l'évolution de ces deux habiletés cognitives. Dans la partie expérimentale de cette thèse, nous avons donc cherché à combler les manques identifiés afin de tester notre hypothèse, à l'aide d'une série de quatre expériences menées chez le primate humain et non-humain.

Chapitre II : Prérequis Cognitifs pour le Traitement de Grammaires Supra-régulières

IV. Apprentissage d'Associations Non-adjacentes

Cette dénomination renvoie à des dépendances entre deux éléments étant séparés par un ou plusieurs autres éléments dans une séquence (également appelées dépendances « à longue distance », ou « disjointes »). On parle d'associations ou de « dépendances » car le premier élément prédit l'apparition du deuxième, avec une probabilité plus ou moins importante (e.g. une probabilité transitionnelle de 1, dans le cas d'une dépendance déterministe). Il y est souvent fait référence par la notation AXB, où A et B sont deux éléments associés tandis que X est un élément variable statistiquement indépendant à la fois de A et de B (i.e. l'élément X ne prédit pas l'apparition de B).

Dans les langues naturelles, on trouve de nombreuses dépendances non-adjacentes, à plusieurs niveaux. Au niveau morphosyntaxique tout d'abord, on retrouve par exemple des dépendances non-adjacentes pour l'accord en genre et en nombre du groupe nominal (e.g. les babouins), ainsi que pour l'accord du verbe (e.g. nous mangeons). Dans ces exemples, une modification de l'un des deux éléments dépendants donne une phrase agrammaticale (e.g. nous mangez, vous mangeons). Il existe également des dépendances non-adjacentes au niveau phonologique. L'harmonie vocalique par exemple, un phénomène observé dans certaines langues comme le turc, implique des dépendances non-adjacentes entre les voyelles d'une même unité (mot ou syntagme), celles-ci devant être « compatibles » entre-elles selon certaines règles spécifiques à chaque langue. On retrouve également des dépendances non-adjacentes entre consonnes. En français par exemple, les séquences consonne-voyelle-consonne au sein d'un mot consistent plus fréquemment en une consonne labiale (e.g. /p/, /b/) suivie d'une coronale (e.g. /d/, /t/), que l'inverse (Nespor, Peña, & Mehler, 2003). Enfin, des dépendances non-adjacentes sont bien entendu également entretenues entre des syntagmes, dans des phrases comportant par exemple des propositions relatives (voir Figure 5, p. 23).

La capacité à détecter des dépendances non-adjacentes est aussi probablement essentielle dans de nombreux autres domaines, comme le soulignent Turk-Browne, Jungé, et Scholl (2005): “*People are constantly bombarded with noise in space and time that needs to be segregated in order to extract a coherent representation of the world, and people rarely encounter a sequence of relevant stimuli without any interruptions.*”.

IV.1 Données chez l’humain

IV.1.1 Méthodes expérimentales

L’apprentissage de dépendances non-adjacentes chez l’humain est étudié de deux façons : en étudiant leur traitement dans les langues naturelles, et en effectuant des expériences d’apprentissage de langages artificiels (*Artificial Language Learning*, ou ALL). Ce paradigme a été initié par Saffran, Aslin, et Newport (1996), pour étudier la segmentation du flux de parole sur la base des probabilités transitionnelles entre éléments adjacents. Son principe est proche des tâches d’AGL : les participants sont exposés, le plus souvent de manière passive, à des séquences d’items dépourvus de sens (syllabes, sons, etc.). Ils sont ensuite testés sur de nouvelles séquences, congruentes ou non avec les séquences familières. Mais à la différence des paradigmes AGL, ici ce n’est pas la capacité des participants à extraire une grammaire sous-jacente qui est évaluée, mais leur sensibilité aux régularités distributionnelles (ou statistiques) présentes dans ces séquences. On peut distinguer deux façons d’étudier l’apprentissage de dépendances non-adjacentes dans des paradigmes ALL (Encadré 3). La première est d’évaluer la capacité des participants à segmenter un flux continu d’éléments (e.g. syllabes) sur la base des probabilités transitionnelles entre éléments non-adjacents (Newport & Aslin, 2004). La seconde consiste à présenter aux participants les dépendances non-adjacentes au sein de courtes séquences explicitement segmentées, et à tester leur capacité à détecter des violations de ces dépendances (Gómez, 2002).

Encadré 3 Les deux principaux paradigmes d'apprentissage de langues artificielles (ALL) pour étudier l'apprentissage de dépendances non-adjacentes

	Familiarisation	Test	
		Congruent	Non-congruent
Structure	$A_1 X_1 B_1 A_2 X_1 B_2 A_1 X_2 B_1$		
Tâche de segmentation ¹	<u>di gu</u> <u>tae ke</u> <u>gu bu</u> di ki tae	di gu tae	tae ke gu
Tâche de détection de violations ²	<u>pel wadim rud</u> -- vot wadim <u>jic</u> -- pel kicey rud	pel wadim rud	pel wadim jic

Les exemples de stimuli sont tirés de ¹Newport, Hauser, Spaepen, & Aslin (2004), et ²Gómez (2002). Les segments soulignés illustrent la manière dont sont construites les séquences présentées congruentes (en vert) et non-congruentes (en rouge) lors du test, pour les deux types de tâches. Les tirets (--) signalent des pauses.

Le principe général de ces tâches est de présenter des séquences de stimuli, composées de triplets (appelés « mots ») de la forme AXB. A et B entretiennent une dépendance déterministe, tandis que les probabilités entre éléments adjacents ne sont pas informatives (i.e. A ne permet pas de prédire X, et X ne permet pas de prédire B). À la suite de cette période d'exposition, les participants sont testés sur des séquences congruentes et non congruentes.

Dans les tâches de segmentation, les éléments (ici, des syllabes) sont diffusés les uns à la suite des autres, sans que les frontières entre les triplets soient indiquées par des indices acoustiques (pas de pause, ni d'indication prosodique). Les probabilités transitionnelles intra- et inter-triplet entre éléments adjacents étant équilibrées, seules les probabilités transitionnelles entre éléments non-adjacents permettent d'effectuer la tâche de segmentation. Dans la phase de test, des triplets congruents et non-congruents sont présentés aux participants. Ces derniers sont constitués du dernier élément d'un mot, suivi des deux premiers d'un autre mot (e.g. $B_1 A_2 X_1$). Si les participants ont été capables d'effectuer la tâche de segmentation, ils les identifient comme n'étant pas des mots du langage auquel ils ont été exposés.

Dans les tâches de détection de violation, des pauses indiquent explicitement les frontières entre les triplets. Les triplets non-congruents présentés lors du test ne comportent pas de violation des dépendances adjacentes, mais impliquent une violation des dépendances non-adjacentes (e.g. $A_1 X_1 B_2$).

IV.1.2 Ontogénie

La capacité de détecter les dépendances non-adjacentes présentes au niveau morphosyntaxique²⁰ dans les langues naturelles n'apparaît chez l'enfant qu'au cours de la

²⁰ La sensibilité aux dépendances non-adjacentes présentes au niveau phonologique, elle, apparaît plus tôt, au cours de la première année : des violations de l'harmonie vocalique sont détectées dès six mois par des enfants turcs (van Kampen, Parmaksiz, Van de Vijver, & Hoehle, 2008), et dès dix mois lorsque les violations concernent des consonnes (Gonzalez-Gomez & Nazzi, 2012).

deuxième année (pour une revue, voir Sandoval & Gómez, 2013). Nazzi, Barrière, Goyet, Kresh, et Legendre (2011) ont montré, par exemple, que des enfants français de 17-18 mois réagissent à des violations de l'accord singulier/pluriel des verbes (voir également van Heugten & Shi, 2010). A 18 mois, des enfants anglais détectent des violations de la dépendance entre l'auxiliaire (*is*) et la terminaison du verbe (*-ing*) formant le présent continu, tandis qu'ils n'y réagissent pas à 15 mois (Santelmann & Jusczyk, 1998). Des enfants néerlandais reconnaissent une dépendance entre un article défini et un suffixe diminutif à 24, mais pas à 17 mois (van Heugten & Johnson, 2010). Ces études montrent également que plusieurs facteurs ont une incidence sur cette capacité : les violations sont détectées seulement si le nombre de syllabes séparant les éléments dépendants est relativement faible (Hohle, Schmitz, Santelmann, & Weissenborn, 2006; Santelmann & Jusczyk, 1998; Soderstrom, White, Conwell, & Morgan, 2007). De plus, toutes les dépendances non-adjacentes d'une langue ne sont pas acquises au même moment. Une analyse de corpus a par exemple montré que les dépendances acquises le plus tôt au cours du développement étaient celles apparaissant le plus fréquemment dans les phrases entendues par les enfants néerlandais (van Heugten & Johnson, 2010).

Ces données convergent avec celles issues d'expériences d'ALL : une étude de Gómez et Maye (2005) montre que la capacité à extraire des dépendances non-adjacentes implémentées dans un langage artificiel apparaissait entre 12 et 15 mois²¹. Pris dans leur ensemble, ces résultats révèlent une différence de trajectoire développementale entre la capacité à extraire des dépendances adjacentes et celle à extraire des dépendances non-adjacentes, cette première étant observée dès la première année (e.g. Saffran et al., 1996; Saffran, Johnson, Aslin, & Newport, 1999).

IV.1.3 Apprentissage de dépendances non-adjacentes chez l'adulte

Un ensemble d'expériences d'ALL indique que chez l'adulte, l'extraction de dépendances non-adjacentes est également plus difficile que celui de dépendances adjacentes (e.g. Cleeremans & McClelland, 1991; Gebhart, Newport, & Aslin, 2009; Newport & Aslin, 2004; Pacton, Sobaco, & Perruchet, 2015; Wilson, Smith, & Petkov, 2015). En effet, tandis que l'extraction de dépendances adjacentes ne présente aucune difficulté particulière (e.g. Fiser & Aslin, 2002; Hunt & Aslin, 2001; Saffran et al., 1999; Saffran, Newport, & Aslin, 1996),

²¹ A noter toutefois que des données EEG suggèrent une sensibilité aux dépendances non-adjacentes beaucoup plus précoce (dès trois mois, Mueller, Friederici, & Männel, 2012).

l'extraction de dépendances non-adjacentes ne s'observe que sous certaines conditions : (1) lorsque les éléments dépendants ont des caractéristiques perceptives communes tandis que l'élément intermédiaire diffère (e.g. dépendances entre consonnes présentées avec une voyelle intermédiaire, et *vice-versa*, Newport & Aslin, 2004 ; voir également Creel, Newport, & Aslin, 2004; Gebhart et al., 2009; Onnis, Monaghan, Richmond, & Chater, 2005), (2) lorsque l'élément intermédiaire est très variable (chez l'adulte comme chez l'enfant : Gómez, 2002; voir également Onnis, Monaghan, Christiansen, & Chater, 2005) ou qu'il est invariable (Onnis et al., 2005), ou encore (3) lorsque les éléments dépendants sont d'abord présentés de manière adjacente (chez l'adulte : Lai & Poletiek, 2011 ; et chez l'enfant : Lany & Gómez, 2008).

Un autre type d'étude s'est penché sur la dynamique temporelle de cet apprentissage, en employant des paradigmes adaptés à partir de la tâche de pointage sériel (SRT, Nissen & Bullemer, 1987) permettant un enregistrement continu de l'apprentissage des participants au cours de l'exposition aux régularités. Ces travaux ont notamment cherché à savoir si lorsqu'un matériel comprenant à la fois des dépendances adjacentes et non-adjacentes était présenté aux participants, ceux-ci pouvaient les apprendre simultanément. Leur résultats apportent une réponse positive à cette question, du moins lorsque ces deux types de régularités sont instanciées par des vocabulaires différents (i.e. les éléments entretenant des dépendances non-adjacentes n'entretiennent pas de dépendances adjacentes, et *vice-versa* (Romberg & Saffran, 2013a; Vuong, Meyer, & Christiansen, 2016). Il est toutefois observé une différence intéressante dans le pattern de résultats obtenu : tandis que l'ensemble des participants parviennent à traiter les dépendances adjacentes présentes dans le matériel expérimental, leurs performances avec les dépendances non-adjacentes suivent une distribution bimodale : seule une partie des participants parviennent à extraire ces dernières (voir également Misyak & Christiansen, 2012; Mueller, Friederici, & Männel, 2012; Romberg & Saffran, 2013; Wilson, Smith, et al., 2015).

L'ensemble de ces données suggère que l'apprentissage de dépendances non-adjacentes est plus difficile chez l'humain que l'extraction de dépendances adjacentes : il est observé plus tardivement au cours du développement, et seulement sous certaines conditions.

IV.2 Données comparatives

IV.2.1 Apprentissage de dépendances adjacentes

De nombreuses espèces sont sensibles aux dépendances adjacentes (pour une revue : ten Cate & Okanoya, 2012; Conway & Christiansen, 2001; Wilson, Marslen-Wilson, & Petkov, 2017). Cette capacité a été observée chez le tamarin pinché (*Saguinus oedipus*, Hauser, Newport, & Aslin, 2001), le macaque rhésus (*Macaca Mulatta*, Wilson, Smith, et al., 2015), le babouin de Guinée (*Papio papio*, Minier, Fagot, & Rey, 2016), le rat (*Rattus norvegicus*, Toro & Trobalón, 2005), le pigeon (*Columba livia domestica* Froehlich, Herbranson, Loper, Wood, & Shimp, 2004), et deux espèces d'oiseaux chanteurs (*Lonchura striata var. domestica*, Takahasi, Yamada, & Okanoya, 2010 ; *Taeniopygia guttata*, Chen & ten Cate, 2015). Les patterns de résultats obtenus chez ces différentes espèces sont très similaires à ceux obtenus chez l'humains²². Par exemple, testés dans une tâche de pointage sériel, pigeons (Froehlich et al., 2004) et humains (Hunt & Aslin, 2001) présentent une corrélation comparable entre leurs temps de réponse et le degré de prédictibilité de l'élément à venir dans la séquence : leurs réponses étaient d'autant plus rapides que les probabilités transitionnelles entre éléments adjacents étaient importantes (Figure 22).

²² Mais des différences peuvent également être observées. Par exemple, Toro & Trobalón (2005) ont montré que les rats se basaient sur des computations statistiques plus simples (fréquence des co-occurrences) que celles effectuées par des enfants de huit mois (probabilités transitionnelles) testés dans une tâche similaire (Aslin, Saffran, & Newport, 1998).

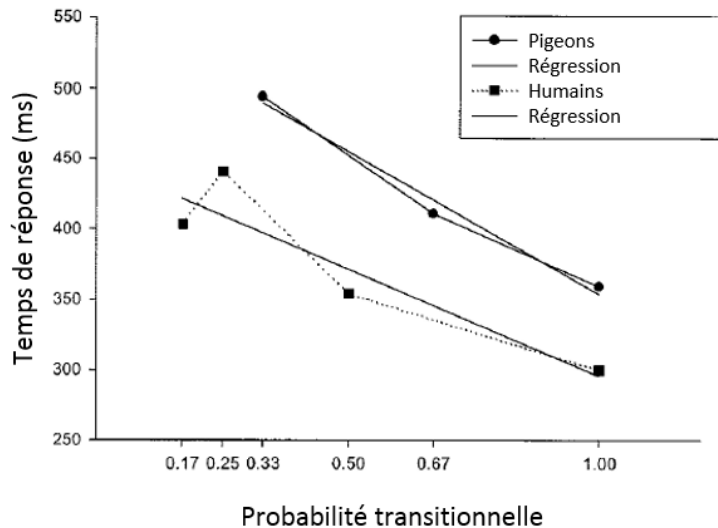


Figure 22 Temps de réponse obtenus chez le pigeon (Froehlich et al., 2004) et l'humain (Hunt & Aslin, 2001) en fonction des probabilités transitionnelles entre éléments adjacents, dans une tâche de pointage sériel. Les pigeons répondent globalement plus lentement que les humains, mais leur pente de régression est comparable. Modifié d'après Froehlich et al. (2004).

La tâche de pointage sériel a également été utilisée pour une étude menée chez le babouin et l'humain au sein de notre laboratoire (Minier et al., 2016 ; Rey et al., 2018). Les participants devaient suivre une cible se déplaçant sur un écran tactile. Une séquence de neuf cibles était présentée à chaque essai, en réalité constituée de trois triplets présentés dans un ordre aléatoire. Aucun indice ne signalait toutefois les frontières entre ces triplets, mis à part les probabilités transitionnelles (qui étaient égales à 1 au sein des triplets). Non seulement les participants des deux espèces sont parvenus à segmenter les séquences, mais ils présentaient également les mêmes dynamiques temporelles. Leurs temps de réponse sur le dernier élément des triplets décroissaient plus rapidement au cours du temps que leurs temps de réponse sur le second élément. Ces résultats suggèrent un effet cumulatif de l'information contextuelle des éléments précédents pour la prédiction de l'élément à venir dans la séquence, chez ces deux espèces.

L'ensemble de ces données suggère que la sensibilité aux dépendances adjacentes est largement répandue au sein de la phylogénèse, et que cette capacité est sous-tendue par des mécanismes communs aux humains et aux autres espèces étudiées.

IV.2.2 Apprentissage de dépendances non-adjacentes

IV.2.2.1 Données comportementales

Un premier type d'expériences menées chez l'animal NH implique des dépendances entre éléments non-adjacents étant perceptivement plus proches l'un de l'autre que de l'élément intermédiaire (sur le même modèle que Creel et al., 2004, et Newport & Aslin, 2004, chez l'humain). Des résultats positifs ont été obtenus en utilisant ce type de matériel chez deux espèces de primates (*Saguinus oedipus*, Newport, Hauser, Spaepen, & Aslin, 2004 ; *Saimiri sciureus*, Ravignani, Sonnweber, Stobbe, & Fitch, 2013). De plus, les rats parviennent à détecter des dépendances entre des consonnes présentées avec une voyelle intermédiaire, et *vice-versa* (de la Mora & Toro, 2013), tandis qu'ils échouent lorsqu'éléments non-adjacents et intermédiaires appartiennent à la même catégorie perceptive (Toro & Trobalón, 2005). Ces données suggèrent que comme chez l'humain, la similarité perceptive entre éléments non-adjacents faciliterait leur association²³.

Un second type d'expériences étudie l'apprentissage de dépendances non-adjacentes en l'absence d'indices perceptifs. Newport et al. (2004) ont étudié la capacité du tamarin (*Saguinus oedipus*) à segmenter un flux continu de syllabes sur la base des probabilités transitionnelles entre syllabes non-adjacentes, dans un paradigme d'habituation-déshabituaton (voir Figure 14). Six « mots » de trois syllabes de la forme AXB (avec trois paires de syllabes A-B différentes, et deux syllabes X) étaient diffusés dans un ordre pseudo-aléatoire pendant plusieurs minutes via un haut-parleur. À la suite de cette phase d'habituation, les triplets de syllabes correspondant à ces « mots » (e.g. A₁XB₁) et des triplets non-congruents (e.g. B₁A₂X ; voir Encadré 2), étaient diffusés un par un, et les réactions des tamarins mesurées. Les tamarins ont davantage réagi (i.e. ils se sont plus fréquemment tournés vers le haut-parleur) lors de la diffusion de ce dernier type de triplets. Ces résultats suggèrent donc que les tamarins ont bien été capables de segmenter un flux continu de syllabes en triplets, sur la seule base des relations déterministes entretenues par les syllabes non-adjacentes.

Sonnweber, Ravignani, et Fitch (2015) ont testé la capacité du chimpanzé à généraliser des dépendances non-adjacentes à de nouveaux éléments intermédiaires. Dans une tâche à choix-forcé, les singes étaient d'abord entraînés à reconnaître cinq paires de stimuli visuels

²³ Voir toutefois de la Mora et Toro (2013), et Toro, Nespors, Mehler, et Bonatti (2008), pour une théorie proposant l'existence de contraintes spécifiques pesant sur l'extraction de régularités linguistiques chez l'humain.

(Figure 23a). Ils devaient ensuite discriminer ces paires de stimuli présentées autour d'un ou deux stimuli intermédiaires (i.e. de la forme AXB ou AXXB), de chaînes de distracteurs de même longueur (XXX ou XXXX). Ils participaient ensuite à une série de tests de généralisation variés, impliquant notamment la présentation de séquences de longueur étendue (e.g. AXXXB ; Test 2), et l'introduction de nouveaux stimuli X (Test 3). L'un des chimpanzés testés a passé avec succès ces deux tests cruciaux. Ces résultats indiquent que cet individu n'a pas appris par cœur chacune des séquences vues au cours de l'entraînement, mais est capable de reconnaître les associations apprises malgré l'introduction de nouveaux éléments intermédiaires. La tâche utilisée avait toutefois pour particularité d'impliquer la présentation simultanée des stimuli à l'écran (voir Figure 23b). Cette étude ne portait donc pas sur l'apprentissage de dépendances non-adjacentes dans le temps, mais dans l'espace. Or il est possible que ce type d'apprentissage engage des processus cognitifs différant au moins partiellement de ceux impliqués dans l'apprentissage de régularités séquentielles. L'apprentissage de dépendances non-adjacentes présentées séquentiellement est notamment susceptible d'impliquer un stockage temporaire en mémoire de travail du premier élément de l'association (e.g. Wilson et al., 2017). Tandis que dans la tâche utilisée par Sonnweber et al. (2015), une fois que les chimpanzés ont appris que seuls les éléments situés aux extrémités des chaînes étaient pertinents pour effectuer la discrimination, il leur est possible d'effectuer de rapides mouvements oculaires de l'un à l'autre, les traitant comme s'ils étaient adjacents malgré la distance spatiale.

La généralisation de ce type de dépendances non-adjacentes à de nouveaux éléments intermédiaires a toutefois été récemment montrée chez le diamant mandarin, dans une expérience impliquant cette fois-ci la présentation séquentielle de stimuli auditifs (Chen & ten Cate, 2017). Comme dans l'étude de Sonnweber et al. (2015), les oiseaux étaient d'abord entraînés à apprendre deux dépendances adjacentes. Mais à la différence de Sonnweber et al. (2015), les participants étaient ensuite directement testés (i.e. sans feedback) sur des séquences congruentes dans lesquelles un élément intermédiaire avait été injecté. L'un des trois oiseaux testés a passé avec succès ce test.

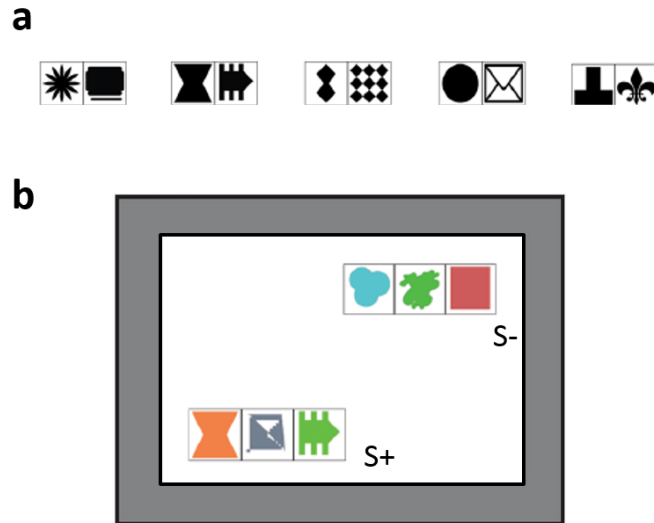


Figure 23 a. Les cinq paires de stimuli que les chimpanzés de l'étude de Sonnweber et al. (2015) devaient apprendre à associer. b. Illustration de la tâche. Les deux chaînes de stimuli étaient présentées simultanément sur un écran tactile, et les chimpanzés devaient sélectionner le stimulus positif pour recevoir une récompense.

Enfin, une étude de Wilson, Smith, et al. (2015) nous renseigne sur l'extraction de dépendances non-adjacentes présentées dans une grammaire régulière plus complexe, comprenant plusieurs dépendances adjacentes et des dépendances non-adjacentes optionnelles (certains éléments pouvant être présentés de manière adjacente ou non ; voir la Figure 13, p. 43, qui représente un diagramme de la grammaire utilisée dans cette étude). Deux macaques et des participants humains ont été exposés à des séquences générées par cette grammaire, puis exposés à deux types de séquences : des séquences impliquant uniquement des violations de relations locales, et des séquences impliquant également des violations de dépendances non-adjacentes. Les résultats obtenus chez les macaques suggèrent que ceux-ci ont bien détecté les violations des transitions entre éléments adjacents, mais sont restés insensibles aux violations des relations non-adjacentes. Un pattern de résultats intéressant est obtenu chez les humains : seule une petite proportion des participants a détecté les violations des dépendances non-adjacentes. Ces résultats concordent avec ceux rapportés dans la section IV.1, et suggèrent que les dépendances adjacentes sont plus faciles à extraire que les dépendances non-adjacentes pour le macaque également.

IV.2.2.2 Corrélats neuronaux

Une seule étude a été consacrée à l'examen des corrélats neuronaux du traitement de dépendances non-adjacentes chez le macaque rhésus (Milne et al., 2016), en répliquant partiellement une étude effectuée chez des participants humains (Mueller et al., 2012). Après une exposition passive à des séquences auditives de la forme AXB, les potentiels évoqués enregistrés par électroencéphalographie (EEG) montrent que le pattern de réponse neuronal des macaques à des séquences violant la dépendance (e.g. A₁XB₂) présente des similarités avec celui observé chez des enfants testés dans des conditions similaires. Chez l'adulte humain, ces violations ne sont pas accompagnées de modifications significatives de l'activité neuronale, dans les mêmes conditions d'exposition passive aux régularités (Mueller et al., 2012). Ces résultats suggèrent que le traitement des dépendances non-adjacentes par le primate non-humain pourrait présenter davantage de similarités avec celui du bébé humain qu'avec celui de l'adulte. Les auteurs de cette étude proposent notamment que chez ces deux premiers types de participants, l'apprentissage de dépendances non-adjacentes pourrait reposer sur un apprentissage implicite des régularités statistiques présentes dans l'entrée sensorielle, tandis que chez l'adulte humain des mécanismes supportant un apprentissage explicite de ces régularités seraient recrutés. Nous reviendrons sur cette discussion dans la section suivante.

IV.3 Mécanismes

IV.3.1 Théories sur l'apprentissage statistique

L'un des postulats de la linguistique générative est que les propriétés distributionnelles des langues naturelles ne peuvent constituer la source des connaissances syntaxiques. On retrouve également cette idée dans les premiers travaux sur l'AGL, qui ont interprété les performances des participants dans des tâches de jugement de grammaticalité comme reflétant l'extraction et l'encodage de règles abstraites des grammaires sous-tendant les séquences présentées (Reber, 1967).

Dans une approche très différente, Skinner (1991) avait proposé que l'acquisition du langage dépend de l'apprentissage de contingences via des mécanismes généraux de conditionnement opérant, comme dans tout autre domaine de la cognition et chez toute espèce.

Les théories de l'apprentissage statistique ont réactualisé ce débat (Bates & Elman, 1996; Seidenberg, MacDonald, & Saffran, 2002).

La notion d'apprentissage statistique fait référence à un ensemble de mécanismes, ou de principes computationnels, permettant l'extraction de régularités distributionnelles par simple exposition à celles-ci. Ces mécanismes sont sensibles aux propriétés redondantes de l'environnement, et permettent par exemple de former des prédictions sur les événements à venir dans une séquence (e.g. Saffran, 2001). Plusieurs approches théoriques peuvent être distinguées au sein de ce courant. Saffran et collaborateurs (1996 ; voir également Aslin, Saffran, & Newport, 1998; Saffran, 2001) ont proposé que l'extraction de régularités distributionnelles soit basée sur l'apprentissage des co-occurrences d'éléments les plus fréquentes dans l'entrée sensorielle, et la computation des probabilités transitionnelles entre ces éléments. L'approche exemplariste propose que les participants mémorisent les exemplaires de séquences auxquels ils sont exposés, et effectuent leurs jugements de grammaticalité en comparant les nouveaux exemplaires rencontrés avec ceux-ci (Vokey & Brooks, 1992). L'approche fragmentariste enfin propose que seuls des fragments (*chunks*) d'éléments présents dans les séquences de familiarisation soient mémorisés (Perruchet & Pacteau, 1990; Servan-Schreiber & Anderson, 1990).

IV.3.2 Apprentissage statistique des dépendances non-adjacentes

Les dépendances non-adjacentes posent en principe un problème pour les théories sur l'apprentissage statistique. Les premiers travaux sur l'apprentissage implicite de régularités statistiques se sont largement concentrés sur l'apprentissage de relations entre éléments adjacents (Stadler & Frensch, 1998). Par ailleurs, supposer que les probabilités transitionnelles entre les éléments adjacents d'une séquence ne seraient pas seulement calculées entre des éléments adjacents, mais également entre éléments plus distants, mène à un problème d'explosion combinatoire (e.g. Levelt, 2008). Les dépendances non-adjacentes posent également un problème pour les approches exemplaristes et fragmentaristes. Selon cette dernière, par exemple, seuls des fragments d'éléments adjacents seraient mémorisés du moins dans les versions initiales de ces théories (Perruchet & Pacteau, 1990; Servan-Schreiber & Anderson, 1990). En effet, ces fragments ne comportent pas d'interruption ou de « blanc ».

Mais l'ensemble des données expérimentales que nous venons de passer en revue indiquent que les dépendances non-adjacentes peuvent être apprises par les humains, enfants et

adultes, ainsi que par des participants appartenant à d'autres espèces, sur la base de régularités distributionnelles implémentés dans des langages artificiels. Ces données indiquent également que l'apprentissage de ce type de dépendances est toutefois plus difficile que l'apprentissage de dépendances adjacentes. Plusieurs théories ont été proposées pour expliquer les contraintes semblant peser sur l'extraction des dépendances non-adjacentes.

Une première théorie propose que l'extraction de dépendances non-adjacentes, à la différence de celle des dépendances adjacentes, reposerait en partie sur des mécanismes non statistiques. Peña et al. (2002) ont observé qu'après une courte familiarisation de 10 min à un flux continu de syllabes organisées en triplets de la forme AXB (tâche de segmentation, voir l'Encadré 2), les participants échouent à généraliser les dépendances non-adjacentes à de nouveaux items intermédiaires (ces derniers étant des syllabes A ou B d'une autre paire, e.g. $A_iA_jB_i$). En effet, ils ne présentent pas de préférence pour ce type de triplet par rapport à des triplets non-congruents (e.g. C_jA_iX). Ils y parviennent cependant lorsque de courtes pauses, dites « subliminales », de 25 ms sont introduites entre les triplets présentés lors de la phase d'exposition. Selon les auteurs, ces résultats indiquent que (1) la segmentation d'un flux de syllabes peut s'effectuer sur la base de mécanismes d'apprentissage statistique calculant des probabilités transitionnelles non-adjacentes, mais que (2) la généralisation de dépendances non-adjacentes à de nouveaux éléments intermédiaires nécessiterait l'extraction de régularités structurelles (A_B) sur la base de mécanismes non statistiques (*rule-based*). Toujours selon les auteurs, l'extraction de ces régularités ne s'observerait que dans la condition où le flux de syllabes est explicitement segmenté par de courtes pauses, car l'extraction rapide de régularités structurelles ne peut se faire qu'une fois que les mots de l'entrée sensorielle ont été identifiés. Cependant, une série de revues critiques accompagnées de répliques partielles des expériences effectuées par Peña et al. (2002) suggère que leurs résultats reflètent un apprentissage des positions des éléments dans les triplets (e.g. les éléments A apparaissent en début de triplet), plutôt qu'un apprentissage des dépendances entre les éléments A et B (Endress & Bonatti, 2007; Perruchet, Tyler, Galland, & Peereman, 2004a). De manière plus générale, le matériel expérimental employé par Peña et al. (2002) rend difficile l'interprétation de leurs résultats. Seidenberg et al. (2002) recensent ainsi dix régularités statistiques différentes dans les séquences de familiarisation qui pourraient avoir été exploitées par les participants. Une étude de Frost et Monaghan (2016), effectuée sur le même modèle que celle de Peña et al. (2002) en

y apportant des ajustements méthodologiques²⁴, montre que dans ces conditions les participants humains peuvent simultanément segmenter un flux de syllabes sur la base des probabilités transitionnelles non-adjacentes, et généraliser ces dépendances non-adjacentes à de nouveaux éléments intermédiaires. La proposition théorique avancée par Peña et al. (2002) n'a donc pas été corroborée par cet ensemble de données.

Un deuxième ensemble de théories avance que l'apprentissage de dépendances adjacentes et non-adjacentes reposent sur les mêmes mécanismes d'apprentissage statistique. Gómez (2002) propose que les contraintes semblant peser sur l'apprentissage de ces dernières découle d'une priorité donnée aux items étant situés à proximité dans l'espace ou le temps. Ce n'est que lorsque les probabilités transitionnelles entre éléments adjacents sont trop faibles, par rapport à celles entre les éléments non-adjacents, que ces dernières seraient exploitées par les participants. Pour tester son hypothèse, elle a exposé différents groupes de participants à un langage artificiel composé de séquences séparées par des silences, chacune de la forme AXB (voir Encadré 2). Le nombre de mots correspondant à l'élément intermédiaire X différait selon les groupes (e.g. 2, 6, 12 ou 24 dans son Expérience 1). Les participants étaient ensuite testés avec une tâche de jugement de grammaticalité, dans laquelle leur étaient présentées ces mêmes séquences *versus* des séquences violant les dépendances non-adjacentes. Seuls les participants du groupe ayant été exposés au langage artificiel impliquant 24 éléments X différents ont correctement rejeté les séquences non-congruentes. Gómez (2002) explique ces résultats de la manière suivante : les participants des autres groupes n'ont pas rejeté les séquences non-congruentes, car ils se seraient basés sur les dépendances adjacentes pour répondre à la tâche. Or, ces dépendances adjacentes étaient bien respectées dans les deux types de séquences test. Les participants exposés au langage comportant la plus grande variabilité d'éléments intermédiaires, eux, sont les seuls à avoir été sensibles aux dépendances non-adjacentes. Les participants se seraient donc focalisés par défaut sur les relations entretenues entre les éléments adjacents, et n'auraient exploité les dépendances non-adjacentes que lorsque ces premières s'avéraient peu informatives (voir également Onnis et al., 2005).

Une théorie explicative complémentaire a été tirée des travaux démontrant le rôle joué par la similarité perceptive des éléments non-adjacents (par rapport aux éléments intermédiaires). Newport et Aslin (2004 ; voir également Creel et al., 2004) proposent que les

²⁴ Cette étude montre notamment que la généralisation à de nouveaux éléments est bien observée lorsque ces éléments sont réellement nouveaux (i.e. des syllabes non présentées au cours de la phase de familiarisation), et non des syllabes A ou B d'une autre paire (e.g. A_iA_jB_i), comme dans Peña et al. (2002).

mécanismes d'apprentissage statistique sous-tendant l'extraction de dépendances non-adjacentes seraient contraints par une loi de similarité, telle que celle proposée par la théorie gestaltiste : des éléments se ressemblant physiquement auraient tendance à être davantage regroupés, ou associés, même s'ils sont relativement distants l'un de l'autre dans le temps ou l'espace.

(Pacton & Perruchet, 2008; Pacton et al., 2015) ont offert une théorie unificatrice, compatible avec l'ensemble des conditions facilitatrices identifiées pour l'extraction des dépendances non-adjacentes, et qui permettrait d'expliquer la priorité donnée aux dépendances adjacentes. Ils proposent que l'extraction de dépendances adjacentes et non-adjacentes est sous-tendue par un seul et même système/mécanisme associatif, soumises à la même contrainte : le traitement attentionnel conjoint de deux éléments serait une condition nécessaire, et suffisante, pour l'apprentissage d'une association. La contiguïté temporelle permettrait ainsi que deux stimuli soient simultanément (ou presque) dans le focus attentionnel. La variabilité de l'élément intermédiaire comme la similarité perceptive augmenteraient la saillance des stimuli non-adjacents, ce qui aurait également pour effet d'augmenter la probabilité qu'ils soient traités simultanément.

Un dernier type de proposition, complémentaire de la précédente, met l'accent sur le rôle de la mémoire de travail dans l'extraction des dépendances non-adjacentes. Celle-ci nécessiterait le maintien en mémoire de travail de l'élément prédictif, afin que sa relation avec l'élément prédit soit apprise (Conway, Deocampo, & Smith, 2016; Wilson, Marslen-Wilson, & Petkov, 2017). Les dépendances adjacentes et non-adjacentes seraient extraites par le même type de mécanismes statistiques, mais des limitations particulières (directement corrélées aux limitations de la WM), et des corrélats neuronaux particuliers (cf section II.2.4.3), devraient être observés spécifiquement lors de l'extraction de ces dernières.

Enfin, des données expérimentales suggèrent l'implication de processus de nature différente chez l'adulte et l'enfant. Dans une étude mesurant les potentiels évoqués enregistrés par électroencéphalographie (EEG), Mueller et al. (2012) ont en effet observé que les conditions d'apprentissage permettant l'apprentissage de dépendances non-adjacentes, ainsi que les corrélats neuronaux de cet apprentissage, différaient entre des participants adultes et des enfants. La tâche consistait en une écoute passive de la majorité des séquences comportant des dépendances non-adjacentes (e.g. A_1XB_1), mêlées à une minorité de séquences violant ces dépendances (e.g. A_1XB_2). Chez les enfants, une réponse de discordance (MMR) est observée

lors de la diffusion des violations, alors qu'aucune réponse différentielle n'est observée chez les adultes. Ce résultat ne refléterait pas une inattention portée à la tâche par ce groupe de participants, car une MMR est bien observée dans une condition contrôle (diffusion de séquences comprenant un élément de hauteur plus élevée, *pitch deviant*). Un deuxième groupe d'adultes a été testé avec un matériel similaire, mais dans une tâche où ils étaient explicitement informés de l'existence de séquences présentant des violations concernant la hauteur ou l'ordre des syllabes. Ces violations devaient être signalées par le sujet en pressant un bouton. Les adultes ayant appris les dépendances non-adjacentes dans ces conditions « d'exposition active », présentent bien une MMR, ainsi qu'une onde négative plus tardive (N2, ou N200). Selon les auteurs, cette dernière pourrait refléter l'engagement de processus de traitements moins automatiques chez l'adulte que chez l'enfant (ainsi que chez le singe, voir section IV.2.2.2), cette composante tardive ayant été reliée dans des études précédentes à un apprentissage explicite de régularités séquentielles (Eimer, Goschke, Schlaghecken, & Stürmer, 1996; Ferdinand, Mecklinger, & Kray, 2008)²⁵.

IV.4 Conclusions

L'ensemble de ces données suggère que l'apprentissage de dépendances non-adjacentes est sous-tendu par des mécanismes d'apprentissage statistique, présents chez un large ensemble d'espèces. Cet apprentissage est sujet à certaines contraintes en partie similaires chez l'humain et les autres espèces : l'exploitation des dépendances adjacentes tend à être privilégiée si celles-ci sont suffisamment informatives, la similarité perceptive entre des éléments distants dans le temps ou l'espace facilite leur association, etc. Plusieurs de ces aspects nécessiteraient cependant des recherches supplémentaires.

Un premier aspect concerne le type de représentations des dépendances non-adjacentes que forment les participants. Apprennent-ils une collection de séquences individuelles (e.g. $A_1X_1B_1$, $A_1X_2B_1$, etc.), ou forment-ils une représentation unifiée de chacune des dépendances non-adjacentes, indépendante des éléments intermédiaires (e.g. $A_1_B_1$) ? Dans les langues naturelles, cette indépendance est centrale, puisque l'accord en genre et en nombre, par

²⁵ Ces résultats font écho avec d'autres données obtenues chez l'adulte : l'application de TMS dans le cortex frontal dorso-latéral (DLPFC) s'accompagne d'une amélioration des performances dans une tâche d'apprentissage statistique (Ambrus, Janacsek, Trimborn, Kovacs, & Nemeth, 2017).

exemple, ou encore la formation du présent continu, implique d'extraire et de reconnaître les dépendances impliquées autour de noms ou de verbes très divers, qui n'auront éventuellement jamais été entendus, ou produits, dans ce contexte. Une manière d'étudier expérimentalement cette question est de tester si les participants continuent à reconnaître la dépendance non-adjacente apprise, malgré l'introduction de nouveaux éléments intermédiaires. La capacité de généraliser des dépendances non-adjacentes à de nouveaux éléments intermédiaires a été confirmée chez l'humain dans des expériences d'ALL (e.g. Frost & Monaghan, 2016). Toutefois, la question de la généralisation n'a fait l'objet que de deux expériences chez l'animal NH et, de plus, dans des conditions expérimentales particulières (Chen & ten Cate, 2017; Sonnweber et al., 2015). En effet, dans ces deux études, les participants ont été préalablement entraînés à associer les éléments en les présentant de manière adjacente²⁶. Cette procédure expérimentale a donc permis de mettre en évidence qu'à la suite d'un entraînement très progressif, deux individus appartenant à des espèces très différentes, le chimpanzé et le diamant mandarin, ont été capables de reconnaître de manière flexible les dépendances apprises, malgré des manipulations des éléments intermédiaires. Ces données suggèrent donc que, dans ces conditions, des animaux NH peuvent former des représentations des dépendances indépendantes de l'élément intermédiaire. Mais dans les langues naturelles par exemple, de nombreuses dépendances sont systématiquement non-adjacentes (comme dans le cas de l'accord en genre et en nombre en français, où les marqueurs sont systématiquement séparés par la racine du nom). Il reste donc à savoir si des animaux seraient capables de former ces représentations de la forme A_B, sans présentation préalable des éléments A et B de manière adjacente.

Par ailleurs, les travaux comparatifs se sont jusqu'à présent essentiellement concentrés sur une mise en évidence de la capacité d'autres espèces à apprendre des dépendances non-adjacentes, avec des résultats essentiellement positifs. Mais il pourrait exister certaines différences entre l'humain et les autres animaux sur des aspects plus fins de cet apprentissage. Notamment car l'apprentissage de dépendances non-adjacentes est susceptible d'engager la mémoire de travail (e.g. Wilson et al., 2017). Or, on sait que ces ressources sont plus limitées chez un certain nombre d'espèces que chez l'humain. Par exemple, une étude a été réalisée au sein de notre laboratoire pour évaluer la taille de l'empan mnésique du babouin de Guinée et de

²⁶ Sonnweber, Ravnani, et Fitch (2015) ne rapportent pas le nombre d'essais qui a été nécessaire pour que cet apprentissage se mette en place, mais Chen et ten Cate (2017) ont fait effectuer à leurs participants plus de 4000 essais pour apprendre deux associations.

l'humain (Joël Fagot & De Lillo, 2011a), à l'aide d'une tâche de rappel spatial immédiat (tâche de Corsi, 1972). Les babouins obtenaient de très bonnes performances lors de la reproduction de séquences de 3 items, mais celles-ci chutaient très fortement dès que des séquences de quatre items étaient présentées. Les humains, eux conservaient de très bonnes performances jusqu'à la longueur maximale testée (6 items). Ces limitations pourraient avoir pour conséquences de contraindre plus strictement le nombre d'éléments intermédiaires pouvant être insérés entre deux éléments non-adjacents, ou encore de nécessiter une durée d'exposition plus importante aux dépendances non-adjacentes (tandis qu'aucune différence, ou une différence moindre, serait observée entre l'humain et une autre espèce pour l'apprentissage de dépendances adjacentes). A ce jour, aucune expérience ne nous renseigne sur ces hypothèses. En effet, le nombre d'études comparatives testant dans des conditions expérimentales similaires des participants humains et NH sur l'apprentissage de dépendances non-adjacentes est très faible. Les quelques études de ce type révèlent pourtant des données intéressantes : elles montrent que les participants des différentes espèces testées parviennent bien à apprendre ces dépendances, mais également que des différences sont observées entre ces espèces. Elles suggèrent notamment : (1) l'existence de contraintes perceptives différentes selon les espèces (de la Mora & Toro, 2013; Newport et al., 2004), (2) qu'en présence de régularités plus locales, l'exploitation des dépendances non-adjacentes est moins privilégiée encore chez le primate non-humain que chez l'humain (Wilson et al., 2013), et que (3) les mécanismes recrutés chez le primate non-humain pour le traitement des dépendances non-adjacentes pourraient différer de ceux engagés chez l'adulte humain (Milne et al., 2016; Mueller et al., 2012).

Résumé

- **Chez l'humain, la capacité à apprendre des dépendances non-adjacentes apparaît plus tardivement au cours du développement que celle à apprendre des dépendances adjacentes, et ne se manifeste chez l'adulte que sous certaines conditions.**
- **Un certain nombre d'expériences d'apprentissage de langages artificiels (ALL) révèlent que l'animal NH est également capable d'apprendre ce type de dépendances.**
- **Des différences se dessinent toutefois dans les quelques expériences comparatives testant dans les mêmes conditions expérimentales des participants humains et NH.**

- **Davantage d'expériences comparatives sont nécessaires pour évaluer plus finement les similarités et différences dans les traitements de dépendances non-adjacentes chez différentes espèces.**

V. Extraction de Propriétés Structurelles

L'extraction de propriétés structurelles communes à un jeu de séquences permet la généralisation de la réponse comportementale à de nouvelles séquences isomorphes, qu'elles soient constituées d'éléments familiers ou nouveaux. Elle implique l'encodage de propriétés relativement abstraites, puisque ce sont les relations entretenues entre les différents éléments de la séquence qui sont à prendre en compte, et non l'identité spécifique de ces éléments. Cette capacité est davantage étudiée comme constituant un prérequis pour l'extraction de régularités structurelles dans les langues naturelles, plutôt que comme ayant des correspondances directes avec des phénomènes linguistiques existants (Corballis, 2009; Marcus, Vijayan, Rao, & Vishton, 1999).

L'apprentissage de ce type de régularités a été essentiellement étudié via l'utilisation de deux types de patterns²⁷ basés sur des répétitions. Un premier type de pattern comporte une répétition immédiate d'un item, tandis que le troisième élément diffère (i.e. les séquences sont de la forme ABB ou AAB). Dans un deuxième type de pattern, le premier et le dernier élément sont identiques et l'élément intermédiaire diffère (i.e. les séquences comportent une répétition non-adjacente, et sont de la forme ABA²⁸). Ces patterns ont pour avantage de correspondre à des séquences relativement courtes, et d'être basés sur des relations de similitude/différence entre des éléments, qui peuvent être reconnues par de nombreuses espèces (e.g. par l'abeille, Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001). Ils sont donc particulièrement appropriés pour étudier l'apprentissage de propriétés structurelles de manière comparative chez l'adulte, l'enfant, et les animaux NH. De plus, discriminer ces différents patterns ne peut s'effectuer en adoptant une stratégie de comptage : ils contiennent tous deux éléments différents, l'un n'apparaissant qu'une fois et l'autre étant répété (Marcus et al., 1999). Cette tâche ne peut donc

²⁷ Les auteurs utilisent souvent le terme de « pattern algébrique » (Marcus, Vijayan, Rao, & Vishton, 1999). L'utilisation de ce terme implique cependant un postulat fort sur la nature des mécanismes en jeu (cf. section V.3). A ce terme, nous préférons dans cette thèse celui de « pattern », plus neutre.

²⁸ Dans cette littérature, on trouve également la notation XXY/XYX pour faire référence à ces différents types de patterns. Cette notation est généralement utilisée pour souligner que les éléments constituant le pattern sont tirés du même lot d'éléments (e.g. que les éléments X et Y n'appartiennent pas à deux classes perceptives distinctes). Par souci de simplification, nous employons invariablement ici la notation AAB/ABB/ABA, et précisons lorsque c'est le cas que « A » et « B » réfèrent à deux classes de stimuli distinctes.

s'effectuer qu'en portant attention aux relations entretenues entre les éléments situés aux différentes positions de la séquence.

Les expériences réalisées sur l'apprentissage de patterns suivent les principes généraux de l'AGL : elles consistent en une phase de familiarisation aux régularités, suivie d'une phase de test. On distingue deux types de tests. Dans un premier type, de nouvelles séquences constituées des mêmes items que ceux ayant servi à la familiarisation sont présentées. Par exemple, si les participants ont été familiarisés avec les séquences suivantes : *121* et *232*, les séquences test pourront être *131* (congruente) et *133* (non-congruente). Dans un deuxième type de test, plus robuste, les individus sont testés sur des séquences constituées de nouveaux items (e.g. *454* et *455*).

V.1 Données chez l'humain

Les travaux menés chez l'humain se sont exclusivement concentrés sur l'ontogénèse de cette capacité. Dans une expérience *princeps*, Marcus et al. (1999) ont testé des bébés de sept mois dans un paradigme d'habituation/déshabituaiton. Les participants étaient d'abord exposés durant deux minutes à des triplets de syllabes, qui suivaient soit un pattern ABB (e.g. « *ga ti ti* ») soit un pattern ABA (« *ga ti ga* »). Ils étaient ensuite exposés à des séquences composées de nouvelles syllabes, la moitié étant congruente avec le pattern familier (e.g. « *wo fe fe* » et « *wo fe wo* », respectivement), et l'autre non (i.e. ces séquences non-congruentes suivaient un pattern ABA pour le groupe familiarisé au pattern ABB, et *vice-versa*). Les bébés des deux groupes ont davantage réagi lors de la diffusion de séquences non-congruentes, que de celle de séquences congruentes, ce qui indiquait, selon les auteurs qu'ils ont bien extrait les patterns sous-jacents.

Des études ont par la suite suggéré que cette aptitude serait facilitée par l'utilisation de patterns constitués de stimuli linguistiques. Marcus, Fernandes, et Johnson (2007) ont par exemple montré que des bébés de sept mois étaient capables d'extraire ces patterns de séquences composées de sons non-linguistiques, mais seulement après une familiarisation à des séquences isomorphes de syllabes. Différentes explications ont été proposées pour expliquer ces résultats : par l'intérêt particulier que suscite les stimuli langagiers pour le bébé (e.g. Vouloumanos & Werker, 2004, 2007), la familiarité de ce type de stimuli (Saffran & Thiessen,

2007; Seidenberg & Elman, 1999), ou encore leur statut de signal de communication (Ferguson & Lew-Williams, 2016).

Par ailleurs, un certain nombre de travaux suggèrent que la capacité à extraire des propriétés structurelles de séquences se développe de manière progressive et incrémentale, les régularités les plus simples étant les premières à pouvoir être détectées. Johnson et al. (2009) ont ainsi montré qu'à huit mois, des bébés sont capables de détecter la présence de répétitions immédiates dans des séquences de formes visuelles, mais n'étaient sensibles ni à la position de ces répétitions dans la séquence (i.e. ils ne distinguent pas entre un pattern AAB et un pattern ABB), ni à des répétitions non-adjacentes (ABA). A onze mois, les participants discriminaient AAB et ABB, mais ne parvenaient toujours pas à extraire un pattern ABA. Des résultats similaires ont été obtenus avec des stimuli linguistiques, mais avec un calendrier développemental différent. A la naissance, les bébés sont capables de discriminer un pattern ABB de séquences aléatoires ABC, mais ne traitent pas différemment un pattern ABA de ces mêmes séquences aléatoires (Gervain, Macagno, Cogoi, Pena, & Mehler, 2008). A sept mois en revanche, ils sont capables d'effectuer les deux types de discrimination (Gervain & Werker, 2013).

V.2 Données comparatives

V.2.1 Données comportementales

Une première série d'étude a mis en évidence la capacité à extraire des patterns chez plusieurs espèces de primates non-humains. Hauser et Glynn (2009) ont mené une expérience de play-back chez le macaque rhésus (*Macaca Mulatta*) en milieu naturel. Les macaques ont tout d'abord été familiarisés avec des séquences de la forme ABB ou AAB, les éléments A et B étant sélectionnés dans huit classes de vocalisations de cette espèce (e.g. *scream-scream-grunt*). Après avoir atteint le critère d'habituation (i.e. deux essais successifs sans réponse d'orientation en direction du haut-parleur), de nouvelles séquences étaient diffusées aux singes. Lors de cette phase de test, les macaques ont exprimé davantage de réponses d'orientation lors de la diffusion de séquences non-congruentes, que lors de la diffusion de séquences congruentes qui consistaient en de nouvelles combinaisons des mêmes classes de vocalisations. Ces résultats suggèrent que le macaque est capable de discriminer les patterns AAB et ABB, et de généraliser

cette discrimination, du moins à de nouvelles réorganisations des mêmes stimuli (voir également Procyk, Ford Dominey, Amiez, & Joseph, 2000). Plus récemment, Sonnweber et al. (2015) ont testé chez deux chimpanzés (*Pan troglodytes*) l'extraction d'un pattern de la forme $AB^{1-5}A$, (i.e. deux éléments A identiques séparés par un à cinq éléments différents, les éléments A et B étaient tirés du même lot de 70 stimuli). Les patterns correspondaient à des chaînes de formes visuelles présentées simultanément sur un écran (cf. Figure 23b, p. 80). Après un entraînement à discriminer ces chaînes d'autres chaînes dont le premier et dernier élément différaient, les chimpanzés ont généralisé leurs performances à des stimuli de couleurs et de formes nouvelles. Ces chimpanzés semblent en outre être capables de généralisation intermodale : dans une expérience ultérieure, leurs temps de réponse dans la tâche de discrimination étaient plus longs lorsqu'une séquence auditive non-congruente (ABB) leur était diffusée simultanément, par rapport aux essais dans lesquels il leur était diffusé une séquence isomorphe ABA (Ravignani & Sonnweber, 2017).

Un autre mammifère, le rat (*Rattus norvegicus*), a également été testé sur cette capacité. Une première expérience a été conduite par Toro et Trobalón (2005) avec les mêmes stimuli (des syllabes) que ceux employés par Marcus et al. (1999, Expérience 3). Cette expérience utilisait une tâche de conditionnement opérant. Les tests de généralisation n'indiquent pas que les rats réussissent à discriminer entre les différents patterns. Des résultats positifs ont toutefois été obtenus chez cette espèce avec d'autres types de stimuli. Murphy, Mondragón, et Murphy (2008, Expérience 2) ont entraîné des rats à discriminer des séquences composées de deux sons purs organisés selon un pattern ABA (e.g. A = 3.2 kHz and B = 9 kHz ou *vice-versa*), et AAB/ABB. Les rats sont ensuite parvenus à généraliser leurs performances à des séquences composées de deux nouveaux sons purs (e.g. C = 12.5 kHz and D = 17.5 kHz). Enfin, de la Mora et Toro (2013) ont entraîné des rats à discriminer un pattern AAB instancié par des voyelles séparées les unes des autres par des consonnes ($C_iVC_iVC_jV$, Expérience 1), ou entre des consonnes séparées par des voyelles ($V_iCV_iCV_jC$, Expérience 2), de séquences de la forme ABC. Dans ces deux expériences, les rats ont ensuite généralisé leurs performances à de nouveaux phonèmes de chacune de ces catégories.

Chez les oiseaux, une première série d'expériences a été menée chez le diamant mandarin (*Taeniopygia guttata*, van Heijningen, Chen, van Laatum, van der Hulst, & ten Cate, 2013). Les oiseaux ont été entraînés dans une tâche Go/No-go à répondre à des séquences de la forme ABA et à négliger de répondre à celles de la forme AAB/ABB, les éléments A et B étant tirés du même lot de vocalisations de cette espèce (cf. Figure 24). Une majorité de participants

ont atteint le critère d'apprentissage. Une série de tests menés ensuite indiquent toutefois que les animaux n'effectuaient pas la tâche en traitant la structure sous-jacente, e.g. en répondant aux séquences dont le premier et dernier élément étaient identiques, mais évitaient en réalité de répondre aux séquences contenant une répétition immédiate d'un élément A ou B. Par exemple, les séquences test de la forme ABAB étaient traitées par les oiseaux comme un stimulus positif, tandis que les séquences test ABBA l'étaient comme un stimulus négatif. De plus, la présentation de séquences non-congruentes de la forme AABB donnait lieu à un taux de réponses (« go ») plus faible encore que lors de la présentation de séquences AAB/ABB. Ce dernier résultat suggère selon les auteurs un effet cumulatif de la présence de répétitions. Des résultats similaires ont été obtenus chez le moineau du japon, dans une étude utilisant un protocole similaire (*Lonchura striata*, Seki, Suzuki, Osawa, & Okanoya, 2013). Ces expériences suggèrent donc que ces deux espèces d'oiseaux chanteurs sont capables de détecter la présence d'éléments répétés, et utilisent préférentiellement cette régularité pour effectuer la tâche de discrimination de patterns. Mais elles ne permettent pas de savoir si cet apprentissage est généralisable à de nouveaux éléments. Dans une expérience complémentaire abordant cette question, des diamants mandarins n'ont pas généralisé la détection de répétitions immédiates à de nouveaux types de vocalisations (Chen, van Rossum, & ten Cate, 2015). Les résultats de cette expérience suggèrent que les diamants mandarins se sont ici basés sur les positions de certains items en particulier (e.g. réponse aux nouvelles séquences terminant par un élément A).

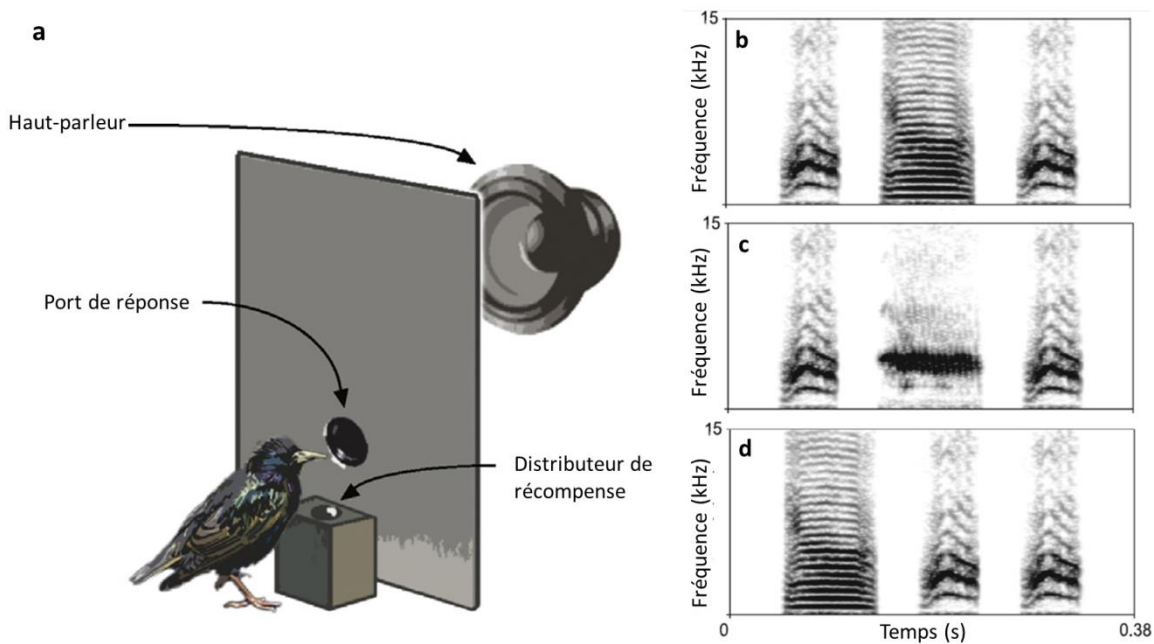


Figure 24 a Représentation schématique d'un dispositif de conditionnement opérant, pour une tâche « Go/No-go » de classification de stimuli auditifs. La réponse « Go » consiste pour l'oiseau à introduire son bec dans le port de réponse. Modifié d'après Comins et Gentner (2013). b, c, d. Exemples de séquences construites à partir de vocalisations de diamants mandarins, suivant un pattern ABA (b, c) ou ABB (d). Modifié d'après van Heijningen et al. (2013).

Récemment, (Spierings & ten Cate, 2016) ont introduit dans leur expérience un contrôle élégant de ce dernier facteur (cf. Table 2). Dans une tâche Go/No-go, ils ont entraîné des diamants mandarins et des perruches (*Melospittacus undulatus*) à discriminer des patterns AAB et ABA, constitués de différents types de vocalisations de diamants mandarins²⁹. Ils ont ensuite présenté aux oiseaux une série de séquences test congruentes et non-congruentes, en faisant varier systématiquement le nombre de positions partagées avec les séquences d'entraînement (Tests 1-5). Leurs résultats indiquent que les diamants ont appris et utilisé les positions des éléments : ils ont traité comme des S+ les séquences test qui comportaient un maximum d'éléments présentés à la même position que dans les séquences d'entraînement, qu'elles soient congruentes ou non avec la structure des séquences d'entraînement. Les résultats des perruches indiquent que celles-ci se sont au contraire basées sur le caractère isomorphe ou non des séquences test. Dans une expérience supplémentaire, les perruches ont par ailleurs été capables de généraliser leurs performances à des séquences composées de nouveaux types de vocalisations.

²⁹ Les diamants mandarins ont également été testés avec des sons artificiels dans cette étude, avec des résultats similaires.

Table 2 **Contrôle du nombre de positions partagées entre séquences d'entraînement et test**

Positions partagées	Test	Entraînement		n
		S+	S-	
Non contrôlées	ACA	<u>A</u> BA	<u>A</u> AB	5 vs 3
		<u>A</u> DA	<u>A</u> AD	
		<u>B</u> CB	<u>C</u> CD	
Contrôlées	FAF	<u>C</u> AC	<u>A</u> AD <u>B</u> BF	1 vs 2

Exemple de séquences présentées à un animal entraîné à répondre ("Go") à un pattern ABA. Les lettres (A-F) correspondent ici à différents types de vocalisations de diamants madarins. Dans la colonne "test" figurent de nouveaux triplets congruents avec le pattern. Dans les colonnes "S+" et "S-" figurent toutes les séquences d'entraînement partageant au moins une position (éléments soulignés) avec la séquence test. La colonne "n" indique le nombre de positions de la séquence test partagées avec les séquences S+ *versus* S-. Typiquement, lorsque les positions partagées ne sont pas contrôlées, les nouveaux triplets congruents ont davantage de positions en commun avec les séquences S+ qu'avec les séquences S- (e.g. ici, 5 *versus* 3). En contrôlant ce facteur, le nombre de positions partagées peut être équilibré, ou même être renversé (ici, 1 *versus* 2). Les exemples donnés ici sont tirés de Chen et al. (2015, Expérience 1b, Test 2), et Spierings & ten Cate (2016, Test 4).

Une autre série d'expériences, menée chez l'étourneau sansonnet (*Sturnus vulgaris*), met en lumière le rôle crucial que peuvent jouer certains choix méthodologiques sur l'extraction de régularités structurelles (l'implémentation des structures par des catégories perceptives, Comins & Gentner, 2013 ; l'utilisation d'un large jeu de séquences d'entraînement, Comins & Gentner, 2014). Toutefois, dans ces expériences, la généralisation à de nouvelles classes d'éléments n'est pas testée, ce qui ne permet pas de conclure quant au caractère généralisable des régularités extraites par ces étourneaux.

Enfin, des résultats intéressants ont été obtenus chez le poussin (*Gallus gallus* ; Santolin, Rosa-Salva, Regolin, & Vallortigara, 2016). Dans une première expérience, les poussins étaient entraînés dans un paradigme de conditionnement opérant à discriminer des patterns AAB et ABA, présentés sous la forme de chaînes de stimuli visuels présentés simultanément sur un écran (cf. Figure 25). A l'issue de cet entraînement, les poussins sont parvenus à généraliser leurs performances à de nouveaux stimuli visuels. Dans une autre expérience, ils parviennent également à discriminer des patterns AAB et ABB. Ces résultats suggèrent que les poussins

sont capables d'extraire des régularités structurelles présentées dans des patterns de stimuli visuels (voir également Versace, Spierings, Caffini, ten Cate, & Vallortigara, 2017).

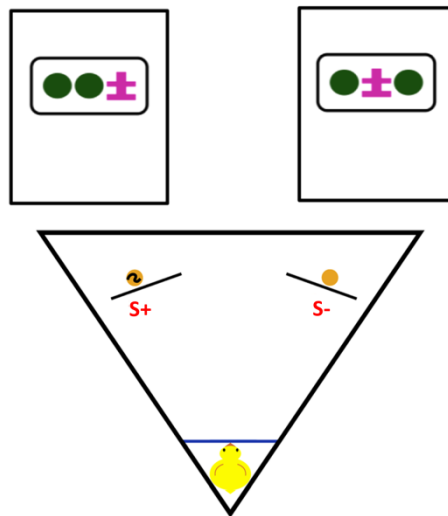


Figure 25 Illustration schématique d'une tâche de conditionnement opérant testant la discrimination de pattern chez le poussin, d'après Santolin et al. (2016). Les images présentées dans la partie haute du schéma représentent les stimuli affichés sur les écrans situés à gauche et à droite de l'arène, indiqués par les lignes obliques. Dans cet exemple, le participant est entraîné à choisir les séquences de la forme AAB, une récompense alimentaire étant dissimulée derrière l'écran correspondant.

V.2.2 *Corrélat neuronal*

Dans une série d'expériences comparatives combinant un paradigme « *oddball* » (stimulus discordant) et les techniques d'imagerie fonctionnelle (fMRI), il a été montré que le traitement de certaines propriétés structurelles implique des réseaux en partie similaires chez l'humain et le macaque. Dans ces expériences, des séquences de tons sont diffusées aux participants, étant tous identiques (AAAA), ou comportant un élément final discordant (i.e. un autre ton, AAAB). Des violations des régularités structurelles (i.e. la diffusion d'une séquence AAAB après une série de AAAA, et *vice-versa*) sont accompagnées d'activations cérébrales dans des régions du cortex préfrontal, notamment certaines régions du VLPFC et de l'insula (ainsi que les régions aSTS et pSTS du lobe temporal), chez les deux espèces (Bekinschtein et al., 2009; Uhrig, Dehaene, & Jarraya, 2014; L. Wang, Uhrig, Jarraya, & Dehaene, 2015). La présentation de nouvelles séquences sonores suivant le même pattern n'est pas accompagnée d'une activation différentielle de ces régions préfrontales, confirmant que ce sont bien les violations de pattern, et non d'exemplaires spécifiques mémorisés, qui sont détectées. Ces résultats indiquent que (1) les macaques encodent le pattern sous-jacent à ces séquences (voir

également Shima et al., 2007), et que (2) cet encodage est sous-tendu par des régions similaires chez l'humain et le macaque. De plus, des régions similaires (e.g. la partie ventrale du sulcus intra pariétal, VIP) s'activent chez les deux espèces lors de violations du nombre d'items dans les séquences (e.g. présenter une séquence AB, ou AAAAAB, après la diffusion de plusieurs séquences AAAB ; (L. Wang et al., 2015). Une différence intéressante est toutefois rapportée : tandis que chez le macaque les réseaux activés par ces deux types de violations sont bien distincts, chez l'humain, deux régions, le gyrus frontal inférieur (IFG) et la partie postérieure du sulcus temporal supérieur (pSTS ; voir la Figure 19, p. 62), s'activent lorsque surviennent l'une ou l'autre de ces violations. Selon les auteurs, ces données suggèrent que les humains formeraient une représentation plus intégrée des différentes propriétés caractérisant des séquences auditives (ici, structurelle et numérique) que les singes.

V.3 Mécanismes

Les données passées en revue dans les deux sections précédentes suggèrent que le bébé humain d'un certain âge, ainsi que différentes espèces NH, sont capables de former des représentations suffisamment abstraites des patterns pour permettre la reconnaissance de séquences isomorphes composées d'éléments nouveaux. Il existe cependant un débat ancien, et encore non résolu, sur la nature exacte des mécanismes sous-tendant ce comportement. On peut distinguer trois grands ensembles de théories, qui proposent respectivement : (1) l'implication de mécanismes basés sur l'utilisation de règles, (2) l'existence de mécanismes spécialisés dans la détection de répétitions, et (3) l'implication de mécanismes d'apprentissage non-symboliques.

V.3.1 Apprentissage basé sur des règles

Il existe une longue tradition de théories proposant que la généralisation de régularités structurelles (ou grammaticales) à un nouveau vocabulaire reposerait sur l'apprentissage de règles abstraites (e.g. Reber, 1969). Cette lignée est reprise par Marcus et al. (1999, voir également Marcus, 2001), qui proposent que l'apprentissage de patterns implique des opérations symboliques d'extraction de règles algébriques. Ces mécanismes supporteraient un encodage des positions sérielles (1^{er} item, 2^{ème} item, etc.) sous forme de variables, et un

encodage des relations entretenues par ces variables sous la forme de relations algébriques (e.g. identique, différent, $>$, $<$). La représentation du pattern résultant de cet encodage serait tout à fait indépendante des exemplaires rencontrés (i.e. abstraite au sens strict). Ces mécanismes seraient généraux, i.e. non restreints à des domaines particuliers, mais ils joueraient un rôle clé dans l'acquisition du langage (Chomsky, 1980; Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995; Pinker, 1991).

V.3.2 *Apprentissage statistique*

Un autre ensemble de théories propose d'expliquer l'extraction de régularités structurelles par des mécanismes d'apprentissage généraux non symboliques. Une première série de travaux de ce type a consisté à chercher à modéliser les performances des participants de Marcus et al. (1999) à l'aide de simulations par des réseaux connexionnistes. Tandis que lors des simulations initiales effectuées par Marcus et al. (1999), le modèle échouait à discriminer les séquences congruentes et non-congruentes, les différents modèles élaborés ultérieurement y parviennent, en implémentant diverses modifications au modèle initial (Altmann, 2002; Altmann & Dienes, 1999; Christiansen & Curtin, 1999; Negishi, 1999; Seidenberg & Elman, 1999; Shultz & Bale, 2001). Par exemple, Altmann (2002) est parti de l'idée que les bébés de sept mois sur lesquels cette expérience a été réalisée pouvaient être considérés comme ayant déjà été familiarisés au cours de leur vie aux syllabes utilisées dans l'expérience. Il a donc entraîné au préalable son modèle sur des phrases simples, avant de l'entraîner sur le jeu de séquences de familiarisation utilisées par Marcus et al. (1999, Expérience 1). Dans un test ultérieur, le modèle est parvenu à discriminer les nouvelles séquences correspondant à chaque pattern. Mais toutes ces simulations ont fait l'objet de critiques (voir les réponses individuelles de Marcus à chacun des articles cités ci-dessus, ainsi qu' Endress et al., 2007; Shultz & Bale, 2001). De plus, une limite majeure de ces travaux est qu'ils tendent davantage à proposer des explications alternatives aux données obtenues par Marcus et al. (1999), qu'à proposer un cadre explicatif général caractérisant des mécanismes statistiques qui permettraient de généraliser sur la base de propriétés structurelles.

Une deuxième lignée de travaux s'est attachée à étendre les théories classiques de l'apprentissage statistique, initialement appliquées à l'apprentissage de propriétés de surface, à l'apprentissage de propriétés structurelles. Nous ne présenterons pas en détail ces différents

modèles, mais utiliserons deux exemples pour souligner les principales différences entre les théories sous-jacentes à ces modèles, et celles basées sur l'utilisation de règles.

Brooks et Vokey (1991) ont proposé que l'on puisse effectuer des jugements de similarité entre des séquences, non pas seulement sur leurs propriétés de surface, mais également sur la base des relations entretenues par les différents items au sein de ces séquences (*relational analogy*, ou *abstract analogy*). Ce mécanisme permettrait par exemple de juger que les séquences de lettres « MXVVVM » et « BDCCCB » sont similaires dans la mesure où, dans chacune de ces séquences, la première et la dernière lettre sont identiques, et que chacune comporte un triplet de lettres identiques vers la fin. Le jugement s'effectue via la comparaison du nouvel exemplaire avec un ou plusieurs exemplaires mémorisés, plutôt qu'avec une représentation symbolique des propriétés structurelles de ces séquences. Autrement dit, la représentation du pattern, ou de la structure, n'est pas indépendante des exemplaires présentés (pour une application plus directe de cette idée à l'acquisition du langage, voir Abbot-Smith & Tomasello, 2006). L'encodage de régularités structurelles et de régularités de surface pourraient ainsi ne pas différer par la nature des mécanismes impliqués, mais par le nombre et le type de propriétés des séquences prises en compte (Pothos, 2005). Les jugements basés sur des règles consisteraient, selon cette vue, en des jugements de similarité ne prenant en compte qu'un petit jeu de propriétés (e.g. structurelles), en négligeant les autres propriétés caractérisant la séquence (e.g. identité des items).

Une deuxième idée ayant été avancée est que les computations statistiques pourraient être effectuées non pas seulement au niveau des mots ou des syllabes, mais également sur les catégories syntaxiques (Seidenberg et al., 2002). Ainsi la sensibilité aux propriétés distributionnelles des différentes catégories, et, pour le cas des patterns, des différentes positions dans la séquence, pourraient constituer la base des jugements de grammaticalité.

V.3.3 Apprentissage par des mécanismes spécialisés

Il a également été proposé que l'extraction de patterns basés sur des répétitions implique des mécanismes non pas généraux, mais *spécialisés* (Endress et al., 2007). D'après cette proposition théorique, la détection de répétitions serait sous-tendue par un mécanisme appartenant à un jeu de primitives perceptives et mnésiques (en anglais, *perceptual or memory primitives* ; Endress, Nespors, & Mehler, 2009). Ces mécanismes seraient automatiques, et prendraient place dans les étapes précoces du traitement (Endress et al., 2007; Gervain et al.,

2008). Ils dériveraient de contraintes générales des systèmes perceptifs et mnésiques, et certains d'entre eux au moins pourraient être partagés avec d'autres espèces (Endress, Carden, Versace, & Hauser, 2010).

Différentes données expérimentales ont amené à postuler l'existence d'un mécanisme spécialisé pour la détection de répétitions. Par exemple, Endress et al. (2007) ont exposé des adultes à des séquences de tons générés par deux types de grammaires : grammaires basées sur des répétitions (ABA, ABB), et grammaires ordinales (tons bas-haut-medium, medium-haut-bas). Leurs prédictions étaient que si seuls des mécanismes symboliques généraux sous-tendent l'extraction de ces grammaires, les performances devraient être équivalentes pour les deux types de grammaires. En effet, d'un point de vue computationnel, les relations d'égalité ($X=Z$) sont implémentées et traitées exactement comme les relations de supériorité et infériorité (e.g. $X < Z$). Or, les participants de cette étude ont eu des performances bien meilleures avec les grammaires basées sur des répétitions qu'avec les grammaires ordinales.

V.4 Conclusions

Les expériences menées chez l'animal NH suggèrent que la capacité à discriminer différents patterns basés sur des répétitions, et de généraliser cette performance à de nouveaux stimuli, est relativement répandue au sein de la phylogénèse. Mais la nature des régularités extraites reste incertaine pour une grande majorité de ces expériences. En effet, discriminer les patterns AAB/ABB et ABA peut s'effectuer en ayant appris à détecter la présence ou l'absence d'une répétition immédiate (ten Cate & Okanoya, 2012; Corballis, 2009)³⁰. La généralisation de cette discrimination à de nouveaux items pourrait ainsi refléter l'extraction de la propriété de répétition immédiate, mais ne refléterait pas l'encodage de *l'ensemble* des propriétés structurelles caractérisant ces différents patterns (e.g. dans une séquence de la forme AAB, les deux premiers éléments sont identiques tandis que le dernier diffère ; dans une de la forme ABA, le premier et dernier élément sont identiques, tandis que le second diffère). En revanche, discriminer les patterns AAB et ABB ne peut s'effectuer via cette même stratégie binaire

³⁰ On peut également citer une stratégie à laquelle nous avons déjà fait référence dans la section II.2.3 (voir Table 1, p. 36) : exploiter la différence de variabilité perceptive entre les différents types de patterns. En effet, une séquence de la forme ABA comporte deux transitions entre des éléments ou des classes d'éléments différents, tandis qu'une de la forme AAB/ABB n'en comporte qu'une.

(détection de la présence/absence de répétition immédiate). Un individu entraîné à discriminer ces deux patterns a toutefois pu apprendre à porter attention à la présence/absence d'une répétition en *début* de séquence par exemple, et à ignorer le dernier élément de la séquence, là encore sans traiter l'ensemble de la structure sous-jacente (i.e. cet individu traiterait une séquence AAA de la même manière ; ten Cate & Okanoya, 2012). Une simple tâche de discrimination entre ces trois patterns, en l'absence de tests supplémentaires (tels que ceux employés par van Heijningen et al., 2013 et Chen et al., 2015), ne permet donc pas de distinguer entre les trois stratégies suivantes : (1) détection d'une répétition immédiate, (2) détection d'une répétition immédiate à une certaine position dans la séquence, et (3) encodage de l'ensemble de la structure sous-jacente.

Or, nous avons vu que les études menées chez le bébé humain ont montré que ces trois stratégies se mettaient en place progressivement au cours du développement (Gervain et al., 2008; Gervain & Werker, 2013 ; Johnson et al., 2009). Une prochaine étape pour l'étude comparative de l'extraction de régularités structurelles est donc de parvenir à identifier plus précisément les types de régularités structurelles exploitées par les différentes espèces dans ces tâches.

Un deuxième aspect encore peu exploré dans une perspective comparative est la caractérisation des mécanismes sous-jacents. A ce stade de l'avancée des débats sur la nature symbolique ou non-symbolique de ces mécanismes, les recherches sur l'animal NH ne semblent pas pouvoir apporter de données conclusives. Les données passées en revue suggèrent seulement que si l'extraction de patterns est sous-tendue chez l'humain par des mécanismes d'apprentissage de règle, ces mécanismes sont susceptibles d'être apparus de manière précoce au cours de la phylogénèse. La question se pose en revanche de savoir si les animaux NH sont dotés de mécanismes spécialisés leur permettant de détecter la présence de répétitions dans des séquences de stimuli.

Résumé

- **Différentes espèces d'animaux NH testés dans des tâches de discrimination de patterns basés sur des répétitions (AAB/ABB/ABA) se révèlent capables de généraliser leurs performances à de nouveaux éléments.**
- **Les stratégies employées par ces animaux pour effectuer la tâche sont toutefois peu claires.**

- **Chez le bébé humain, la capacité à détecter des répétitions immédiates apparaît plus tôt au cours de l'ontogenèse que celle à détecter des répétitions non-adjacentes.**
- **Les animaux NH pourraient se baser sur la détection de la présence/absence de répétitions immédiates pour effectuer ces tâches, plutôt que d'extraire l'ensemble du pattern sous-jacent.**

Chapitre III : Problématique

Nous avons vu dans le Chapitre I qu'une hypothèse dominant actuellement les théories sur l'évolution des capacités syntaxiques est que la capacité à traiter des grammaires supra-régulières serait apparue récemment dans la lignée humaine. Cette hypothèse tend à être confirmée par les données expérimentales actuellement disponibles, qui suggèrent une incapacité des animaux NH testés à traiter ce type de grammaire. Toutefois, nous avons vu qu'il était difficile de tirer une conclusion définitive sur la base de ces expériences, du fait de divers problèmes méthodologiques. Le principal problème que nous avons identifié est que les matériels expérimentaux employés comportaient diverses régularités pouvant être exploitées par les animaux pour effectuer la tâche, sans traiter la grammaire sous-jacente.

Un examen complet de la littérature révèle par ailleurs que les différentes espèces testées, y compris l'humain, rencontrent également des difficultés dans l'encodage de régularités séquentielles de complexité intermédiaire qui constituent des prérequis cognitifs au traitement de grammaires supra-régulières (Chapitre II). La détection de dépendances non-adjacentes, qu'elles soient implémentées dans des séquences d'éléments associés (Chapitre II.1), ou dans des patterns basés sur des répétitions (i.e. répétitions non-adjacentes, Chapitre II.2), est plus difficile pour l'humain que l'extraction d'associations ou de répétitions adjacentes (i.e. elle apparaît plus tard au cours de l'ontogénie, et tend à n'être observée que sous certaines conditions). Les expériences menées chez l'animal NH ne traitent pour la plupart (à l'exception de Wilson, Smith, et al., 2015) qu'indirectement cette question.

Or, nous proposons que les échecs obtenus dans les expériences testant la capacité des animaux NH à traiter de grammaires supra-régulières puissent être imputés à un biais vers l'exploitation d'informations locales, et une difficulté à traiter les régularités à plus longue distance. Dans ce cadre, l'objet de la présente thèse est double :

- (1) Evaluer directement l'hypothèse selon laquelle la détection de régularités entre des éléments non-adjacents présente une difficulté particulière pour le primate NH, à la fois pour l'apprentissage d'associations et l'extraction de patterns basés sur des répétitions.**

(2) Tester la capacité du primate NH à traiter des grammaires supra-régulières dans des conditions expérimentales prévenant l'exploitation de régularités locales.

Les travaux réalisés au cours de cette thèse pour remplir ces deux objectifs ont donné lieu à cinq articles expérimentaux, et un chapitre théorique.

Dans la première étude, nous avons examiné l'apprentissage d'associations non-adjacentes de manière comparative chez le babouin de Guinée et l'humain. Ce travail avait pour but de combler deux manques identifiés dans cette littérature : (1) l'absence de données comparatives sur les dynamiques temporelles de cet apprentissage, et (2) une méconnaissance des représentations que forment les animaux NH testés de ces dépendances non-adjacentes. Nous avons pour cela employé une tâche de pointage sérielle, qui offre une mesure continue de l'apprentissage.

Pour la seconde étude, nous avons évalué la sensibilité du babouin de Guinée aux répétitions adjacentes (ABB) et non-adjacentes (ABA), après une exposition répétée au pattern correspondant (Expérience 1), et sans exposition préalable (Expérience 2). Ces expériences ont été réalisées à l'aide d'une adaptation de la tâche de pointage sérielle que j'ai élaborée au cours de cette thèse.

La troisième étude présente une analyse corrélationnelle des performances individuelles des babouins dans les deux études précédentes, et dans des études réalisées par d'autres membres de l'équipe. Les données récoltées nous ont notamment permis d'examiner si les performances des singes pour l'extraction de dépendances adjacentes et non-adjacentes corrélaient entre elles, ainsi que d'une tâche à l'autre.

Dans la dernière étude, nous rapportons les résultats d'une expérience dans laquelle nous avons contrasté chez le babouin l'apprentissage de deux grammaires supra-régulières, l'une impliquant des dépendances centrées-emboîtées, et l'autre des dépendances croisées.

Dans une dernière section théorique, nous résumons et discutons ces travaux expérimentaux. Nous replaçons enfin ces données au regard des théories existantes sur l'évolution des capacités syntaxiques, et proposons des perspectives pour une meilleure compréhension de l'évolution du langage à travers l'étude comparée des fonctions cognitives générales.

VOLET EXPERIMENTAL



I. Méthodologie Générale

I.1 Expériences chez le primate non-humain

I.1.1 Le babouin de Guinée

Toutes les expériences menées sur le primate non-humain dans le cadre de ma thèse ont été conduites sur des babouins de Guinée (*Papio papio*). Cette espèce appartient à la famille des *cercopithecidae*, qui sont des singes de l'ancien monde (cf. Figure 26). D'après nos connaissances actuelles sur l'histoire évolutive des primates, on estime que la séparation entre les cercopithécidés et les hominidés remonte à 30-35 millions d'années (Ciochon & Fleagle, 1987; Steiper, Young, & Sukarna, 2004).

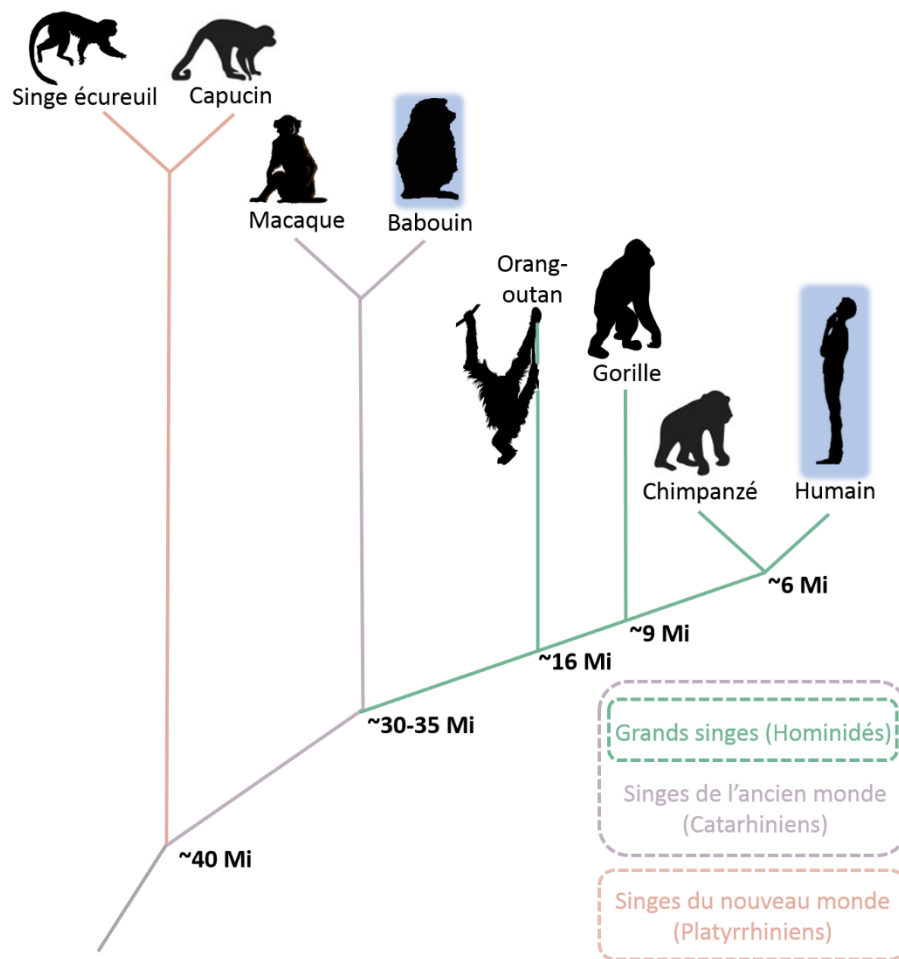


Figure 26 Illustration de la place du babouin de Guinée et de l'humain au sein de l'arbre phylogénétique de l'ordre des primates.

Le genre *Papio* comprend cinq espèces différentes de babouins, dont l'aire de répartition et la structure sociale varie d'une espèce à l'autre. L'aire de répartition du babouin de Guinée se situe en Afrique de l'ouest, et s'étend sur une petite zone comprenant la Guinée, la Gambie, le Sénégal, le sud de la Mauritanie et l'ouest du Mali. Son organisation sociale se caractérise par des groupes de taille relativement importante, étant pour une majorité d'entre eux multi-mâle et multi-femelle (avec 23 individus en moyenne dans ces groupes, et un sexe-ratio moyen chez les adultes de 1:2.5 (Patzelt et al., 2011). Des analyses génétiques réalisées sur des populations sauvages indiquent un biais vers une dispersion des femelles et une philopatrie des mâles (Kopp, Fischer, Patzelt, Roos, & Zinner, 2015).

Les recherches menées sur la cognition de cette espèce au sein du laboratoire ont permis de caractériser ses similitudes et différences avec la cognition humaine dans des domaines aussi

divers que la perception visuelle (e.g. Parron, Deruelle, & Fagot, 2007; Parron & Fagot, 2007), la mémoire (Joël Fagot & Cook, 2006; Joël Fagot & De Lillo, 2011a), ou encore la catégorisation (Joël Fagot, Goldstein, Davidoff, & Pickering, 2006; Martin-Malivel & Fagot, 2001). Mais ce sont plus particulièrement les résultats de deux séries de travaux effectués sur ces babouins, l'une portant sur l'apprentissage de régularités statistiques et l'autre sur l'extraction de propriétés relationnelles, qui rendent ce modèle particulièrement pertinent pour notre problématique.

Plusieurs études menées au sein du laboratoire ont en effet montré que ces singes possèdent de bonnes capacités d'apprentissage statistique dans la modalité visuelle. Une étude de Goujon et Fagot (2013) a ainsi montré qu'ils étaient capables d'apprendre rapidement des configurations spatiales permettant de prédire la localisation d'une cible sur un écran. Dans une tâche de recherche visuelle reposant sur le paradigme de *contextual cueing* (Chun & Jiang, 1998), les babouins devaient détecter un stimulus cible (un « T ») qui apparaissait parmi des distracteurs (« L ») sur un écran, et toucher ce premier stimulus. Six configurations différentes de distracteurs étaient présentées, chacune associée à une localisation spécifique de la cible. Les résultats montrent que les babouins étaient sensibles à cette information contextuelle, leurs temps de réponse sur la cible diminuant rapidement au cours de leur exposition aux configurations. Toujours dans la modalité visuelle, Grainger, Dufau, Montant, Ziegler, et Fagot (2012) ont montré que les babouins étaient capables d'apprendre à classifier des chaînes de lettres affichées sur un écran en mots et non-mots (e.g. « casino » et « caviro »), et de généraliser leur comportement à de nouvelles chaînes. Les données indiquent que les singes ont discriminé les mots des non-mots en détectant les bigrammes (5/6 babouins) et trigrammes de lettres (1/6 babouin) qui se trouvaient plus fréquemment dans les mots par rapport aux non-mots (Hannagan, Ziegler, Dufau, Fagot, & Grainger, 2014). Les travaux réalisés par Laure Minier-Munding, qui a effectué sa thèse au sein de l'équipe, ont par la suite montré que les babouins étaient également capables d'apprendre des régularités statistiques séquentielles et qu'ils pouvaient notamment segmenter une séquence en sous-unités sur la seule base de ses propriétés distributionnelles (Minier-Munding, 2015). Pour effectuer ces recherches, elle a opérationnalisé une tâche de pointage sériel dont le principe a été conservé pour les expériences menées dans le cadre de ma thèse. Cet ensemble de résultats suggère que ces babouins sont sensibles aux régularités statistiques entre élément adjacents (dans l'espace ou dans le temps), ce qui est un prérequis à l'extraction des régularités plus complexes qui sont l'objet de ma thèse.











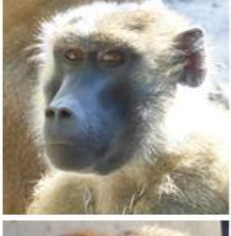


Dans une autre série de travaux, les babouins se sont montrés capables de catégoriser des paires de stimuli en fonction de la similarité des relations qu'ils entretiennent entre eux (Joël Fagot & Maugard, 2013; Joël Fagot & Thompson, 2011; Flemming, Thompson, & Fagot, 2013; Maugard, Wasserman, Castro, & Fagot, 2014). Dans ces tâches, les babouins étaient d'abord exposés à une paire de deux stimuli visuels identiques (e.g. une paire de triangles), ou différents (e.g. un triangle et un cercle). Deux autres paires de stimuli leur étaient ensuite présentées, l'une entretenant le même type de relation, et la deuxième correspondant à l'autre type de relation. Les babouins étaient récompensés s'ils sélectionnaient la paire qui présentait la même relation que la paire initiale. Une partie des babouins testés ont complété avec succès ces tâches, et généralisé leur comportement à de nouvelles paires de stimuli. Ces résultats indiquent que les babouins sont capables de traiter des relations entre relations pour des paires (e.g. (Joël Fagot & Thompson, 2011) ou des chaînes (Flemming et al., 2013) de stimuli présentés simultanément. Ces travaux suggèrent donc que les babouins sont capables d'extraire les propriétés structurelles de ces ensembles de stimuli. L'un des objectifs de ma thèse est donc de vérifier dans quelle mesure cette capacité est également rencontrée dans le cas d'une présentation séquentielle des stimuli.

1.1.2 Participants

Les expériences de cette thèse ont été réalisées chez un groupe de babouins comprenant 26 individus (Table 3). Chaque expérience a été proposée à l'ensemble du groupe, mais seuls sont retenus pour l'analyse des résultats ceux ayant achevé l'expérience dans le temps imparti.

Table 3 Caractéristiques des babouins ayant participé aux expériences : sexe, famille, âge (en années, au 01/01/2016).

	NOM	Sex.	Fam.	Age		NOM	Sex.	Fam.	Age
	ANGELE	F	1	10		BOBO	M	5	9
	ARIELLE	F	2	10		BRIGITTE	F	2	19
	ARTICHO	M	3	10		CAUET	M	2	8
	ATMO- -SPHERE	F	7	17		CLOCLO	M	1	8
	B06	F	10	31		DAN	M	1	7
	BARNABE	M	4	9		DORA	F	2	7

	NOM	Sex.	Fam.	Age		NOM	Sex.	Fam.	Age
	DREAM	F	5	7		HARLEM	M	1	3
	EWINE	F	4	6		HERMINE	F	6	3
	FANA	F	4	5		KALI	F	4	20
	FELIPE	M	1	5		PETOU-LETTE	F	1	16
	FEYA	F	2	5		PIPO	M	8	16
	FILO	M	1	5		VIOLETTE	F	6	11
	FLUTE	F	4	5		VIVIEN	M	9	11

I.1.3 Dispositif expérimental

Tous les babouins ayant participé aux expériences sont logés à la station de primatologie du CNRS de Rousset sur Arc (France), au sein de la plateforme *Comportement et Cognition du Primate*. La plateforme a pour particularité de permettre un logement des babouins en groupe social, avec un accès libre 24h/24 et 7j/7 à dix box expérimentaux (dispositif ALDM, pour *Automated Learning Device for Monkeys*, (Joël Fagot & Bonté, 2010a; Joël Fagot & Paleressompoulle, 2009). En 2017, ce groupe comprenait 32 singes. La plateforme comprend un enclos extérieur de 750 m² et une animalerie de 20 m², ainsi que deux bungalows dans lesquels se trouvent les box de test (voir Figure 27 a-b). Les singes peuvent entrer et sortir librement des box de test, permettant une participation à volonté aux expériences. Chaque box est équipé d'un écran tactile (19 pouces, EloTouch, 1915L Intellitouch), dont le babouin est séparé par un panneau comprenant trois ouvertures (l'une pour les yeux, et deux pour les mains, voir Figure 27 c-d), ainsi que d'un dispositif de récompense délivrant des grains de blé. Chaque babouin est équipé d'une puce dans chaque main, permettant sa reconnaissance via un système de radio-identification (RFID). Ainsi, les babouins peuvent interrompre leur session à tout moment, et reprendre là où ils en étaient à leur prochaine venue / initier l'essai suivant. Les expériences sont programmées par Joël Fagot sur Eprime Pro 2.0 (Pittsburgh: Psychology Software Tools, Inc, 2012).

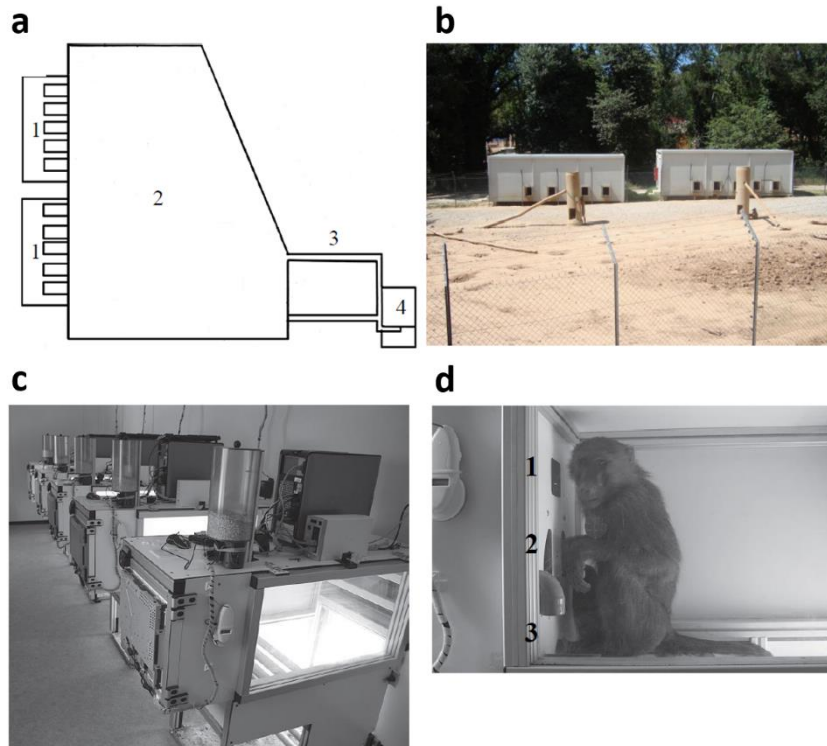


Figure 27 **a.** Illustration de l'organisation spatiale de la plateforme Comportement et Cognition du Primate de Rousset. (1) Bungalows hébergeant les dix box expérimentaux, (2) enclos extérieur, (3) tunnels connectant l'enclos à (4) l'animalerie. **b.** Vue des bungalows et de l'enclos extérieur depuis le toit de l'animalerie. **c.** Vue de l'intérieur de l'un des bungalows. **d.** Babouin dans un box de test. (1) Ouverture par laquelle le babouin peut voir l'écran, (2) ouvertures par lesquels le babouin peut passer ses mains pour toucher l'écran (noter que la présence de deux ouvertures, sur la gauche et la droite du panneau, permettant au babouin d'utiliser ses deux mains pour répondre), (3) dispositif de récompense délivrant des grains de blés.

Le système expérimental développé dans le laboratoire offre la possibilité aux babouins de participer aux expériences à volonté, sans qu'il soit nécessaire de les isoler de leur groupe social et/ou de leur imposer des sessions de travail à des heures prédéfinies. Dans ces conditions, les babouins effectuent un nombre considérable d'essais par jour (e.g. 1520 essais en moyenne par jour et par singe en juillet 2012, (Joël Fagot, Marzouki, Huguet, Gullstrand, & Claidière, 2015) ; nombre pouvant varier selon la période considérée et le type de tâche). Les données relatives à l'utilisation des box indiquent que les babouins ont tendance à travailler par courtes sessions de 3.13 minutes en moyenne (± 3 secondes), pour une durée d'utilisation journalière de 177 minutes en moyenne par individu ((Joël Fagot et al., 2015). Des observations éthologiques réalisées sur le budget-temps des babouins suggèrent que seules les périodes d'inactivité se trouvent réduites par la présence de ce dispositif, tandis que le temps alloué aux déplacements et aux interactions sociales s'en trouve relativement inchangé, avec une tendance

à une augmentation (Joël Fagot et al., 2015). Par ailleurs, un important recouvrement est observé entre le réseau social mesuré dans l'enclos et la répartition des babouins dans les box : les babouins qui se livrent régulièrement à des interactions affiliatives ont tendance à occuper au même moment des box adjacents (Claidière, Gullstrand, Latouche, & Fagot, 2017). L'ensemble de ces données suggère donc que ce dispositif offre à la fois un enrichissement à ces animaux captifs, et permet d'obtenir un nombre important de données de qualité.

I.2 Expériences chez l'humain

Une partie des expériences réalisées sur le babouin ont également été réalisées chez des participants humains, afin d'offrir des données comparatives.

I.2.1 Participants

Les participants humains ont été recrutés à l'université d'Aix-Marseille. Pour être inclus dans notre échantillon, les étudiants devaient convenir à la liste de conditions suivante : (1) être droitiers, (2) avoir une vue normale ou corrigée, et (3) être naïfs par rapport aux buts de l'expérience (i.e. pour chaque expérience, nous nous assurons que les candidats n'avaient participé ni à l'une de mes expériences antérieures, ni à l'une de celles de Minier-Munding, 2015). Les participants étaient dédommagés 10 euros par heure.

I.2.2 Matériel et procédures

Les humains étaient testés dans un box expérimental du plateau technique *Comportement et cognition humaine* du laboratoire de psychologie cognitive de l'Université d'Aix-Marseille. L'écran tactile utilisé était identique à ceux utilisés par les babouins (19 pouces, EloTouch, 1915L Intellitouch), et les expériences étaient programmées par mes soins sur Eprime 2.0.

Afin de tester les participants dans des conditions les proches possibles de celles des babouins, les instructions qui leur étaient données étaient réduites à leur strict minimum. A la différence des babouins, les humains ne recevaient pas de récompense alimentaire. Toutefois, la durée de la pénalité de temps avant la présentation de l'essai suivant imposée après une erreur

(5 secondes) était similaire pour les babouins et les humains, procédure permettant de renforcer les réponses correctes.

II. ETUDE I – Traitement des Dépendances Non-adjacentes chez le Primate et Humain et Non-humain : Dynamiques Temporelles et Généralisation

Malassis, R., Rey, A., & Fagot, J. (*sous presse*). Non-adjacent Dependencies Processing in Human and Non-Human Primates. *Cognitive Science*.

L'apprentissage de dépendances non-adjacentes nécessite d'apprendre puis de reconnaître des relations prédictives entre des éléments séparés par un ou plusieurs éléments intermédiaires variables. Chez l'humain, cette spécificité des dépendances non-adjacentes par rapport aux dépendances adjacentes se traduit par leur apparition plus tardive au cours du développement ontogénétique (Gómez & Maye, 2005), et la nécessité de certains facteurs facilitants pour en réussir l'extraction (e.g. comme une forte similarité perceptive entre les éléments non-adjacents, Onnis, Monaghan, Richmond, & Chater, 2005). Des travaux expérimentaux menés au cours des dix dernières années suggèrent que certaines espèces animales partagent avec l'humain la capacité à extraire des dépendances non-adjacentes présentées dans des séquences de stimuli dans des grammaires de type AXB (A et B étant les deux éléments associés, et X un élément variable). Ces travaux suggèrent donc une similarité entre l'humain et l'animal en ce qui concerne les processus d'apprentissage sous-tendant l'extraction de dépendances non-adjacentes.

Une des limitations de cette littérature est qu'elle ne nous fournit aucune donnée comparative nous renseignant sur les dynamiques temporelles de cet apprentissage. De plus, aucune expérience n'a examiné si l'animal Non-humain est, comme l'humain, capable de détecter des dépendances non-adjacentes présentées dans un langage artificiel, et de les généraliser à de nouveaux éléments intermédiaires (Frost & Monaghan, 2016). Nous avons donc cherché à répondre à ces questions par une série de deux expériences.

Dans la pratique, nous avons utilisé une tâche de pointage sériel dans ces deux expériences, dans laquelle les sujets devaient toucher une cible apparaissant successivement à différentes localisations sur un écran tactile. Les déplacements de la cible suivaient une règle AXB : la localisation de la première cible permettait de prédire la localisation de la troisième cible, tandis que la localisation intermédiaire variait. Les temps de réponses obtenus sur la troisième cible fournissent une mesure continue de l'apprentissage, le principe de cette tâche

étant que les temps de réponse diminuent lorsque la localisation de la prochaine cible peut être anticipée par le participant.

La première expérience, conduite chez le babouin ($n = 16$) et l'humain ($n = 10$), a permis de mettre en évidence des différences interspécifiques dans les dynamiques de cet apprentissage : Pour apprendre ces régularités, les babouins ont eu besoin d'une exposition aux dépendances non-adjacentes bien plus importante que les humains testés dans les mêmes conditions. Dans la deuxième expérience, les babouins ont cependant été capables de généraliser les dépendances apprises lorsqu'elles ont été présentées au sein de nouvelles séquences.

Ainsi, les résultats de ces deux expériences indiquent que si les babouins sont capables d'apprendre et de généraliser des dépendances non-adjacentes, ces opérations présentent une difficulté plus importante pour eux par rapport aux humains. Ces résultats sont discutés au regard des données obtenues dans une autre expérience comparative babouins-humains menée par notre équipe, dans laquelle il n'était pas observé de différences entre ces deux espèces quant à la quantité d'exposition nécessaire pour extraire des dépendances *adjacentes*, dans une tâche similaire (Rey, Minier, Malassis, Bogaerts, & Fagot, 2018 ; en annexe p. 324).

Non-adjacent Dependencies Processing
in Human and Non-human Primates

Raphaëlle Malassis, Arnaud Rey and Joël Fagot

Laboratory of Cognitive Psychology
Institute for Language, Communication, and the Brain
Aix-Marseille University & CNRS
Marseille, France

Corresponding author: Raphaëlle Malassis, raphaelle.malassis@univ-amu.fr

Laboratoire de Psychologie Cognitive – CNRS

Aix-Marseille Université

3, place Victor Hugo - Case D

13331 Marseille Cedex 03 – FRANCE

Abstract

Human and non-human primates share the ability to extract adjacent dependencies and, under certain conditions, non-adjacent dependencies (i.e., predictive relationships between elements that are separated by one or several intervening elements in a sequence). In this study, we explore the online extraction dynamics of non-adjacent dependencies in humans and baboons using a serial reaction time task. Participants had to produce three-target sequences containing deterministic relationships between the first and last target locations. In Experiment 1, participants from the two species could extract these non-adjacent dependencies, but humans required less exposure than baboons. In Experiment 2, the data show for the first time in a non-human primate species the successful generalization of sequential non-adjacent dependencies over novel intervening items. These findings provide new evidence to further constrain current theories about the nature and the evolutionary origins of the learning mechanisms allowing the extraction of non-adjacent dependencies.

Keywords: language evolution, statistical learning, sequence learning, long-distance dependencies, animal cognition

Statistical learning can be defined as the implicit learning of regularities embedded in the environment, which has been proposed to play an important role in language acquisition (for a review, see Romberg & Saffran, 2010) and many other aspects of cognition (e.g., in visual perception, Fiser & Aslin, 2001). Several experimental paradigms have been used to study these fundamental learning mechanisms such as the Artificial Grammar Learning paradigm (AGL, Reber, 1967), the Artificial Language Learning paradigm (ALL, Saffran et al., 1996) or the Serial Response Time paradigm (SRT, Nissen & Bullemer, 1987). In these experiments, participants are typically exposed to sequences of nonsense stimuli (e.g., syllables, tones or visual shapes) that are organized with a specific grammar. Learning of the predictive relationships embedded in the input can then be assessed by presenting novel sequences that are either consistent or inconsistent with this grammar.

Comparative human/non-human studies can inform us about the nature and dynamic of these learning mechanisms, and their occurrence during the evolution, by distinguishing domain-general and evolutionary old processes from those that might have appeared more recently in the human lineage. Comparative experiments have demonstrated so far that animals, like humans, are highly proficient in learning predictive relationships between *adjacent* elements (i.e., elements that are presented one after the other in a sequence, without any delay or element in between; see for a review: (ten Cate & Okanoya, 2012; Conway & Christiansen, 2001; Wilson et al., 2017). This ability has been reported in several primates (tamarins: Hauser et al., 2001; macaques: Wilson et al., 2013; Wilson, Smith, & Petkov, 2015; and baboons: Minier, Fagot, & Rey, 2016), and non-primate species (pigeons: Froehlich et al., 2004; rats: Toro & Trobalón, 2005; and songbirds: Takahasi, Yamada, & Okanoya, 2010). Along with neuroimaging data (e.g., Wilson, Kikuchi, et al., 2015), these behavioral data have led to the hypothesis that the extraction of adjacent dependencies relies on evolutionarily conserved mechanisms (Friederici, 2004; Wilson et al., 2017).

In contrast, the evolutionary origins of the capacity to extract non-adjacent dependencies (i.e., predictive relationships between elements that are separated by one or several intervening elements in a sequence, hereafter “NADs”) remain a matter of debate. A standard example of NADs in English is the relationship between auxiliaries and inflectional morphemes that are separated in the speech stream by the verbal root (e.g., *is reading*). Crucially, extracting and recognizing these dependencies requires generalization over a variable verb. Studies conducted in humans indicate that NADs extraction is more challenging for humans than the extraction of regularities between adjacent elements (Cleeremans & McClelland, 1991; Gebhart et al., 2009;

Newport & Aslin, 2004; Pacton et al., 2015; Wilson, Smith, et al., 2015). This capacity develops later in human infancy (Gómez & Maye, 2005) than the sensitivity to adjacent dependencies (Saffran, Aslin, & Newport, 1996; but see Marchetto & Bonatti, 2013, for different findings). Moreover, the extraction of NADs looks especially challenging in human adults outside some specific facilitative contexts. For instance, it is facilitated when the non-adjacent elements have a high degree of perceptual (e.g., phonological) similarity (Creel, Newport, & Aslin, 2004; Gebhart et al., 2009; Onnis, Monaghan, Richmond, & Chater, 2005), when the intervening elements are highly variable (Gómez, 2002; Onnis, Monaghan, Christiansen, & Chater, 2005), or when the non-adjacent elements are located at the edge of the sequences (Peña, Bonatti, Nespó, & Mehler, 2002). Studies also revealed similar performances and constraints on NADs learning in experiments using linguistic and nonlinguistic stimuli, such as tones (Creel et al., 2004; Endress, 2010; Gebhart et al., 2009), or actions (Endress & Wood, 2011), suggesting that the learning mechanisms involved in this ability are not language-specific. Several experiments aimed to test whether animals can also extract NADs. Wilson et al. (2015) exposed human participants and macaques to a complex grammar involving multiple adjacent and non-adjacent dependencies. All participants from both species detected violations of the adjacent dependencies. By contrast, half of the tested humans, and none of the monkeys were sensitive to the NADs. The authors concluded that when multiple regularities are present, monkeys rely preferentially on local ones, whereas humans exhibit more flexibility. Human learners can encounter some difficulties with the learning of several regularities simultaneously (e.g., Kovács & Mehler, 2009), but Wilson et al. (2015) suggests that this might be even more difficult for non-human learners. Animals' focus on various local cues rather than on NADs has also been reported when multiple embedded NADs must be processed, as in the so-called center-embedded grammars (e.g., van Heijningen, Visser, Zuidema, & ten Cate, 2009). However, these results may point to a limitation in the monkeys' ability to track several NADs simultaneously, rather than an inability to extract NADs *per se*.

Complementary information on animals' ability to learn NADs has been obtained from experiments using simpler grammars of the form $AX^{(n)}B$. In these grammars, A and B are two paired elements, with a non-adjacent transitional probability of 1, and $X^{(n)}$ is one or several (n) variable intervening elements. Two main types of NAD have been investigated thus far using these grammars: dependencies between elements that are perceptually more similar compared to the interspersed element(s) (i.e., that belong to the same perceptual category, hereafter

“feature-based NAD”), and learning of NADs without such perceptual cues (often called “arbitrary associations”).

Successful learning of feature-based NADs has been recently reported in two primate species, with visual (Sonnweber, Ravignani, & Fitch, 2015) and auditory stimuli (Ravignani, Sonnweber, Stobbe, & Fitch, 2013). Successful generalization across isomorphic visual and auditory sequences was also demonstrated recently (Ravignani & Sonnweber, 2017). In addition, positive results were obtained in rats, with NADs between elements that were phonologically similar (i.e., vowels vs. consonants, and *vice-versa*; de la Mora & Toro, 2013), whereas previous experiments conducted on NADs extraction in this species led to negative results in the absence of phonological similarity (Toro & Trobalón, 2005). Perceptual similarity therefore appears to facilitate the extraction of NADs in some non-human animal species, as it does in humans (e.g., Onnis et al., 2005). These findings support the hypothesis that NADs extraction is sustained by a general-purpose learning mechanism, interacting with general perceptual constraints (i.e., extra-linguistic, such as Gestalt principle of similarity, Creel et al., 2004), both being shared by phylogenetically distant species (Newport & Aslin, 2004; see however de la Mora & Toro, 2013, and Toro, Nespors, Mehler, & Bonatti, 2008, for a discussion of potential human-specific constraints on speech processing).

Evidence for successful NADs learning in the absence of such perceptual cues can also be found in non-human animals. To date, three distinct experiments were conducted on that topic. Newport, Hauser, Spaepen, and Aslin (2004) assessed the ability of tamarins (*Saguinus oedipus*) to extract dependencies between non-adjacent syllables in a speech stream, using a habituation/dishabituation procedure. Six 3-syllable nonsense words of the form AXB (three instantiations of A-B pairs and one of two X syllables inserted in the middle) were presented in a continuous stream during several minutes via a loudspeaker. After this familiarization phase, the tamarins’ reactions to words (e.g., A₁XB₁) *versus* part-words (B₁A₂X) were tested. Tamarins exhibited a higher rate of orientation responses towards the speaker after a part-word than after a word, suggesting that they learned the NADs. The authors concluded that the tamarins accurately segmented this continuous stream into triplets of syllables based on these NADs. However, one limitation of this study is that the tamarins’ ability to generalize these dependencies over novel intervening X syllables was not tested. Therefore, this experiment leaves unexplored the question of how the NADs were processed and stored in their memory. Indeed, they could have extracted the three A-B non-adjacent relationships, or instead memorized the six AXB “words” as wholes. The same limitation also applies to Milne et al.

(2016) who used a passive listening method in rhesus macaques (*Macaca mulatta*) and the recording of event-related potentials. Despite the results suggesting the learning of NADs, this study failed to provide generalization tests to confirm that the learned NADs could be maintained over novel intervening events.

To our knowledge, Sonnweber et al. (2015) is the only study proposing this kind of test in a non-human species. These data were obtained in one subgroup of chimpanzees (*Pan troglodytes*) referred to as the “arbitrary associative dependencies” group. These chimpanzees were presented on each trial with two “strings” of visual shapes depicted concurrently on a screen in a two-alternative forced choice task. The task required to choose the strings of the form AXⁿB (with five instantiations of A-B pairs and 60 instantiations of X items, with $n = 1$ or 2) over XXⁿX strings. What the participants precisely learned was evaluated by a series of tests proposed after the training phase. The two main generalization tests (Test 2: extension of the number of X elements, and Test 3: introduction of novel X shapes) were successfully completed by one participant over the three tested chimpanzees. However, the fact that the items were presented concurrently rather than sequentially raises potential issues. The processing of associations between spatially distant elements, especially at rather small distance (up to 12 cm from each other during training, with a maximum of 30 cm in Test 2, according to the reported pixel measures), can arguably rely on strategies different from those available for the extraction of sequential NADs. For instance, once the chimpanzees have learned that only the left- and right-most shapes of each string matter in this task, they can make quick eye-movements from one to the other and process them as if they were adjacent, despite the spatial distance. Extraction and recognition of sequential NADs, in contrast, may require to hold an element in working memory and relate it to another element occurring later in the sequence (Wilson et al., 2017). Therefore, Sonnweber et al. (2015) offers promising findings about non-human primates’ ability to generalize NADs, but these researches would need to be extended to the case of sequential NADs. Another limitation of the comparative literature is that it contains no comparative data on the temporal dynamics of NADs learning in human and non-human animal species. Earlier, we reported that baboons and humans learned adjacent dependencies at the same speed (Minier et al., 2016; Rey, Minier, Malassis, Bogaerts, & Fagot, in press). A remaining issue is whether this is also true for NADs.

The aim of the present study was twofold in this context. It was firstly to provide human and baboon comparative data on the fine-grained dynamics of NADs extraction. It was secondly to test whether baboons can generalize sequential NADs over novel intervening elements. In

Experiment 1, baboons and humans performed a serial response time task (SRT, Nissen & Bullemer, 1987), requiring to follow a stimulus appearing sequentially at different locations on a touchscreen. In a subset of the presented sequences, the first target location predicted the last target location, while the second target location varied. Response times provided an online behavioral metric of NADs learning. In Experiment 2, generalization of the NADs over novel intervening locations was tested. During this test, novel consistent sequences were contrasted with novel sequences violating the dependencies. This method was inspired from Gómez (2002). Only baboons took part in this second part of the study, as generalization has already been demonstrated in humans (Endress & Bonatti, 2007; Frost & Monaghan, 2016; Marchetto & Bonatti, 2015).

Experiment 1

Method

Baboons

Participants and apparatus

Participants are 16 Guinea baboons (*Papio papio*, 11 females, age range 4-21 years) from the CNRS primate facility (Rousset-sur-Arc, France). They live within a larger group of 26 individuals, within a 700 m² outdoor enclosure and had a permanent access to ten Automated Learning Devices for Monkeys (ALDM, for a detailed description, see (Joël Fagot & Bonté, 2010b; Joël Fagot & Paleressompoulle, 2009) equipped with a 19-inch touch screen and a food dispenser. The main feature of ALDM equipment is that a radio frequency identification reader (RFID) identifies each baboon via a microchip implanted in each arm. The baboons can therefore participate to the research at will, without being captured, as the test programs recognize them automatically. All baboons had previously participated to numerous computerized experiments using the ALDM test systems, including experiments on sequence processing (Minier et al., 2016), but they have never been exposed to tasks requiring the learning of NADs. The experiment was controlled by the EPrime software (Version 2.0, Psychology Software Tools, Pittsburgh).

Procedure

Participants had to perform a serial response time (SRT) task which required following and touching a dot moving in a pre-defined order on a touch screen. Each trial began by the display of a fixation cross (120x120 pixels) presented at the bottom of the screen (see Fig. 1a). Touching this stimulus triggered the display of a matrix of 9 locations (3x3) which contained 8 white crosses (60x60 pixels each) and a red circle target (80x80 pixels). Participants had to touch the target. The target then disappeared and moved to the next location on the screen. There was no delay between the offset of the target and its next display. A trial consisted in a sequence of three targets. An accurate completion of the 3-target sequence delivered grains of dry wheat inside the ALDM test unit. An incorrect response (i.e., selection of an incorrect location) stopped the trial and triggered a 5-sec timeout without food reward. Trials in which the participants failed to select a stimulus within 5-sec after the target's appearance were aborted and presented again in the next trial. The time elapsed between the appearance of the target (i.e., the red circle) and the baboon's touch on this target was recorded as the response time. Response times were recorded for each response in the 3-target sequence.

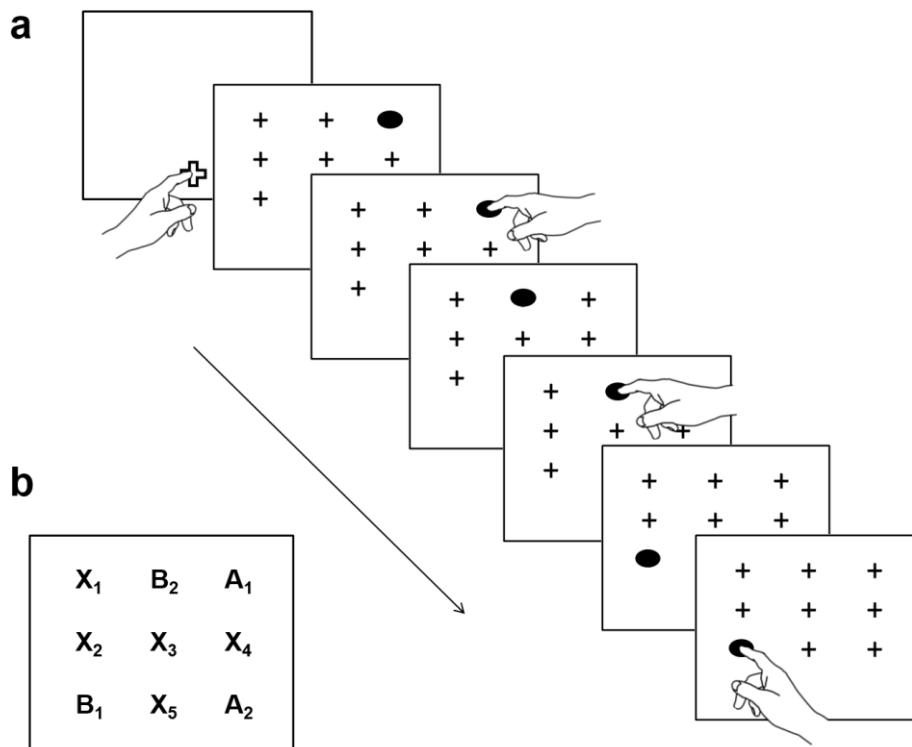


Figure 1: (a) General procedure of the serial response time task in Experiment 1. After touching the fixation cross, participants had to touch the red circle target that appeared successively at three different locations. (b) Locations on the touchscreen. Letters are given for illustrative purposes but were not displayed.

Two types of 3-target sequences were presented. In the first type (“NAD” condition), the first location of the target systematically predicted its last location, leading to a non-adjacent Transitional Probability (TP) of 1. Two pairs of A-B locations were used (A_1 - B_1 and A_2 - B_2), and a variable X location was inserted between A and B locations. The sequences were therefore of the form A_1XB_1 and A_2XB_2 . The full set of 9 possible locations consisted in 4 locations for A_1 , B_1 , A_2 and B_2 , and 5 locations for X elements (i.e., X_1 , X_2 , X_3 , X_4 , X_5 ; see Fig. 1b). These locations were chosen pseudo-randomly among the nine possible spatial locations of the matrix, with the sole constraint that neither A_1 and B_1 , nor A_2 and B_2 , could be located next to each other. As the five X locations were presented at equal frequencies, the second location was uninformative about the third target location. The TP between the first and the second target locations was equal to 0.2, and the TP between the second and the third was equal to 0.5. A total of ten (two A-B pairs * five X locations) NAD sequences were used (see Table 1).

The second type of sequences had an XXB_1 and XXB_2 structure. These sequences were used in the control condition. To equate for the number of sequences per condition, 10 control sequences were constructed, as illustrated in Table 1. TPs between the first and third locations, as well as between the second and third locations, were equal to 0.5 in these sequences. It was therefore impossible to use adjacent or non-adjacent dependencies to predict the third target location. NAD and control sequences were presented at equal frequencies across the experiment. This equates motor practice in the two conditions. Experiment 1 involved a total of twenty 100-trial blocks. Each block contained five randomly intermixed presentations of the 10 NAD and 10 control sequences described above. All baboons were presented with the same sequences. Custom-written Python code (www.python.org) was used to generate the sequences.

Table 1 The 20 sequences presented in Experiment 1.

NAD	Control
A ₁ X ₁ B ₁	X ₃ X ₁ B ₁
A ₁ X ₂ B ₁	X ₅ X ₂ B ₁
A ₁ X ₃ B ₁	X ₂ X ₃ B ₁
A ₁ X ₄ B ₁	X ₁ X ₄ B ₁
A ₁ X ₅ B ₁	X ₄ X ₅ B ₁
A ₂ X ₁ B ₂	X ₃ X ₁ B ₂
A ₂ X ₂ B ₂	X ₅ X ₂ B ₂
A ₂ X ₃ B ₂	X ₂ X ₃ B ₂
A ₂ X ₄ B ₂	X ₁ X ₄ B ₂
A ₂ X ₅ B ₂	X ₄ X ₅ B ₂

Training

The baboons were familiar with the general principle of the task at the beginning of the experiment, as they previously performed a SRT task involving 9-target sequences (Minier et al., 2016). Before the experimental phase described above, the baboons received an initial training phase of twenty 100-trial blocks in which the first, second and third target locations were selected randomly in the matrix. These random sequences were constructed with the constraint to systematically avoid the repetition of any given target location within the 3-target sequence. The aim of the training phase was twofold: (1) to familiarize the baboons with the 3-target sequences (instead of 9 as in Minier et al., 2016), and (2) to limit any long-term effect of the regularities previously learned.

Data analyses

Our set of statistical analyses compared response times on the third target (RTs) in the NAD and control conditions. Remember that the NAD and control sequences were presented at equal frequencies across the experiment. This procedure allowed us distinguishing response time accelerations due to a practice effect, as participants performed the control and NAD sequences an equal number of times, and those attributable to the learning of the predictive relationships present in the NAD sequences. Importantly, the final transitions between the

second and the third target locations were the same across conditions (i.e., X_{1-5-B1} and X_{1-5-B2}). Motor constraints were therefore controlled for that last transition. Shorter RTs in the NAD compared to the control condition would therefore indicate successful learning of the NADs.

For statistical analyses, the data of the exposure phase were grouped in five blocks of 400 trials. For each participant, mean response times on the third target (RTs) was computed for each condition and block. These data were then analyzed using repeated-measures ANOVA involving the Condition (NAD and Control) and Block (1-5) as within-participant factors. Post-hoc comparisons were performed using Tukey Honestly tests ($p < .05$)

On average, the sixteen baboons required five days of testing (range 3-7 days) to complete the experiment. Incorrect trials were removed from the data set (1.6%), as were removed all the trials with response times greater than two standard deviations from the mean (5.0% of the remaining trials, computed for each participant and each block).

Humans

Participants and apparatus

Ten human participants participated in this experiment (five females, age range 19-24). All participants were students of Aix-Marseille University and were paid for their participation. All participants were right-handed, had normal or corrected-to-normal vision, and were naïve as to the goal of the study. Participants performed the task while sited in front of a 19-inch touch screen. The experiment lasted approximately three hours per participant, and was divided in two sessions. Testing sessions were separated by two days at most.

Procedure and data analyses

The task and stimuli were similar in their general principles for humans and for baboons, with only slight differences between species. First, the number of blocks was reduced by half in humans ($n = 10$), in comparison to baboons ($n = 20$). This reduction was motivated by the long duration of the experiment (i.e., three hours with the current design). Second, humans did not receive a food reward at the end of each correct trial, contrary to baboons. The duration of the time-out and inter-trial interval were otherwise similar in the two species. Third, humans were instructed at the beginning of the experiment to follow the red target as quickly and accurately as possible. They were also told that a green screen means that they did not correctly touch the target. Participants were interviewed at the end of the experiment to assess their

declarative knowledge of the regularities embedded in the sequences. They were asked: “Did you notice any regularity, or rule, in some of the sequences?”

Regarding the training phase, human participants were first trained on the SRT task using the same procedure as for baboons, but their number of training trials was reduced to ten 100-trial blocks. Statistical analyses were the same as for baboons, except that blocks of the exposure phase were grouped in five blocks of 200 trials (instead of 400).

Incorrect trials were removed from the data set (7.3%), as well as trials with response times greater than two standard deviations from the mean (3.3% of the remaining trials, computed for each participant and each block).

Results

Baboons

Mean RTs obtained in baboons are reported for each condition and block in Fig. 2a. The repeated measures ANOVA revealed an effect of Condition close to significance level, $F(1, 15) = 4.21, p = .06, \eta_p^2 = .22$. It indicated that the baboons tended to respond faster in the NAD (Mean \pm SD = 400 ± 41 ms) compared to the control condition (404 ± 39 ms). No main effect of Block was found³¹, $F(1, 15) = 1.21, p = .29$, but there was a significant interaction between these two factors, $F(1, 15) = 13.73, p < .01, \eta_p^2 = .48$. Post-hoc analyses revealed that RTs were significantly shorter in NAD compared to the control trials for the last block (Block 5, Cohen's $d = .29$), but not for the previous blocks (1-4). Altogether, the findings suggest that baboons successfully learned the NADs in our task, but that this learning took a substantial number of trials.

³¹ As noted by some reviewers, an increase in RTs for both conditions is observed between Blocks 3 and 4 (see Fig. 2a). We have no definitive explanation to this increase, which can be explained by a series of factors not directly related to the task, such as an interfering social event in the group of baboons, or a change in weather conditions.

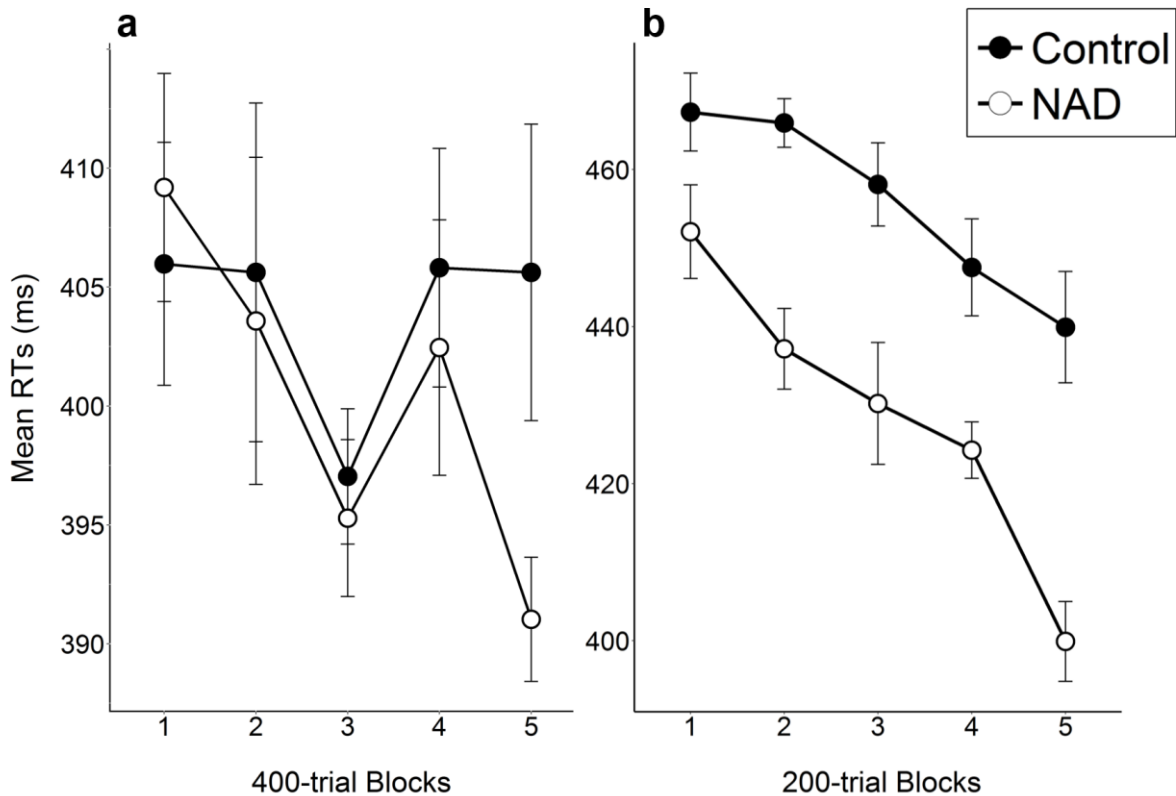


Figure 2. Response times for the third target (RTs) depending on the block, for the control and the non-adjacent dependencies (NAD) conditions in Experiment 1, for baboons (a) and for humans (b). The bars represent standard errors from normalized data.

Humans

Mean RTs are reported for each condition and block in Fig. 2b. Repeated measure ANOVAs revealed a main effect of Condition, $F(1, 9) = 38.60, p < .001, \eta_p^2 = .81$, indicating faster responses in the NAD (429 ± 65 ms) compared to the control condition (456 ± 58). A reliable main effect of Block was found as well, $F(1, 9) = 30.99, p < .001, \eta_p^2 = .77$, as was found a significant Block*Condition interaction, $F(1, 9) = 11.87, p < .01, \eta_p^2 = .57$. Post hoc analyses on this interaction indicated a RTs advantage in the NAD condition compared to the control condition that was significant from the first block ($d_s =$ Block 1: .25, Block 2: .42, Block 3: .39, Block 4: .39, Block 5: .70). These results suggest that humans began to learn the NADs earlier than baboons, during the first block, and that this learning was amplified throughout the course of the experiment.

Post-experiment interviews revealed that none of the participants reported the existence of a predictive relationship between the first and the last locations (i.e., neither A₁-B₁ nor A₂-B₂ relationships were reported). However, 4 out of 10 participants correctly reported that two locations only were used at the last step of the sequence. In addition, four participants could explicitly verbalize some of the triplets (from one to four triplets, depending on the participant). Six of these triplets of locations were NAD sequences, and four were control sequences. To assess whether the results we obtained were triggered by the participants' declarative knowledge of this subset of sequences, we replicated the analysis described above after removing from the dataset the corresponding trials, and these exclusions did not change the pattern of statistical results. Overall, these data suggest that human participants could extract the NADs. They further suggest that while human participants exhibited learning of the regularities, as inferred from their response times, this learning was implicit.

Discussion

In Experiment 1, baboons and humans could extract the NADs, as both species responded faster to the third target when its location was predicted by the first target location (NAD condition) than when it was not (control condition). Fig. 1 suggests that this difference progressively increased in both species, with substantial differences between them. A significant difference in RTs between the two conditions was found from the first block of exposure in humans. This effect only emerged after five blocks in baboons, revealing a substantial difference in the amount of exposure necessary for each species to exhibit successful learning of the NADs. Considering that each of the two NADs was presented 100 times within a 400-trial block, we estimate that baboons demonstrated learning after 400 (100 * 4 blocks) presentations of each NAD. In contrast, this advantage in response times reached significance since the first 200-trial block in humans, meaning that learning took place within the first 50 presentations of these NADs. Moreover, effect size of the difference between the two conditions on the last block was small to medium in baboons ($d = .29$), while it was medium to large in humans ($d = .70$).

However, a discrepancy in task practice between the two species might partly explain these differences between humans and baboons. Indeed, all our baboons had already participated to a study involving a SRT task (Minier et al., 2016), and performed 20 blocks of random sequences before completing the current experiment. In contrast, human participants only performed 10 random blocks and never had been trained to the SRT task before. This

discrepancy in task practice between the two species might explain why humans' response times decreased progressively throughout the experiment in both the control and the NAD conditions, whereas baboons only showed this acceleration in the NAD condition. Importantly, the baboons responded faster on average than humans on the first block (Mean \pm SD = 406 \pm 43 and 460 \pm 62 ms, respectively, $t(24) = 2.58, p = .02, d = .99$), probably as a consequence of their extensive practice with the task, while the two species show comparable RTs on the last block (400 \pm 54 and 420 \pm 56 ms, respectively, $t(24) = 0.89, p = .38$). These data suggest that a ceiling effect might have limited baboon's acceleration in the NAD condition.

One limitation of this first experiment is that it does not rule out the possibility that the participants progressively learned the NAD sequences by rote, or a subset of them, and responded faster at the third location because of this rote learning, but without extracting the NADs embedded in these sequences. Experiment 2 addresses this hypothesis in baboons by testing their ability to generalize the NADs over novel intervening locations.

Experiment 2

Experiment 2 tested whether baboons merely learned the ten A_1XB_1 and A_2XB_2 sequences as wholes, or extracted instead the $A_1_B_1$ and $A_2_B_2$ non-adjacent relationships. Baboons were exposed successively to two different "languages"³² in a within-participant design, and two experimental phases were proposed for each language. The first phase corresponded to the exposure phase of Experiment 1, except that the set of possible locations was increased from 9 to 16 locations. Six of these locations remained unused during that phase. During the second phase, which is referred to as the test phase, the baboons were presented with two types of novel sequences. In half of these sequences, the NADs learnt during the exposure phase were presented over the six locations not used during the exposure phase. The dependencies were inconsistent in the remaining half of the sequences (i.e., A_1 was presented with B_2 and A_2 with B_1). Importantly, the consistent sequences in Language 1 were the inconsistent sequences in Language 2, and *vice-versa*. This two-language design therefore guaranties that any difference in response times between the conditions will reflect the learning and generalization of the NADs, rather than idiosyncratic sequence features related to differences in motor constraints.

³² We borrow here the term "language" to the literature on artificial language learning, the set of sequences being considered here as a set of "words".

Methods

Participants and apparatus

Experiment 2 involved 14 (10 females, age range 4-21 years) out of 16 baboons of Experiment 1, because we were unable to test two of the previous participants for reasons unrelated to our research. The apparatus was the same as in the previous experiment. The general testing procedure was also the same as in Experiment 1, except that the display contained 16 (4x4) possible locations rather than only 9 (3x3).

Procedure

Table 2 provides the full lists of sequences used in Languages 1 and 2. The first exposure phase was identical to Experiment 1 in both languages. In Language 1, sequences took the form A_1XB_1 and A_2XB_2 for the NAD trials (50% of the trials), and XXB_1 and XXB_2 for the control trials (50% of the trials; see Table 2), with six different instantiations of the X locations. In Language 2, the sequences took the form A_1XB_2 and A_2XB_1 for the NAD trials, and control sequences were the same as in Language 1. The remaining six locations (noted Y_{1-6}) were never used during the exposure phase. Locations have been chosen pseudo-randomly, with the sole constraint that A_1 and B_1 , as well as A_2 and B_2 , could not be located next to each other in the matrix. Fig. 3 shows the sixteen locations on the touchscreen that were used in Experiment 2. Note that the locations were the same in both languages, and that only the NADs differed between them.

Baboons then received three types of sequences during the test phase. The first type (50% of the trials) corresponded to the NAD sequences of the exposure phase. For the two other types of trials, sequences were either consistent with the NADs of the exposure phase (“Consistent” condition, 25% of the test trials), or inconsistent with those dependencies (“Inconsistent” condition, 25%). In these two latter conditions, the second location of the sequence was drawn from the set of Y locations never used during the exposure phase. Therefore, sequences were novel in these two conditions. In Language 1, consistent sequences took the form A_1YB_1 and A_2YB_2 , while inconsistent sequences took the form A_1YB_2 and A_2YB_1 . In Language 2, the consistent sequences were the inconsistent sequences of Language 1 (i.e., A_1YB_2 and A_2YB_1), while the inconsistent sequences were the consistent ones of Language 1 (i.e., A_1YB_1 and A_2YB_2). Twelve sequences were constructed for each condition.

NAD, consistent and inconsistent trials were randomly intermixed within each block of the test phase.

Each baboon participated successively to the two language sessions. For each language, the baboons received 40 blocks of 96 trials during the exposure phase, and 20 blocks of 96 trials in the test phase. Note that the amount of exposure changed from 20 blocks in Experiment 1 to 40 in Experiment 2, for both languages, because we hypothesized that learning was still in progress during the last blocks of Experiment 1. To limit potential influences of Language 1 acquisition on the learning of Language 2, six months elapsed between the two language sessions, during which the baboons were exposed to different unrelated tasks. Furthermore, a random phase consisting of 35 blocks of 96 trial blocks was given to the baboons at the beginning of the Language 2 session.

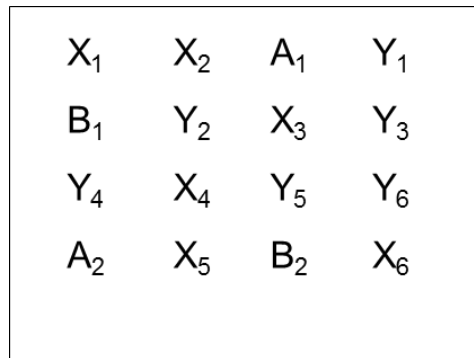


Figure 3. Locations on the touchscreen in Experiment 2. Letters are given for illustrative purposes but were not displayed. Note that the locations were the same for Languages 1 and 2.

Table 2. The 48 sequences presented during Experiment 2, for Language 1 (L1) and Language 2 (L2).

NAD (L1)	NAD (L2)	Control (L1 & L2)	Consistent (L1) / Inconsistent (L2)	Inconsistent (L1) / Consistent (L2)
A ₁ X ₁ B ₁	A ₂ X ₁ B ₁	X ₂ X ₁ B ₁	A ₁ Y ₁ B ₁	A ₂ Y ₁ B ₁
A ₁ X ₂ B ₁	A ₂ X ₂ B ₁	X ₃ X ₂ B ₁	A ₁ Y ₂ B ₁	A ₂ Y ₂ B ₁
A ₁ X ₃ B ₁	A ₂ X ₃ B ₁	X ₆ X ₃ B ₁	A ₁ Y ₃ B ₁	A ₂ Y ₃ B ₁
A ₁ X ₄ B ₁	A ₂ X ₄ B ₁	X ₁ X ₄ B ₁	A ₁ Y ₄ B ₁	A ₂ Y ₄ B ₁
A ₁ X ₅ B ₁	A ₂ X ₅ B ₁	X ₄ X ₅ B ₁	A ₁ Y ₅ B ₁	A ₂ Y ₅ B ₁
A ₁ X ₆ B ₁	A ₂ X ₆ B ₁	X ₅ X ₆ B ₁	A ₁ Y ₆ B ₁	A ₂ Y ₆ B ₁
A ₂ X ₁ B ₂	A ₁ X ₁ B ₂	X ₂ X ₁ B ₂	A ₂ Y ₁ B ₂	A ₁ Y ₁ B ₂
A ₂ X ₂ B ₂	A ₁ X ₂ B ₂	X ₃ X ₂ B ₂	A ₂ Y ₂ B ₂	A ₁ Y ₂ B ₂
A ₂ X ₃ B ₂	A ₁ X ₃ B ₂	X ₆ X ₃ B ₂	A ₂ Y ₃ B ₂	A ₁ Y ₃ B ₂
A ₂ X ₄ B ₂	A ₁ X ₄ B ₂	X ₁ X ₄ B ₂	A ₂ Y ₄ B ₂	A ₁ Y ₄ B ₂
A ₂ X ₅ B ₂	A ₁ X ₅ B ₂	X ₄ X ₅ B ₂	A ₂ Y ₅ B ₂	A ₁ Y ₅ B ₂
A ₂ X ₆ B ₂	A ₁ X ₆ B ₂	X ₅ X ₆ B ₂	A ₂ Y ₆ B ₂	A ₁ Y ₆ B ₂

Data analyses

Two separated sets of analyses were conducted on the exposure and test phases. Blocks were grouped for these analyses in 10 (Exposure) and 5 (Test) blocks of 384 trials each. Learning of NADs during the exposure phase was evaluated with repeated-measures ANOVA performed on RTs and involving the Condition (NAD, Control), Language (1, 2) and Block (1-10) as within-participant factors. Post-hoc comparisons were performed using Tukey Honestly tests ($p < .05$).

Generalization was evaluated in the test phase with repeated-measures ANOVA on RTs involving the Condition (NAD, Consistent, Inconsistent), Language (1, 2) and Block (1-5) as within-participant factors. RTs in the consistent and inconsistent conditions were further compared for the first test block using repeated-measures ANOVA involving the Condition (Consistent, Inconsistent) and Language (1, 2) as within-participant factors. RTs in the consistent and inconsistent conditions were further compared for the first test block using a

repeated-measures ANOVA involving the Condition (Consistent, Inconsistent) and Language (Language 1, Language 2).

We also investigated potential effects of spatial factors on the generalization of NADs. This analysis was conducted with an ANOVA using the Condition (Consistent, Inconsistent) and Pair (A₁-B₁, A₂-B₂, A₁-B₂, A₂-B₁) as within-participant factors, and the RTs obtained in the first test block as dependent variable. An additional set of analyses tested whether generalization occurred for every novel Y location. One-tailed paired t-tests ($p < .05$) were used to assess if RTs were reliably longer in the inconsistent compared to the consistent condition for each Y location.

To complete Language 1 session, the fourteen baboons required 8 days (range 5-12) on average. The same baboons required on average 10 days (range 5-17) of testing to complete Language 2 session. Incorrect trials were removed from the data set (Language 1: 1.7%, Language 2: 2.1%), as well as trials with RTs greater than two standard deviations from the mean (4.9 % and 5.1%, respectively, computed for each participant and each block).

Results

Main Results

Mean RTs are reported in Fig. 4 for each language, phase and condition. For the exposure phase, the main effect of Condition was significant, $F(1, 13) = 16.28, p = .001, \eta_p^2 = .56$, corresponding to faster response times in the NAD (Mean \pm SD = 422 ± 44 ms) than in the control condition (430 ± 43 ms). Also significant was the main effect of Block, $F(1, 13) = 16.88, p = .001, \eta_p^2 = .56$, showing that RTs decreased with practice. Finally, the main effect of Language was also significant, $F(1, 13) = 6.60, p < .01, \eta_p^2 = .34$, indicating faster response times on average for the second (414 ± 32 ms) compared to the first language (438 ± 50 ms). The Condition*Block interaction was close to significance level, $F(1, 13) = 3.21, p = .10, \eta_p^2 = .20$. There were no other interactions (all $ps > .42$). The main effect of condition, along with the absence of Condition*Language interaction, indicate that the baboons were able to learn the NADs in both languages.

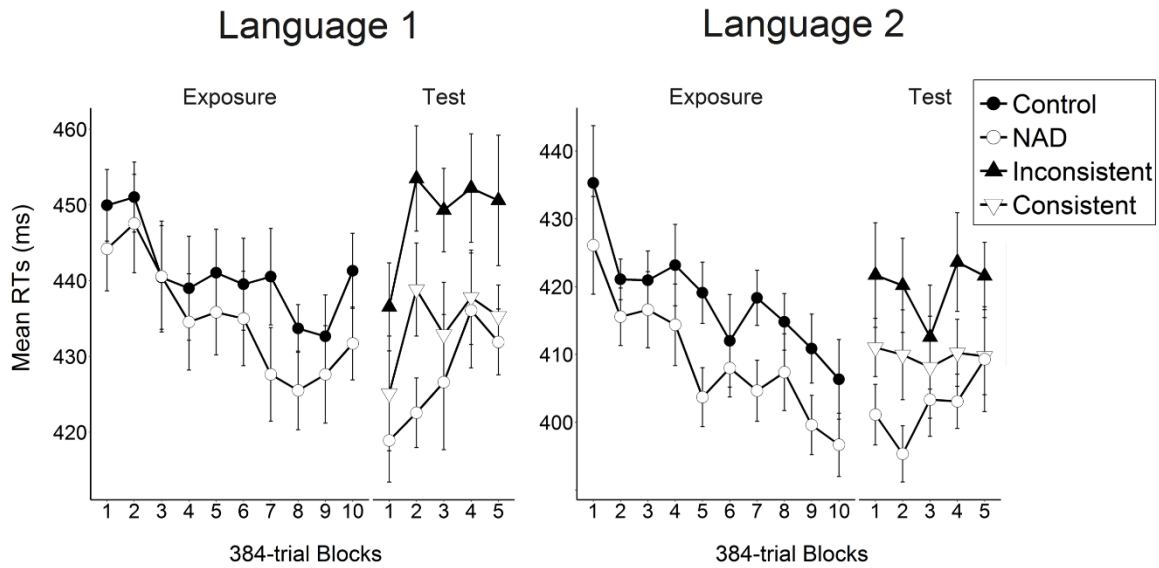


Figure 4. Mean response times (RTs) on the third target obtained in each condition of the exposure and test phases of Experiment 2 and for each language. Bars represent standard errors from normalized data.

Analyses conducted on the test phase revealed a significant effect of Condition, $F(2, 26) = 13.39, p < .001, \eta_p^2 = .51$. Post hoc analyses revealed that the three conditions differed significantly on average from each other. Baboons were faster to respond to the third target in the NAD condition (415 ± 46 ms) than in the two other conditions (Inconsistent condition: 434 ± 47 ms, $d = .41$; Consistent condition: 441 ± 46 ms, $d = .15$). Critically, they were also significantly slower in the inconsistent compared to the consistent condition ($d = .26$). A main effect of Language was found, $F(1, 13) = 8.75, p = .011, \eta_p^2 = .40$, RTs being faster in the second compared to the first language session. No reliable effect of Block was found, $F(1, 13) = 1.52, p = .24$, but the Block*Condition interaction was significant, $F(2, 26) = 3.82, p < .05, \eta_p^2 = .23$, RTs in the NAD condition only progressively increasing throughout the test phase. No other reliable interactions among these factors were found (all $ps > .54$). An ANOVA conducted on the first test block confirmed that a significant difference between the consistent and inconsistent conditions was already present at the onset of that phase, as revealed by a main effect of Condition, $F(1, 13) = 9.53, p < .01, \eta_p^2 = .42$. No other main effect or interaction were found ($ps > .12$).

Spatial factors

An ANOVA revealed a significant interaction between Condition and Pair, $F(3, 39) = 4.56$, $p < .01$, $\eta_p^2 = .26$. Post-hoc tests revealed significantly longer RTs in the inconsistent compared to the consistent condition for two pairs: A₂-B₂ (416 ± 68 and 389 ± 44 ms, respectively, $d = .44$) and A₂-B₁ (467 ± 61 and 438 ± 53 ms, $d = .50$), but not for the A₁-B₁ (458 ± 38 and 463 ± 58 ms) and A₁-B₂ (379 ± 45 and 384 ± 43 ms) pairs. These results suggest that baboons have been able to generalize two out of the four NADs.

Table 3 reports the results of the one-tailed paired t-tests evaluating whether the generalization obtained with the two NADs mentioned above occurred on every novel Y location. RTs were longer in the inconsistent compared to the consistent condition for all Y locations, and this difference was significant for four out of the six Y locations (i.e., Y₂, Y₃, Y₄, Y₆), and approached significance for a fifth one (i.e., Y₅). Therefore, these results suggest that the baboons generalized the NADs over several novel Y locations, with only one exception for Y₁.

Table 3 Response times (Mean \pm SD, ms) and results of the one-tailed paired t-tests assessing generalization over each novel Y locations, for the A₂-B₁ and A₂-B₂ NADs during the first test block. *dfs* were all equal to 13.

	Y ₁	Y ₂	Y ₃	Y ₄	Y ₅	Y ₆
Inconsistent	426 \pm 33	445 \pm 69	434 \pm 37	472 \pm 76	435 \pm 54	436 \pm 49
Consistent	419 \pm 39	391 \pm 40	413 \pm 35	423 \pm 49	419 \pm 65	415 \pm 38
t	0.91	3.51	3.36	2.72	1.32	1.80
Cohen's <i>d</i>	.15	.96	.60	.77	.27	.48
<i>p</i>	.19	<.01	<.01	<.01	.10	<.05

Discussion

Two results were obtained in Experiment 2. First, baboons successfully learned the NADs during the exposure phase, replicating the findings already obtained in Experiment 1. Second, for two NADs response times were slower for inconsistent than for consistent test trials. This finding suggests generalization of these NADs over the novel intervening locations. Importantly, these results suggesting generalization cannot be accounted for by learning of the consistent sequences, since each beginning of a sequence (i.e., A₁-Y₁₋₆ and A₂-Y₁₋₆) was

followed by a B₁ or B₂ location at equal frequency within each language session. Analyses moreover confirmed that the speed advantage for consistent sequences was already present at the onset of the test phase.

Interestingly, reliable generalization was only observed for the two NADs beginning by A₂ which was located in the lower-left corner of the matrix (see Fig. 3). By contrast, the baboons failed to properly generalize the two NADs beginning at the A₁ location, which was not located in a corner. This result suggests that the saliency of the predictive target had an impact on baboons' ability to generalize NADs in this task. We also note that generalization of the above two NADs was not observed for the only Y element located in a corner of the matrix (Y₁), possibly because of its greater saliency, or eccentricity in the matrix, compared to the other Y locations which were more central. Taken together, these findings suggest that the baboons can generalize NADs over novel locations, but that this ability may be hindered by spatial factors.

General discussion

In Experiment 1, humans and baboons responded faster to the last target in the NAD compared to the control sequences. Importantly, participants did not merely learn that the sequences always finished by one of two locations (with a probability of .5 for each), but processed the conditional probabilities of the third location, given the first location. This result was replicated in baboons in Experiment 2. These data overall provide evidence of baboons' ability to learn NADs. However, while humans presented a significant difference in response times between the two conditions within the first exposure block, this difference appeared after approximately 400 presentations of each NAD in baboons. This difference in learning speed differs from what has been found for adjacent dependencies (Minier et al., 2016; Rey et al., submitted). In these earlier studies, baboons and humans learned at similar speed (in less than 200 trials) the predictive relationships between two locations presented at the end of three different triplets. This suggests that learning of NADs is more difficult for baboons than humans, while these two species seem equally fast for learning adjacent dependencies. Importantly, the amount of exposure required by the baboons to extract NADs in the current experiment is in the same range as previously found with habituation-dishabituation paradigms in other primate species (e.g., 672 presentations in Newport et al., 2004). This amount may serve as a reference for designing future experiments investigating related topics.

For two NADs we found that when presented with novel intervening locations the baboons were slower to respond when the sequences were inconsistent with the learned NADs

than when the dependencies were preserved. To our knowledge, this result is the first demonstration in a non-human species of a successful generalization of sequentially non-adjacent dependencies over novel intervening elements. The current results therefore extend to the case of sequential NADs Sonnweber et al.'s findings (2015), in which generalization of NADs was tested between visual shapes presented concurrently. Along with previous findings in humans (e.g., Frost & Monaghan, 2016), our study indicates that this ability is shared across different primate species. However, successful generalization of the NADs in the current SRT study appeared to be modulated by spatial features of the stimuli. Indeed, the baboons only generalized the two NADs whose predictive target was located in a corner of the matrix and generalization was not observed for one of the novel intervening location which was also located in a corner. These data present some similarities with previous findings in the auditory domain. For instance, Endress, Scholl, & Mehler (2005, see also Peña et al., 2002) reported that human participants could properly extract some specific regularities from auditory sequences only when the relevant stimuli were presented at an edge of the sequences. Moreover, several studies manipulating the phonological parameters of the stimuli suggested that the relative saliency of the non-adjacent and intervening elements impacts the extraction of NADs (Creel et al., 2004; Onnis et al., 2005; Toro & Trobalón, 2005). Taken together, these findings suggest (1) that perceptual factors constrain the extraction of NADs, and (2) that this effect occurs in both the auditory and visuo-spatial domains.

The exact nature of the processes involved in NADs extraction and recognition is still debated in the literature, and at least three theories are proposed. According to a first theory, the mechanisms that support NADs extraction are the same associative mechanisms as those allowing the learning of adjacent dependencies (Perruchet, Tyler, Galland, & Peereman, 2004b). This single-process account proposes that the joint attentional processing of two elements is a necessary and sufficient condition to associate them, even when these are distant in space or time (Pacton & Perruchet, 2008). A second and complementary theory proposes that the extraction of adjacent and non-adjacent dependencies involves similar statistical learning mechanisms, but emphasize the role played by working memory for NADs extraction (Conway, Deocampo and Smith, 2016; Wilson et al., 2017). According to this dual-system theory, the detection of adjacent relationships would be sustained by modality-specific processes involving perceptual-motor neural networks (Conway et al., 2016), or ventral regions of the frontal cortex (Wilson, Kikuchi, et al., 2015), while the extraction of non-adjacent relationships would additionally recruit more dorsal areas (Wilson et al., 2017). According to a last hypothesis

(Endress & Bonatti, 2007; Endress, Nespors, & Mehler, 2009; Peña et al., 2002), humans and possibly other animals (Endress, Carden, Versace, & Hauser, 2010) may be endowed with some rule-based mechanisms dedicated to fast learning of “classes” of elements occurring at the edges of sequences (e.g., here, learning that A_1 and A_2 always occur in the first position), as well as statistical learning mechanisms for the computation of transitional probabilities (e.g. $A_1_B_1$). The current experiment did not aim at assessing the edge-based positional learning hypothesis, but examined the extraction time course of those transitional probabilities.

In humans, learning occurred early in the experiment, and was strengthened throughout the entire exposure phase. In baboons, learning of the NADs occurred later in the experiment, but a similar progressive emergence was observed. In previous SRT experiments, we found that the learning of adjacent dependencies in triplets (Minier et al., 2016), or pairs of shapes (Joël Fagot, Malassis, & Medam, 2017) gave rise to similar progressive (albeit earlier) decrease in response times throughout the exposure to these regularities. These data, along with previous findings (Frost & Monaghan, 2016; Romberg & Saffran, 2013; Vuong et al., 2016), suggest that adjacent and non-adjacent dependencies might be extracted by similar statistical learning mechanisms. However, they do not allow disentangling between the single-process and the dual-system proposals described above. Indeed, the difference in speed learning that is observed between baboons and humans might reflect an involvement of working memory in NADs extraction, as working memory capacity differ between these species (Joël Fagot & De Lillo, 2011b). However, addressing this question would require additional behavioral and brain imaging studies. Regarding behavioral studies, correlational designs in humans and non-human primates would allow examining individual differences in statistical learning ability across different types of dependencies (i.e., adjacent and non-adjacent, Siegelman, Bogaerts, Christiansen, & Frost, 2017; Siegelman & Frost, 2015), as well as exploring the relationships between working memory capacity and NADs extraction. Comparative neuroimaging studies have already suggested that adjacent dependencies are processed by evolutionary conserved brain structures, including the frontal operculum and anterior insula (Wilson, Kikuchi, et al., 2015). However, further researches are also needed in this domain to discover whether NADs processing is sustained by similar or different neural substrates in human and non-human primates (Milne et al., 2016; Wilson et al., 2017).

Another important question is the domain generality of the mechanisms involved during the learning of NADs. For instance, are the extraction of NADs in the current SRT task and the extraction of NADs in natural languages (such as in *is reading*) supported by the same learning mechanisms? Evidences accumulated so far reveal a contrasted pattern (R. Frost, Armstrong,

Siegelman, & Christiansen, 2015a). On the one hand, individuals with specific language impairment have been found to exhibit poor performance in auditory statistical learning tasks (e.g., Evans, Saffran, & Robe-Torres, 2009) as well as in SRT tasks (Lum, Conti-Ramsden, Morgan, & Ullman, 2014), compared with typically-developing individuals. They also express greater difficulties in processing NADs in an AGL-SRT task (Hsu, Tomblin, & Christiansen, 2014a). In another study, Misyak, Christiansen, and Tomblin (2010) found a positive correlation between typically-developing individuals' ability to process long-distance dependencies from a natural language and NADs in an AGL-SRT task. On the other hand, statistical learning appears to be subject to modality and stimulus specificity (e.g., Conway & Christiansen, 2005, 2006). Siegelman and Frost (2015) reported for instance that the performance of their participants did not correlate between any of four statistical learning tasks they have used (auditory-verbal, auditory-nonverbal, visual, and SRT tasks). These results suggest the existence of modality and stimulus specific constraints on statistical learning mechanisms, which prevent drawing strong conclusions regarding the similarity between NADs extraction in natural languages and in our visuo-spatial task.

From an evolutionary standpoint, the ability to keep track of predictive relationships between non-adjacent events might present various functional advantages, for instance for planning complex behavioral or motor sequences, or for monitoring social interactions and their outcomes (Sonnweber et al., 2015). In baboons, this process may allow for instance to detect that a threat behavior from an individual A is followed by a scream from an individual B, even if other individuals emit unrelated vocalizations in between. One general advantage of the learning of NADs might also be a reduction in memory demands. For instance in song birds, admitting an optional or variable song element between two others (thereby being non-adjacent) could limit the number of pairwise transitions needed to be memorized (Petkov & Wilson, 2012). In our experiments, the extraction of two NADs may be cognitively less demanding than the learning and storage of ten sequences.

Acknowledgments

This work, carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX). This research was supported by the Premilang2 ANR-project (#ANR-13-BSH2-0002) and the Chunked ANR-project (#ANR-17-CE28-0013-02).

References

- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology. General*, *120*(3), 235-253.
- Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences*, *5*(12), 539-546. [https://doi.org/10.1016/S1364-6613\(00\)01800-3](https://doi.org/10.1016/S1364-6613(00)01800-3)
- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(1), 24-39. <https://doi.org/10.1037/0278-7393.31.1.24>
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science*, *17*(10), 905-912. <https://doi.org/10.1111/j.1467-9280.2006.01801.x>
- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1119-1130. <https://doi.org/10.1037/0278-7393.30.5.1119>
- de la Mora, D. M., & Toro, J. M. (2013). Rule learning over consonants and vowels in a non-human animal. *Cognition*, *126*(2), 307-312. <https://doi.org/10.1016/j.cognition.2012.09.015>
- Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica*, *135*(2), 182-190. <https://doi.org/10.1016/j.actpsy.2010.06.005>
- Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*(2), 247-299. <https://doi.org/10.1016/j.cognition.2006.09.010>
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal Cognition*, *13*(3), 483-495. <https://doi.org/10.1007/s10071-009-0299-8>
- Endress, A. D., Nespors, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, *13*(8), 348-353. <https://doi.org/10.1016/j.tics.2009.05.005>

- Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, *134*(3), 406-419. <https://doi.org/10.1037/0096-3445.134.3.406>
- Endress, A. D., & Wood, J. N. (2011). From movements to actions: two mechanisms for learning action sequences. *Cognitive Psychology*, *63*(3), 141-171. <https://doi.org/10.1016/j.cogpsych.2011.07.001>
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical Learning in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *52*(2), 321-335. [https://doi.org/10.1044/1092-4388\(2009/07-0189\)](https://doi.org/10.1044/1092-4388(2009/07-0189))
- Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: use of a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio papio*). *Behavior Research Methods*, *42*(2), 507-516. <https://doi.org/10.3758/BRM.42.2.507>
- Fagot, J., & De Lillo, C. (2011). A comparative study of working memory: Immediate serial spatial recall in baboons (*Papio papio*) and humans. *Neuropsychologia*, *49*(14), 3870-3880. <https://doi.org/10.1016/j.neuropsychologia.2011.10.003>
- Fagot, J., Malassis, R., & Medam, T. (2017). The processing of positional information in a two-item sequence limits the emergence of symmetry in baboons (*Papio papio*), but not in humans (*Homo sapiens*). *Learning & Behavior*. <https://doi.org/10.3758/s13420-017-0290-1>
- Fagot, J., & Paleressompouille, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, *41*(2), 396-404. <https://doi.org/10.3758/BRM.41.2.396>
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499-504.
- Friederici, A. D. (2004). Processing local transitions versus long-distance syntactic hierarchies. *Trends in Cognitive Sciences*, *8*(6), 245-247. <https://doi.org/10.1016/j.tics.2004.04.013>
- Froehlich, A. L., Herbranson, W. T., Loper, J. D., Wood, D. M., & Shimp, C. P. (2004). Anticipating by Pigeons Depends on Local Statistical Information in a Serial Response Time Task. *Journal of Experimental Psychology: General*, *133*(1), 31-45. <https://doi.org/10.1037/0096-3445.133.1.31>

- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117-125. <https://doi.org/10.1016/j.tics.2014.12.010>
- Frost, R. L. A., & Monaghan, P. (2016). Simultaneous segmentation and generalisation of non-adjacent dependencies from continuous speech. *Cognition, 147*, 70-74. <https://doi.org/10.1016/j.cognition.2015.11.010>
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review, 16*(3), 486-490. <https://doi.org/10.3758/PBR.16.3.486>
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*(5), 431-436.
- Gómez, R., & Maye, J. (2005). The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy, 7*(2), 183-206. https://doi.org/10.1207/s15327078in0702_4
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition, 78*(3), B53-B64. [https://doi.org/10.1016/S0010-0277\(00\)00132-3](https://doi.org/10.1016/S0010-0277(00)00132-3)
- Hsu, H. J., Tomblin, J. B., & Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Frontiers in Psychology, 5*, 175. <https://doi.org/10.3389/fpsyg.2014.00175>
- Kovács, Á. M., & Mehler, J. (2009). Flexible Learning of Multiple Speech Structures in Bilingual Infants. *Science, 325*(5940), 611-612. <https://doi.org/10.1126/science.1173947>
- Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 51*(100), 1-10. <https://doi.org/10.1016/j.cortex.2013.10.011>
- Marchetto, E., & Bonatti, L. L. (2013). Words and possible words in early language acquisition. *Cognitive Psychology, 67*(3), 130-150. <https://doi.org/10.1016/j.cogpsych.2013.08.001>

- Marchetto, E., & Bonatti, L. L. (2015). Finding words and word structure in artificial speech: the development of infants' sensitivity to morphosyntactic regularities. *Journal of Child Language*, 42(4), 873-902. <https://doi.org/10.1017/S0305000914000452>
- Milne, A. E., Mueller, J. L., Männel, C., Attaheri, A., Friederici, A. D., & Petkov, C. I. (2016). Evolutionary origins of non-adjacent sequence processing in primate brain potentials. *Scientific Reports*, 6, 36259. <https://doi.org/10.1038/srep36259>
- Minier, L., Fagot, J., & Rey, A. (2016). The Temporal Dynamics of Regularity Extraction in Non-Human Primates. *Cognitive Science*, 40(4), 1019-1030. <https://doi.org/10.1111/cogs.12279>
- Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J. (2010). Sequential Expectations: The Role of Prediction-Based Learning in Language. *Topics in Cognitive Science*, 2(1), 138-153. <https://doi.org/10.1111/j.1756-8765.2009.01072.x>
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127-162.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, 49(2), 85-117. <https://doi.org/10.1016/j.cogpsych.2003.12.002>
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1-32. [https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)
- Onnis, L., Monaghan, P., Christiansen, M. H., & Chater, N. (2005). Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. In *PROCEEDINGS OF THE TWENTY-SIXTH ANNUAL CONFERENCE OF THE COGNITIVE SCIENCE SOCIETY*.
- Onnis, Luca, Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, 53(2), 225-237. <https://doi.org/10.1016/j.jml.2005.02.011>
- Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 34(1), 80-96. <https://doi.org/10.1037/0278-7393.34.1.80>

- Pacton, S., Sobaco, A., & Perruchet, P. (2015). Is an attention-based associative account of adjacent and nonadjacent dependency learning valid? *Acta Psychologica*, *157*, 195-199. <https://doi.org/10.1016/j.actpsy.2015.03.002>
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science (New York, N.Y.)*, *298*(5593), 604-607. <https://doi.org/10.1126/science.1072901>
- Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: no need for algebraic-like computations. *Journal of Experimental Psychology. General*, *133*(4), 573-583. <https://doi.org/10.1037/0096-3445.133.4.573>
- Petkov, C. I., & Wilson, B. (2012). On the pursuit of the brain network for proto-syntactic learning in non-human primates: conceptual issues and neurobiological hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 2077-2088. <https://doi.org/10.1098/rstb.2012.0073>
- Ravignani, A., & Sonnweber, R. (2017). Chimpanzees process structural isomorphisms across sensory modalities. *Cognition*, *161*, 74-79. <https://doi.org/10.1016/j.cognition.2017.01.005>
- Ravignani, A., Sonnweber, R.-S., Stobbe, N., & Fitch, W. T. (2013). Action at a distance: dependency sensitivity in a New World primate. *Biology Letters*, *9*(6), 20130852. <https://doi.org/10.1098/rsbl.2013.0852>
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855-863. [https://doi.org/10.1016/S0022-5371\(67\)80149-X](https://doi.org/10.1016/S0022-5371(67)80149-X)
- Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley interdisciplinary reviews. Cognitive science*, *1*(6), 906-914. <https://doi.org/10.1002/wcs.78>
- Romberg, A. R., & Saffran, J. R. (2013). All Together Now: Concurrent Learning of Multiple Structures in an Artificial Language. *Cognitive Science*, *37*(7), 1290-1320. <https://doi.org/10.1111/cogs.12050>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science (New York, N.Y.)*, *274*(5294), 1926-1928.
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal*

- Society of London. Series B, Biological Sciences*, 372(1711).
<https://doi.org/10.1098/rstb.2016.0059>
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105-120.
<https://doi.org/10.1016/j.jml.2015.02.001>
- Sonnweber, R., Ravignani, A., & Fitch, W. T. (2015). Non-adjacent visual dependency learning in chimpanzees. *Animal Cognition*, 18(3), 733-745.
<https://doi.org/10.1007/s10071-015-0840-x>
- Takahasi, M., Yamada, H., & Okanoya, K. (2010). Statistical and Prosodic Cues for Song Segmentation Learning by Bengalese Finches (*Lonchura striata* var. *domestica*). *Ethology*, 116(6), 481-489. <https://doi.org/10.1111/j.1439-0310.2010.01772.x>
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1598), 1984-1994.
<https://doi.org/10.1098/rstb.2012.0055>
- Toro, J. M., Nespore, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream: functional differences between vowels and consonants. *Psychological Science*, 19(2), 137-144. <https://doi.org/10.1111/j.1467-9280.2008.02059.x>
- Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, 67(5), 867-875.
<https://doi.org/10.3758/BF03193539>
- van Heijningen, C. A. A. van, Visser, J. de, Zuidema, W., & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*.
<https://doi.org/10.1073/pnas.0908113106>
- Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2016). Concurrent Statistical Learning of Adjacent and Nonadjacent Dependencies. *Language Learning*, 66(1), 8-30.
<https://doi.org/10.1111/lang.12137>
- Wilson, B., Kikuchi, Y., Sun, L., Hunter, D., Dick, F., Smith, K., ... Petkov, C. I. (2015). Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. *Nature Communications*, 6, 8901.
<https://doi.org/10.1038/ncomms9901>

- Wilson, B., Marslen-Wilson, W. D., & Petkov, C. I. (2017). Conserved Sequence Processing in Primate Frontal Cortex. *Trends in Neurosciences*, 0(0).
<https://doi.org/10.1016/j.tins.2016.11.004>
- Wilson, B., Slater, H., Kikuchi, Y., Milne, A. E., Marslen-Wilson, W. D., Smith, K., & Petkov, C. I. (2013). Auditory Artificial Grammar Learning in Macaque and Marmoset Monkeys. *The Journal of Neuroscience*, 33(48), 18825-18835.
<https://doi.org/10.1523/JNEUROSCI.2414-13.2013>
- Wilson, B., Smith, K., & Petkov, C. I. (2015). Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *The European Journal of Neuroscience*, 41(5), 568-578. <https://doi.org/10.1111/ejn.12834>

III. ETUDE II – Extraction de Régularités Structurelles chez le Babouin : Contraster Répétitions Adjacentes et Non-adjacentes

Malassis, R., & Fagot, J. (*under revision*). Extraction of Structural Regularities by Baboons (*Papio papio*): Adjacent and Non-Adjacent Repetition Patterns Differ in Learnability.

L'extraction de patterns relève d'un niveau de difficulté supplémentaire car cette opération ne se limite pas à l'apprentissage d'associations fixes entre les éléments, mais implique la généralisation du pattern appris à tous nouveaux éléments entretenant les mêmes relations structurelles. Deux patterns, tous deux basés sur des répétitions, ont été majoritairement utilisés à ce jour pour l'étude de ce phénomène : AAB (ou ABB), qui implique une répétition adjacente, et ABA, avec une répétition non-adjacente (e.g. Marcus, Vijayan, Rao, & Vishton, 1999). Classiquement, les animaux sont entraînés à discriminer ces deux patterns dans une tâche à choix forcé ou une tâche « Go/ No-go » (e.g. Spierings & ten Cate, 2016; van Heijningen et al., 2013). Leur capacité de généralisation de cet apprentissage à des séquences composées de nouveaux stimuli (e.g., CDC & CCD) est ensuite évaluée.

Chez le bébé humain, la capacité à détecter des répétitions adjacentes apparaît plus tôt au cours de l'ontogénèse que celle à détecter des répétitions non-adjacentes (Johnson et al., 2009; Gervain et al., 2008), suggérant qu'elles constituent une régularité plus facile à extraire (voir Chapitre II, section V.1 du volet théorique). A ma connaissance, aucune expérience comparative n'a étudié cette question chez l'animal NH. Nous avons donc cherché dans cette étude à évaluer si les performances du babouin différaient entre ces deux types de pattern.

Pour cette étude, nous avons conçu une nouvelle tâche, consistant en une adaptation de la tâche de pointage sériel dans laquelle sont présentées des séquences de formes visuelles. Cette tâche a également été utilisée pour une autre étude ; fruit d'une collaboration au sein de l'équipe (Joël Fagot, Malassis, & Medam, 2018) ; en annexe p. 310).

Dans la première expérience, les babouins étaient exposés de manière répétée soit à un pattern AA, soit à un pattern ABA. Ils étaient ensuite testés (1) avec des séquences congruentes constituées de nouvelles formes visuelles, et (2) avec des séquences non congruentes (de la

forme AB et ABC, respectivement). Les résultats de cette expérience révèlent que ces babouins ont été capables d'apprendre les deux types de patterns. Ils sont parvenus de plus à généraliser cet apprentissage à de nouveaux stimuli. Toutefois, l'extraction du pattern à répétition immédiate a donné lieu à de meilleures performances chez le babouin que celle du pattern à répétition non-adjacente.

Dans une seconde expérience, les babouins ont été exposés à des séquences de la forme ABB, ABA et ABC, présentées à fréquences égales et aléatoirement mélangées. Cette expérience avait pour but de tester si les babouins se révélaient spontanément sensibles aux répétitions dans ces conditions. Les résultats révèlent que les babouins répondent plus rapidement aux répétitions adjacentes du pattern ABB, tandis qu'aucune différence n'est observée entre les patterns ABA et ABC. Ces données confirment le caractère particulièrement saillant des répétitions adjacentes.

Extraction of Structural Regularities by Baboons (*Papio papio*):

Adjacent and Non-adjacent Repetition Patterns Differ in Learnability

Raphaëlle Malassis and Joël Fagot

Laboratory of Cognitive Psychology, University of Aix-Marseille, CNRS, France

Corresponding author: raphaelle.malassis@univ-amu.fr

Abstract

Several animal species can discriminate sequential patterns containing repeated items (e.g., ABB *versus* ABA). This achievement suggests an ability to extract the structural properties of the sequences, but it remains possible that the subjects process only the presence/absence of an immediate repetition within the sequences, rather than the whole pattern. We studied this question with 17 Guinea baboons (*Papio papio*), using a novel procedure inspired from the serial response time task. Unlike previous comparative research, Experiment 1 compared adjacent and non-adjacent repetition patterns to random sequences. Baboons showed faster responses to isomorphic adjacent and non-adjacent repetition sequences made of novel items, compared to random sequences, demonstrating successful generalization of both patterns. Sequences of the form ABB, ABA and ABC were presented at equal frequencies in Experiment 2. Baboons showed faster response times in the case of adjacent repetitions, while remaining stably slower in other conditions. This result confirms that adjacent repetitions are a particularly salient feature. Overall, these findings suggest that the mechanisms underlying the extraction of the structural properties of sequences, as well as those allowing fast detection of adjacent repetitions, two important abilities for language acquisition, are shared by human and non-human primates.

Key words: rule learning, artificial grammar learning, serial response time task, monkey

When reading the following three-syllable sequences, « ga ti ga » and « li na li », we can memorize each of them individually, but can as well notice that they share a common structure: their first and last syllables are identical while their second syllable differs. The extraction of such « ABA » pattern requires the processing of the structural properties of the sequences, beyond the exemplars, allowing generalization to novel, isomorphic, sequences. In humans, sensitivity to the structural organization of words within sentences has been proposed as one of the core mechanisms underlying the formation of grammatical knowledge (e.g., Gomez & Gerken, 1999; Marcus, Vijayan, Rao, & Vishton, 1999). The comparative study of this ability therefore offers an entry into the assessment of non-human animals' syntactic abilities.

The extraction of sequential patterns has been extensively studied in laboratories using AAB (or ABB) and ABA patterns, which correspond to two simple grammars based on repetitions. In the AAB (or ABB) pattern, there is an immediate repetition of an item within the triplet (« adjacent repetition patterns »). In the ABA pattern, the first and last elements of the triplets are identical, while the element in the middle differs (« non-adjacent repetition pattern »). The grammars underlying those patterns are not as complex as those found in human language. They generate however sequences of minimal length, and are based on simple same/different relations which can be recognized by a wide range of species (e.g., by bees, Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001). These patterns are therefore perfect candidates to comparatively assess structure processing in human adults, infants, and non-human animals.

In a seminal experiment conducted by Marcus, Vijayan, Rao, and Vishton (Experiments 1 and 2, 1999), seven month old infants were exposed to triplets of syllables following either an ABB (e.g., « ga ti ti ») or an ABA pattern (« ga ti ga »). They were then tested with pattern-following sequences, and pattern-violating sequences made of novel syllables (of the form ABA for the group exposed to ABB sequences, e.g., « wo fe wo », and *vice-versa*). Infants from both groups paid more attention to the pattern-violating sequences compared to the pattern-following ones, suggesting that they have been able to learn and generalize the underlying structures.

An extensive number of studies were conducted in non-human animals following Marcus et al. (1999). Most of these studies used operant conditioning procedures. Typically, subjects are first trained to discriminate two different patterns in a Go/No-go (e.g., Murphy,

Mondragón, & Murphy, 2008) or a two-alternative forced choice task (e.g., Santolin, Rosa-Salva, Regolin, & Vallortigara, 2016). Individuals reaching the learning criterion are then tested on novel exemplars of the two patterns. Importantly, we can distinguish two types of generalization tests. In the first type, the novel sequences consist in new combinations of the elements already used in the training sequences. The second type of generalization test uses sequences entirely made of novel elements.

The use of novel combinations of the training elements revealed successful generalization in rhesus macaques (*Macaca Mulatta*, Hauser & Glynn, 2009). However, one problem with this testing procedure is that it does not rule out the possibility that the participants process some surface features of the sequences, such as a greater perceptual similarity between the training and the novel consistent sequences, compared with the inconsistent ones, rather than the structural properties of the sequences (Beckers, Berwick, Okanoya, & Bolhuis, 2016; ten Cate, 2014). Recently, Spierings and ten Cate (2016) introduced in their design the control of one of these surface cues. They first trained budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*) to discriminate auditory sequences following AAB and ABA patterns. In a series of tests presenting novel combinations of familiar elements, they controlled for the number of elements in the novel sequences which shared their position with the training sequences. The behavior of Zebra finches was controlled by the position of the items rather than the sequences structures (for similar results obtained with this species, see Chen, van Rossum, & ten Cate, 2015, Experiment 1). Budgerigars, in contrast, correctly classified the novel sequences, even when the number of shared positions was greater in the inconsistent sequences compared to consistent sequences. In another transfer test, budgerigars were moreover able to generalize their behavior to isomorphic sequences constructed with novel stimuli.

The use of sequences made of novel elements revealed successful generalization in several other species. Rats could discriminate an AAB pattern from random ABC sequences, with the A and B items being either vowels (Experiment 1) or consonants (Experiment 2) presented in auditory sequences, and transferred this knowledge to novel phonemes (*Rattus norvegicus*, de la Mora & Toro, 2013; see also Murphy et al., 2008). Chicks successfully discriminated ABA and ABB patterns, as well as ABB and AAB patterns, in a task relying on a non-sequential presentation of strings of visual shapes (*Gallus gallus*, Santolin, Rosa-Salva, Regolin, & Vallortigara, 2016; see also Versace, Spierings, Caffini, ten Cate, & Vallortigara, 2017). In an experiment using a similar procedure, two chimpanzees correctly generalized an

ABⁿA pattern to novel colors and shapes (with one to five intervening B elements; *Pan troglodytes*, Sonnweber, Ravignani, & Fitch, 2015). These chimpanzees also exhibited interferences between patterns presented in the visual and auditory modalities, suggesting cross-modal generalization of these patterns (Ravignani & Sonnweber, 2017).

All these results suggest that the ability to discriminate different patterns, based on non-adjacent repetition (ABA), initial adjacent repetition (AAB), and final adjacent repetition (ABB), is widespread among animal species. However, one remaining alternative account of these results is that participants discriminated these sequences by considering the presence/absence of an immediate repetition within the sequence, instead of processing the structural properties of the whole sequences (ten Cate & Okanoya, 2012; Corballis, 2009). Indeed, one procedural difficulty with the previous comparative studies is that most of them only compared the responses to adjacent and nonadjacent patterns, without using a third control condition involving for instance random ABC sequences. When the test only uses ABA and AAB (or ABB) sequences, the sole detection of an adjacent repetition can lead to successful discrimination. This is especially the case when using conditioning tasks, as participants can perform equally well by avoiding the S- or choosing the S+ (ten Cate & Okanoya, 2012). We present below three sets of evidence suggesting that adjacent repetitions are particularly salient cues for both human and non-human animals.

First, Gervain, Macagno, Cogoi, Pena, and Mehler (2008) used optical brain imaging techniques to test in newborns the discrimination between adjacent and non-adjacent repetition patterns against random ABC pattern. In this study, the neural activity elicited by ABB sequences containing adjacent repetitions was greater from the start of the experiment in certain regions of interest, including auditory and left frontal areas, and increased with repeated exposure, while it remained lower and stable for the ABC sequences (see also Bouchon, Nazzi, & Gervain, 2015). In another experiment, neural activity did not differ between ABA and ABC sequences (Gervain et al., 2008, Experiment 2), suggesting that the ability to detect adjacent and non-adjacent repetitions follow different developmental trajectories, the first one only being available at birth in humans.

Second, studies conducted in older infants further support the hypothesis that patterns containing adjacent repetition are easier to detect than nonadjacent repetition patterns. Successful generalization of both types of patterns was observed in seven month-olds with

sequences composed of syllables (Gervain & Werker, 2013; Marcus et al., 1999). However, the extraction of non-adjacent repetition patterns made of non-linguistic auditory stimuli, such as tones, appears more difficult than the extraction of adjacent repetition patterns. It requires for instance to present the tones simultaneously with visual stimuli following the same pattern (Thiessen, 2012). Moreover, when presented with sequences of visual shapes, eleven month-old infants fail to learn an ABA pattern, while they succeed with AAB/ABB repetition patterns (Johnson et al., 2009, however see Saffran, Pollak, Seibel, & Shkolnik, 2007 for positive results with more familiar visual stimuli). In addition, after exposure to both ABB and ABA sequences randomly interleaved, twelve-month-old bilingual infants could generalize both patterns, while monolinguals could only generalize the adjacent repetition pattern (Kovács & Mehler, 2009).

Finally, studies on pattern learning in animals also suggest that the extraction of adjacent repetition patterns is easier than the extraction of non-adjacent repetition ones. Using a Go/No-go task, van Heijningen et al. (2013) trained zebra finches to respond to ABA (S+), but not to ABB/AAB (S-) sequences of song elements. The finches responded by attending to the presence/absence of adjacent repetitions (Experiment 1), or by rote memorization of the training triplets when they were prevented to use this first strategy (Experiment 2). For instance, in Experiment 1 the finches treated ABBA sequences as S- stimuli, but treated ABAB sequences as S+ stimuli, showing that the presence of adjacent repetitions within the sequences controlled their behavior. Additional evidence comes from a study aimed at testing finches' ability to master the so-called "center-embedded" grammars (Heijningen, Visser, Zuidema, & ten Cate, 2009). In this experiment, finches were trained to discriminate sequences of song elements of the form $A^n B^n$ from sequences of the form $(AB)^n$. Probe tests revealed that individuals relied on the detection of either "AA" or "BB" adjacent repetitions to perform the task, instead of processing the full underlying grammar.

Overall, these data suggest that animals, including humans, are particularly prone to attend to and exploit adjacent repetitions. The aim of the present experiments was to determine if baboons can generalize repetition patterns in test conditions in which the presence/absence of an adjacent repetition does not allow an accurate discrimination of the different patterns. This was achieved in Experiment 1 by comparing baboons' responses to (adjacent and non-adjacent) repetition patterns to their responses to random sequences that contain no repetition, after exposure to these patterns. In Experiment 2, adjacent repetition (ABB), non-adjacent

repetition (ABA) and random (ABC) patterns were presented at equal frequencies. This procedure aimed at assessing whether baboons would detect repetitions without training.

Experiment 1

Experiment 1 aimed at testing baboons' ability to learn and generalize an adjacent (AA) and a non-adjacent repetition pattern (ABA) by contrasting these patterns with random sequences (AB and ABC, respectively). The baboons were first exposed to sequences of visual shapes instantiating each kind of repetition pattern, and then tested on sequences made of novel shapes either following or violating (random sequences) the repetition pattern.

Methods

Subjects and apparatus

Participants were 17 Guinea baboons (*Papio papio*, 11 females, age range 4-21 years). Six other baboons were tested but not included because they did not complete the experiment in the time allowed to this experiment. The baboons live within a larger group of individuals, within a 700 m² outdoor enclosure and had a permanent access to ten Automated Learning Devices for Monkeys (ALDM, for a detailed description, see (Joël Fagot & Bonté, 2010b; Joël Fagot & Paleressompouille, 2009), each equipped with a 19-inch touch screen and a food dispenser. The main feature of ALDM equipment is that a radio frequency identification reader (RFID) identifies each baboon via a microchip implanted in each arm. The baboons can therefore participate to the research at will, as the test programs recognize them automatically. The experiment was controlled by the EPrime software (Version 2.0, Psychology Software Tools, Pittsburgh).

Prior to the experiment, the baboons have never been tested on sequential pattern learning, but they had previously participated to numerous computerized experiments using the ALDM test systems, including some experiments on statistical learning, involving chunking of triplets of stimuli (Minier, Fagot, & Rey, 2016), and learning of non-adjacent dependencies (Malassis, Rey, & Fagot, *submitted*).

Stimuli

They were 192 different yellow visual shapes (available in Supplementary Information) selected from the Unicode 9.0 repertoire considering their apparent discriminability. Each shape measured 180x180 pixels maximum.

Task

We used an adaptation of the serial response time task (SRT, Nissen & Bullemer, 1987), inspired from Misyak, Christiansen, and Tomblin (2010). This task involves self-paced presentation of sequences of cued stimuli. Each trial (see Figure 1) began by the display of a fixation cross (120x120 pixels) presented at the bottom of the screen. Touching this stimulus triggered the display of two shapes, presented one above the other (upper/lower location balanced) on the left side of the screen. One of the stimulus was surrounded by a square (empty fill, white border; hereafter “cue”) which disappeared after 350 ms. This cue designated the positive stimulus (S+) that the baboon had to touch, while avoiding the other shape (S-). A correct touch of the S+ immediately triggered a second display, which again contained two stimuli presented one above the other (upper/lower location balanced), on the right side of the screen. After 100 ms, the cue appeared around the S+ for 350 ms and the baboon was expected to touch the S+. A correct consecutive selection of the two S+ delivered grains of dry wheat. An incorrect selection of S- on either the first or the second display immediately stopped the trial and triggered a 3-sec timeout without food reward. Trials in which the baboon failed to select a stimulus within five seconds were aborted and represented immediately. The baboons could respond to each display before, during or after the appearance of the cue. Response times collection therefore provides a continuous measure of the participants expectations about the incoming stimulus. The procedure outlined above was followed for sequences of two items. The same procedure was used in non-adjacent repetition pattern trials, with the exception that these trials involved three instead of two displays, which were presented sequentially in left, center and right part of the screen.

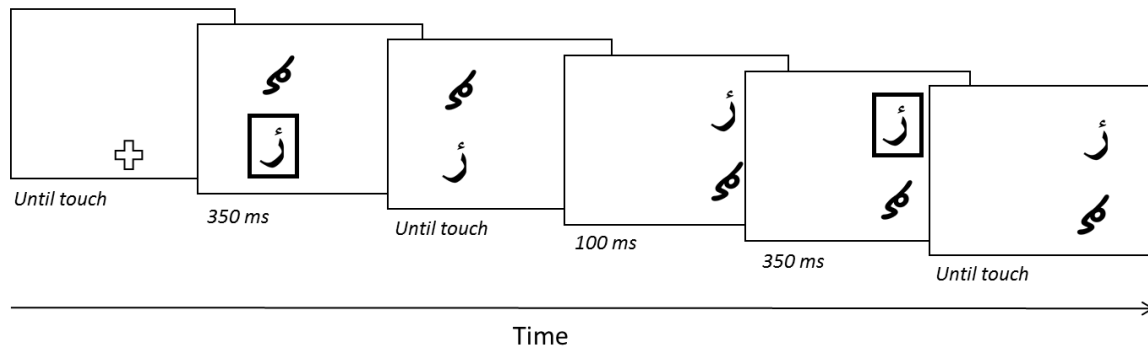


Figure 1 Timeline of an AA trial. The procedure was similar for ABA trials, except that three instead of two displays were presented.

Procedure

Our procedure aimed at comparing baboons' responses to repetition patterns to their responses to random sequences (for a similar approach with humans, see Gervain & Werker, 2013, Experiment 1). Table 1 provides an overview of the experimental phases and conditions. The baboons were tested with the two repetition patterns, one after the other, in a within-subjects design (order counterbalanced across participants). For each pattern, the baboons performed three phases: an exposure phase, a generalization test, and a violation test. These phases were presented sequentially, the order of the two test phases being counterbalanced across baboons. Our collection of 192 shapes was divided in six sets of 32 shapes each (see Supplementary Information). Each set served to generate the sequences for a given phase and pattern (the sets involved in generalization and violation phases being counterbalanced across baboons, see Table 1).

Two types of sequences were presented to the baboons during the exposure phase: baseline (95% of the trials) and violation (5%) sequences. For the adjacent repetition pattern, baseline sequences were of the form AA, and violation sequences were of the form AB, with the A and B stimuli being drawn from the same set of shapes. For the non-adjacent pattern, these two types of sequences were of the form ABA and ABC, respectively, A, B and C stimuli being all drawn from the same set of shapes. In violation trials, the cue appeared around the distractor (i.e., around "B" instead of "A" for the adjacent repetition pattern, and around "C" instead of "A" for the non-adjacent repetition pattern). Baboons were expected to select the cued stimuli. Because 95% of the sequences followed an adjacent or a non-adjacent repetition

pattern, it was possible to predict the stimulus to be selected in the last display. Patterns learning therefore predicts a decrease in response times on the last stimulus of baseline sequences throughout exposure. The introduction of a small amount of violation sequences aimed at disentangling learning of repetition patterns from a mere effect of task practice. We assumed that repetition pattern learning will be indicated by slower response times on violation trials compared to baseline trials.

During the test phases (i.e., the generalization and violation tests), sequences of novel shapes were presented in 20% of the trials. In the generalization test, these sequences followed the repetition pattern. In the violation test, these were random sequences (i.e., of the form AB for adjacent pattern, and ABC for non-adjacent repetition pattern). The remaining 80% of the trials were baseline sequences, randomly intermixed with the novel sequences. Reinforcement schedules were similar to those of the exposure phase: baboons were expected to select the cued stimuli and were reinforced to do so. Note that novel pattern-following and pattern-violating sequences were presented at similar frequencies in generalization and violation tests, respectively. Rote learning of the novel sequences therefore predicts no difference in response times between tests, while generalization of the learned patterns predicts slower response times in the violation compared to the generalization test.

For each pattern, the exposure phase involved 25 sessions of 160 trials (five repetitions of each shape as S+ within a block), and each test phase involved five sessions of 160 trials (four repetitions of each exposure shape and one presentation of each novel shape as S+ within a block). To complete the experiment, the baboons required 18 days on average (range 10-26). Custom-written Python codes (www.python.org) were used to generate the sequences.

Table 1 Overview of all the phases and conditions involved in Experiment 1

Pattern	Phase	Condition	Sequence	Percentage of Trials	Stimuli Set (n°)
Adjacent Repetition	Exposure	Baseline	AA	95%	1
		Violation	AB	5%	
	Generalization Test	Baseline	AA	80%	1
		Generalization	AA	20%	2 or 3
	Violation Test	Baseline	AA	80%	1
		Violation	AB	20%	2 or 3
Non-adjacent Repetition	Exposure	Baseline	ABA	95%	4
		Violation	ABC	5%	
	Generalization Test	Baseline	ABA	80%	4
		Generalization	ABA	20%	5 or 6
	Violation Test	Baseline	ABA	80%	4
		Violation	ABC	20%	5 or 6

Note that the two pattern phases were conducted one after the other (order counterbalanced across participants). The order of test phases and stimuli sets used in those phases were counterbalanced as well.

Data analyses

Sessions were grouped in five (Exposure) and two (Test) blocks of 800 trials each for statistical analyses. All the analyses used as dependent variable the response times (RTs) for touching the last stimulus (i.e., the time elapsed between the apparition of that stimulus and the participant's touch). Incorrect trials were removed from the data set (5.76%), as well as trials with any abnormally long RT (i.e., with a RT > 3 s on either the first, second or third target; 0.22% of the trials). Statistical analyses were done with R (V3.1.1, www.R-project.org).

Learning during the exposure phase was first assessed by comparing baboons' RTs on baseline and violation trials, for each pattern. This was evaluated with a repeated-measures ANOVA involving the Condition (Baseline, Violation), Pattern (Adjacent, Non-adjacent) and Block (1-5) as within-participant factors. Post-hoc comparisons were performed using Tukey's HSD test ($p < .05$).

Two statistical analyses were conducted on the test phases. The first one aimed at verifying that baboons discriminated the novel shapes from the shapes used in baseline sequences, which predicts an increase in RTs. One-sided paired t-tests assessed whether the RTs were greater for the novel compared to baseline sequences. The second analysis assessed whether the increase in RTs (if any) was greater for violation sequences, compared to the generalization sequences. For each participant, mean difference in RTs between the novel and the baseline sequences was computed for each test. These data were then analyzed using one-sided paired t-tests.

Results and Discussion

Exposure phase

Mean RTs are reported in Figure 2 for each pattern, condition and block. The main effect of condition was significant, $F(1, 16) = 15.95$, $p = .001$, partial $\eta_p^2 = .50$, corresponding to overall faster RTs in the baseline (Mean \pm SD = 512 ± 50 ms) than in the violation condition (533 ± 69 ms). Also significant was the main effect of Block, $F(1, 16) = 17$, $p < .001$, partial $\eta_p^2 = .52$, showing that RTs decreased with practice. No main effect of Pattern was found, $F(1, 16) = 0.03$, $p = .87$ (adjacent repetition pattern: 524 ± 58 ms; non-adjacent repetition pattern: 521 ± 64 ms). A significant Condition \times Pattern interaction was however found, $F(1, 16) = 7.31$, $p = .02$, partial $\eta_p^2 = .31$. Post-hoc tests revealed that mean RTs for the different types of sequences differed significantly from each other, in the following order: AA (507 ± 43 ms) < ABA (516 ± 57 ms) < ABC (527 ± 71 ms) < AB (540 ± 67 ms). No other interaction was found (all $ps > .44$). This significant increase in response time observed on the violation sequences for the adjacent and non-adjacent repetition patterns suggests that baboons successfully learned both repetition patterns. However, baboons showed differential responses to the two patterns: effect size of the difference between baseline and violation conditions was moderate to high for the adjacent repetition pattern (Cohen's $d = 0.58$) and small for the non-adjacent repetition

pattern ($d = 0.17$). Further, baboons were faster to respond to baseline sequences of the adjacent repetition pattern than of the non-adjacent pattern, and showed their slowest RTs in reaction to violations of the former.

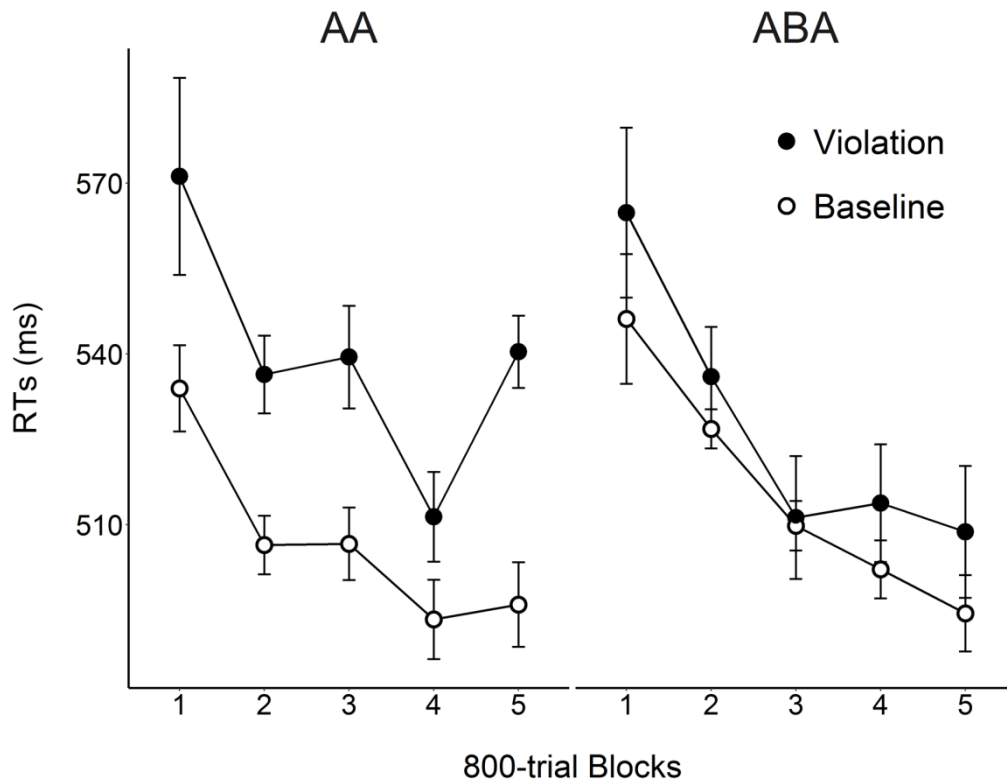


Figure 2 Response times (RTs) obtained in the exposure phase of Experiment 1. Left: adjacent repetition pattern; right: non-adjacent repetition pattern. Bars represent standard errors after subtraction of the individuals' mean RTs.

Test phases

Mean differences in RTs between the novel and baseline sequences are reported in Figure 3 for each pattern and test. A significant positive difference was observed for every pattern and test type: for the adjacent repetition pattern, RTs were greater in generalization trials compared to baseline trials of the generalization test ($M \pm SD = 538 \pm 68$ ms, and 487 ± 33 ms, respectively, $t(16) = 4.01$, $p < .001$, Cohen's $d = 0.96$), as well as in violation trials compared to baseline trials of the violation test (569 ± 61 ms, and 486 ± 37 ms, respectively, $t(16) = 6.31$, $p < .001$, $d = 1.64$). For the non-adjacent repetition pattern, a similar increase in RTs was

observed (generalization vs baseline trials of the generalization test: 535 ± 54 ms, and 500 ± 36 ms, respectively, $t(16) = 4.66$, $p < .001$, $d = 0.76$; violation vs baseline trials of the violation test: 542 ± 51 ms, and 492 ± 33 ms, $t(16) = 6.01$, $p < .001$, $d = 1.18$). These results confirmed that the baboons discriminated the novel shapes used to construct the generalization and violation sequences from the shapes instantiating baseline sequences.

We then assessed whether this difference in RTs was greater for the violation compared to the generalization test. A significant difference was found for the adjacent repetition pattern (Mean difference \pm SD = 83 ± 54 ms, and 52 ± 53 ms, respectively, $t(16) = 2.75$, $p < .01$, $d = 0.60$), as well as for the non-adjacent repetition pattern (50 ± 34 ms, and 35 ± 31 ms, respectively, $t(16) = 2.12$, $p = .03$, $d = 0.48$).

In sum, all the data reported above show generalization of the behavioral response to the two patterns, following the presentation of novel isomorphic sequences. This finding suggests that these two patterns have been successfully learned and generalized by the baboons.

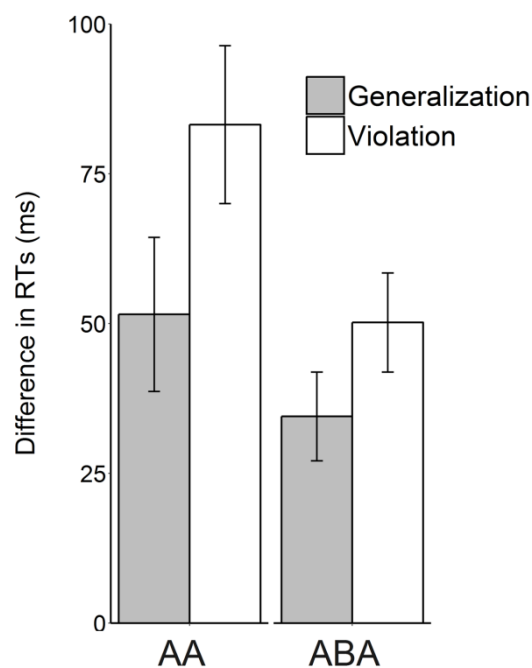


Figure 3 Mean response times (RTs) difference between the novel and baseline conditions of each test phase of Experiment 1. Left: adjacent repetition pattern; right: non-adjacent repetition pattern. Positive values indicate longer RTs in the novel condition compared to the baseline condition. Bars represent standard errors.

Experiment 2

Experiment 1 has suggested that the baboons learned and generalized both adjacent and non-adjacent repetition patterns. However, the possibility remains that the differences in response times observed between the repetition and random patterns derived from the novelty of the last element within these sequences. This issue was addressed in Experiment 2. This second experiment used the same patterns as in Experiment 1 (i.e., adjacent and non-adjacent repetition patterns, in addition with random sequences), but presented these three types of sequences at equal frequency. This procedure allowed us to assess whether any difference in processing speed would be observed between the patterns without learning. To limit potential influences of Experiment 1 on Experiment 2, six months elapsed between the two experiments, during which the baboons were exposed to different and unrelated tasks.

Methods

Subjects, stimuli and task

Data were collected on a total of 13 baboons (six females, age range 4-21 years). Among them, nine had participated in Experiment 1. Eleven other baboons were tested but not included in the analyses because they did not complete the experiment in the allotted time. The apparatus was the same as in the previous experiment.

Experiment 2 used 32 different shapes which have served previously to construct test sequences in Experiment 1 (Set 6). The use of shapes already presented in test sequences was not problematic in itself, because the low frequency of these sequences in Experiment 1 (five repetitions) prevented their learning.

In Experiment 2, the cue was no longer a square surrounding the positive stimulus (S+). It was instead a short disappearance of the stimulus giving the impression that it blinks once. This change was introduced because another experiment (Joël Fagot et al., 2017) suggested that baboons' attention to the task was increased with this procedure. On each display, the two shapes were first presented during 200 ms, after which the S+ disappeared for a short period of 80 ms. The two shapes then remained displayed on the screen until the participant touched one of them. Spatial and temporal parameters of the trials were identical to Experiment 1.

Procedure

Three 3-shapes patterns were presented to the baboons at equal frequencies: random pattern (ABC), pattern involving a final adjacent repetition (ABB) and pattern involving a non-adjacent repetition (ABA). A, B and C stimuli were all drawn from the same set of 32 shapes (Set 6). Ninety-six sequences were created for this design. Thirty-two A-B pairs of shapes were chosen pseudo-randomly, and used as sequences' beginning for the three patterns. They were followed at equal frequencies by an A, B or C shape, rendering the third shape unpredictable. The frequency of each shape within each position was balanced among the three patterns. Experiment 2 involved a total of ten sessions of 96 randomly intermixed trials. Each session comprised 32 sequences for each of the three patterns described above. The 13 baboons required 3.5 days on average to complete Experiment 2 (range 1-6).

Results and Discussion

Sessions were grouped for the analysis in five blocks of 192 trials each. Incorrect trials were removed from the data set (35 %), as well as trials with any abnormally long RT (i.e., with a $RT > 3$ s on either the first, second or third target; 0.5%). Mean RTs (i.e., the time elapsed between the apparition of the last stimulus and the participant's touch) are reported in Figure 4 for each pattern and block. A two-way repeated-measures ANOVA revealed a main effect of the pattern on response times, $F(2, 24) = 11.79, p < .001, \text{partial } \eta_p^2 = .50$. Tukey's HSD test indicated significantly faster RTs for the ABB pattern ($M \pm SD = 660 \pm 66$ ms) compared to the ABA (715 ± 77 ms, Cohen's $d = 0.76$), and ABC patterns (707 ± 62 ms, $d = 0.72$). No significant difference was observed between the ABA and ABC patterns ($p = .41$). The main effect of block was close to significance, $F(1, 12) = 3.90, p = .07, \text{partial } \eta_p^2 = .50$, as was the Pattern x Block interaction, $F(2, 24) = 2.55, p = .10, \text{partial } \eta_p^2 = .50$, suggesting greater acceleration throughout the task for the ABB pattern, compared to the two others.

To analyze the response times obtained at the onset of the experiment, we computed two-tailed paired t-tests comparing the mean RTs obtained in the ten first trials for the different patterns. These analyses indicated a tendency for shorter RTs for the ABB pattern (689 ± 90 ms) than for both the ABA pattern (729 ± 123 ms), $t(12) = 1.86, p = .09, d = 0.37$, and the ABC pattern (723 ± 67 ms), $t(12) = 1.89, p = .08, d = 0.43$. Here again, no difference was found between the ABA and ABC patterns, $t(12) = 0.20, p = .84$.

These results therefore suggest an initial speed advantage for the adjacent repetition pattern (ABB), which increased throughout the experiment. The progressive acceleration observed for the adjacent repetition pattern, but not for the two others, suggests that repeated exposure to adjacent repetitions may prompt their subsequent recognition, even when transitional probabilities are equated (the last element being a repetition of either the first or second shape, or being new, with a probability of .33). Conversely, no difference was found between ABA and ABC patterns, confirming that the differences in RTs observed between the two patterns in Experiment 1 reflect learning of the ABA pattern and its successful generalization.

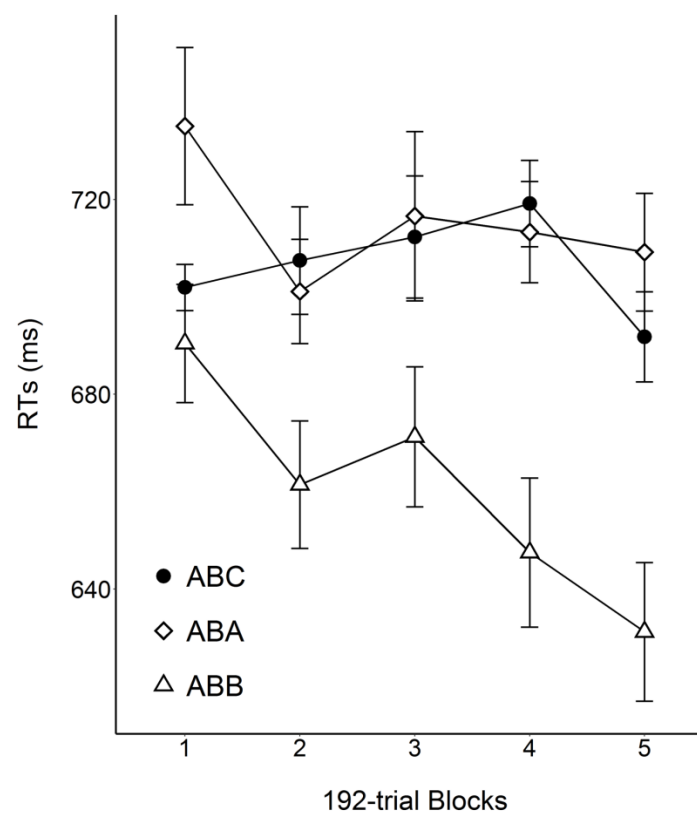


Figure 4 Response times (RTs) for each pattern across blocks in Experiment 2. Bars represent standard errors after subtraction of the individuals' mean RTs.

General Discussion

Our study demonstrates that adjacent repetitions are especially easy to detect by baboons. In Experiment 1, the baboons reacted more strongly to violations of the adjacent

repetition pattern than to violations of the non-adjacent repetition pattern. Moreover, in Experiment 2, baboons showed faster response times in the case of adjacent repetition pattern, while their response times were higher and stable for the non-adjacent repetition and random patterns. These data mirror the results of Gervain et al.'s study (2008, Experiment 1), in which ABB and ABC were presented to human newborns at equal frequencies. Our results therefore support the idea that adjacent repetitions can be detected in absence of training, and that repeated exposures to them enhance their subsequent recognition even when they are interleaved with other types of regularities presented at equal frequencies.

Endress, Nespors, and Mehler (2009) proposed the existence of a set of « perceptive and memory primitives », and that one of these primitives is specialized for the detection of repetitions in a stream of stimuli. These mechanisms would be specialized, automatic, and would require little or no exposure to the stimuli to be recruited. They would furthermore take place in early stages of processing (Endress, Dehaene-Lambertz, & Mehler, 2007; Gervain et al., 2008). Endress, Carden, Versace, and Hauser (2010) proposed that some of these perceptual and memory primitives may be neither language- nor human-specific. Results from our second experiment are consistent with this last hypothesis, as baboons demonstrated a strong sensitivity to adjacent repetitions. However, this effect disappeared for non-adjacent repetitions, probably due to an increased cognitive load. Consequently, future studies assessing the extraction of structural information in non-human animals should guaranty that the subject's strategies do not only rely on the simple detection of repetitions, even more so if their sequences contain adjacent repetitions.

Another important finding of the current study is that baboons' response times were longer for violation sequences of the form ABC, compared to ABA sequences, whether these were made of familiar or novel shapes (Experiment 1). Crucially, no speed difference was observed between these two patterns when their frequencies were equated (Experiment 2). This last result ensures that the findings obtained in Experiment 1 reflect baboons' extraction of the ABA pattern involving non-adjacent repetition. Earlier studies supporting this conclusion compared the ABA to AAB (or ABB) repetition patterns, but this comparison is unable to rule out the hypothesis that the subjects only processed the presence/absence of adjacent repetitions. Comparison of ABA to ABC patterns in our study suggests that baboons can extract more complex structural knowledge from sequences, and that these animals are not exclusively focused on the presence/absence of adjacent repetitions. These results are reminiscent of

Santolin et al. (2016), who found that chicks can discriminate AAB and ABB patterns, and generalize their performance to sequences made of novel stimuli. These authors propose that the chicks considered the presence of adjacent repetitions and their position in the sequence (initial *vs* final position) for responding, but it remains unclear whether chicks relied on the full underlying grammars (e.g., “the first and second elements should be identical while the third differs”). This limitation is also present in our research, as baboons may have learned that the first and last element are identical in the ABA patterns, without processing the fact that second element differs from the first and last ones. Addressing this issue would also require additional experiments, comparing for instance ABA to AAA sequences. Another limitation of the current study is that the non-adjacent repetitions were only separated by one intervening element in the ABA pattern, and it remains unknown if similar findings would persist with larger numbers of intervening elements.

The fact that the baboons processed the non-adjacent repetition pattern in the current study should not hide potential differences with humans. First, humans might differ from other species in the number of exemplars required for exploiting the structural properties of the sequences. A comparative assessment of training set size requirements for successful pattern extraction in human and some other species would inform us on potential differences in constraints. Second, non-human species might further differ from humans regarding the type of elements over which structural regularities can be extracted. In the current study, patterns were based on identity relationships at the token level (i.e., “A” being an arbitrary element which is repeated). In human language, however, structural regularities can also be found over *functional* categories (ten Cate & Okanoya, 2012; Corballis, 2009) and it remains unknown if nonhuman species have the capacity of extracting such language-like patterns. Finally, results from a recent neuroimaging study suggest that the brain areas involved in sequential patterns encoding partly differ between humans and monkeys (L. Wang et al., 2015).

The nature of the mechanisms underlying the processing of the structural properties of sequences is a matter of a longstanding, and still ongoing debate. Marcus et al. (1999; see also Marcus, 2001) argued that generalization to novel isomorphic sequences involves rule-governed mechanisms and symbolic representations of the patterns. However, numerous alternative accounts were proposed for this behavior (e.g., Altmann, 2002; Altmann & Dienes, 1999; Brooks & Vokey, 1991; Christiansen & Curtin, 1999; McClelland & Plaut, 1999; Negishi, 1999; Pothos, 2005; Seidenberg & Elman, 1999). Most of these accounts argue that statistical learning mechanisms can explain generalization in these tasks, and do not postulate

the existence of rule-based mechanisms. In a complementary perspective, Brooks and Vokey (1991) proposed that the subjects may compare the novel exemplars to a set of memorized exemplars, instead of comparing them to an abstract, symbolic, representation of these sequences' structure. Although such alternative accounts are still debated (for a discussion, see e.g., Endress et al., 2007), they may ultimately offer a more parsimonious explanation of generalization in pattern learning studies, than the hypothesis of rule learning.

Acknowledgments

This work was carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002). It has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX). This research was also supported by the Premilang2 ANR-project (#ANR-13-BSH2-0002) and the Chunked ANR-project (#ANR-17-CE28-0013-02).

References

- Altmann, G. T. M. (2002). Learning and development in neural networks-the importance of prior experience. *Cognition*, *85*(2), B43-50.
- Altmann, G. T. M., & Dienes, Z. (1999). Rule Learning by Seven-Month-Old Infants and Neural Networks. *Science*, *284*(5416), 875–875.
<https://doi.org/10.1126/science.284.5416.875a>
- Beckers, G. J. L., Berwick, R. C., Okanoya, K., & Bolhuis, J. J. (2016). What do animals learn in artificial grammar studies? *Neuroscience & Biobehavioral Reviews*.
<https://doi.org/10.1016/j.neubiorev.2016.12.021>
- Bouchon, C., Nazzi, T., & Gervain, J. (2015). Hemispheric Asymmetries in Repetition Enhancement and Suppression Effects in the Newborn Brain. *PLoS ONE*, *10*(10).
<https://doi.org/10.1371/journal.pone.0140160>
- Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology. General*, *120*(3), 316–323. <http://dx.doi.org/10.1037/0096-3445.120.3.316>
- Chen, J., van Rossum, D., & ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition*, *18*(1), 151–164.
<https://doi.org/10.1007/s10071-014-0786-4>
- Christiansen, M. H., & Curtin, S. (1999). Transfer of learning: rule acquisition or statistical learning? *Trends in Cognitive Sciences*, *3*(8), 289–290.
- Corballis, M. C. (2009). Do rats learn rules? *Animal Behaviour*, *78*(4), e1–e2.
<https://doi.org/10.1016/j.anbehav.2009.05.001>
- de la Mora, D. M., & Toro, J. M. (2013). Rule learning over consonants and vowels in a non-human animal. *Cognition*, *126*(2), 307–312.
<https://doi.org/10.1016/j.cognition.2012.09.015>
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal Cognition*, *13*(3), 483–495.
<https://doi.org/10.1007/s10071-009-0299-8>

- Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, *105*(3), 577–614.
<https://doi.org/10.1016/j.cognition.2006.12.014>
- Endress, A. D., Nespors, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, *13*(8), 348–353.
<https://doi.org/10.1016/j.tics.2009.05.005>
- Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: use of a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio papio*). *Behavior Research Methods*, *42*(2), 507–516.
<https://doi.org/10.3758/BRM.42.2.507>
- Fagot, J., Malassis, R., & Medam, T. (2017). The processing of positional information in a two-item sequence limits the emergence of symmetry in baboons (*Papio papio*), but not in humans (*Homo sapiens*). *Learning & Behavior*. <https://doi.org/10.3758/s13420-017-0290-1>
- Fagot, J., & Paleressompouille, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, *41*(2), 396–404.
<https://doi.org/10.3758/BRM.41.2.396>
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences*, *105*(37), 14222–14227. <https://doi.org/10.1073/pnas.0806530105>
- Gervain, Judit, & Werker, J. F. (2013). Learning non-adjacent regularities at age 0 ; 7. *Journal of Child Language*, *40*(4), 860–872. <https://doi.org/10.1017/S0305000912000256>
- Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of “sameness” and “difference” in an insect. *Nature*, *410*(6831), 930–933.
<https://doi.org/10.1038/35073582>
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109–135.
- Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., & Slemmer, J. A. (2009). Abstract Rule Learning for Visual Sequences in 8- and 11-Month-Olds. *Infancy: The Official Journal of the International Society on Infant Studies*, *14*(1), 2–18. <https://doi.org/10.1080/15250000802569611>

- Kovács, Á. M., & Mehler, J. (2009). Flexible Learning of Multiple Speech Structures in Bilingual Infants. *Science*, 325(5940), 611–612.
<https://doi.org/10.1126/science.1173947>
- Marcus, G. F. (2001). *The Algebraic Mind*. MIT Press.
- Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, 283(5398), 77–80.
<https://doi.org/10.1126/science.283.5398.77>
- McClelland, J. L., & Plaut, D. C. (1999). Does generalization in infant learning implicate abstract algebra-like rules? *Trends in Cognitive Sciences*, 3(5), 166–168.
[https://doi.org/10.1016/S1364-6613\(99\)01320-0](https://doi.org/10.1016/S1364-6613(99)01320-0)
- Minier, L., Fagot, J., & Rey, A. (2016). The Temporal Dynamics of Regularity Extraction in Non-Human Primates. *Cognitive Science*, 40(4), 1019–1030.
<https://doi.org/10.1111/cogs.12279>
- Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J. (2010). Sequential Expectations: The Role of Prediction-Based Learning in Language. *Topics in Cognitive Science*, 2(1), 138–153. <https://doi.org/10.1111/j.1756-8765.2009.01072.x>
- Murphy, R. A., Mondragón, E., & Murphy, V. A. (2008). Rule Learning by Rats. *Science*, 319(5871), 1849–1851. <https://doi.org/10.1126/science.1151564>
- Negishi, M. (1999). Do infants learn grammar with algebra or statistics? *Science (New York, N.Y.)*, 284(5413), 435; author reply 436-437.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32.
[https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)
- Pothos, E. M. (2005). The rules versus similarity distinction. *The Behavioral and Brain Sciences*, 28(1), 1–14; discussion 14-49.
- Ravignani, A., & Sonnweber, R. (2017). Chimpanzees process structural isomorphisms across sensory modalities. *Cognition*, 161, 74–79.
<https://doi.org/10.1016/j.cognition.2017.01.005>
- Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, 105(3), 669–680.
<https://doi.org/10.1016/j.cognition.2006.11.004>

- Santolin, C., Rosa-Salva, O., Regolin, L., & Vallortigara, G. (2016). Generalization of visual regularities in newly hatched chicks (*Gallus gallus*). *Animal Cognition*, *19*(5), 1007–1017. <https://doi.org/10.1007/s10071-016-1005-2>
- Seidenberg, M. S., & Elman, J. L. (1999). Do Infants Learn Grammar with Algebra or Statistics? *Science*, *284*(5413), 433–433. <https://doi.org/10.1126/science.284.5413.433f>
- Sonnweber, R., Ravignani, A., & Fitch, W. T. (2015). Non-adjacent visual dependency learning in chimpanzees. *Animal Cognition*, *18*(3), 733–745. <https://doi.org/10.1007/s10071-015-0840-x>
- Spierings, M. J., & ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(27), E3977–3984. <https://doi.org/10.1073/pnas.1600483113>
- ten Cate, C. (2014). On the phonetic and syntactic processing abilities of birds: from songs to speech and artificial grammars. *Current Opinion in Neurobiology*, *28*, 157–164. <https://doi.org/10.1016/j.conb.2014.07.019>
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *367*(1598), 1984–1994. <https://doi.org/10.1098/rstb.2012.0055>
- Thiessen, E. D. (2012). Effects of Inter- and Intra-modal Redundancy on Infants' Rule Learning. *Language Learning and Development*, *8*(3), 197–214. <https://doi.org/10.1080/15475441.2011.583610>
- van Heijningen, C. A. A. van, Visser, J. de, Zuidema, W., & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, pnas.0908113106. <https://doi.org/10.1073/pnas.0908113106>
- van Heijningen, C. A. A., Chen, J., van Laatum, I., van der Hulst, B., & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? *Animal Cognition*, *16*(2), 165–175. <https://doi.org/10.1007/s10071-012-0559-x>

- Versace, E., Spierings, M. J., Caffini, M., ten Cate, C., & Vallortigara, G. (2017). Spontaneous generalization of abstract multimodal patterns in young domestic chicks. *Animal Cognition*, 20(3), 521–529. <https://doi.org/10.1007/s10071-017-1079-5>
- Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of Numerical and Sequential Patterns in Macaque and Human Brains. *Current Biology*, 25(15), 1966–1974. <https://doi.org/10.1016/j.cub.2015.06.035>

IV. ETUDE III – Etude Corrélacionnelle : Performances Individuelles des Babouins pour l’Extraction de Régularités Adjacentes et Non-adjacentes

Malassis, R., Bogaerts, L., Minier-Munding, L., Fagot, J., & Rey, A. (*en préparation*).

Cette étude a bénéficié d’une collaboration avec Louisa Bogaerts, qui a été post-doctorante au sein du laboratoire de psychologie cognitive de l’université d’Aix-Marseille (à présent post doctorante à l’université hébraïque de Jérusalem), et travaille notamment sur les différences de performances individuelles dans des tâches d’apprentissage statistique chez l’humain (Bogaerts, Siegelman, Ben-Porat, & Frost, 2018; Siegelman, Bogaerts, Christiansen, et al., 2017; Siegelman, Bogaerts, & Frost, 2017). Il s’agit d’une analyse corrélacionnelle des performances individuelles des babouins dans différentes tâches d’apprentissage de régularités séquentielles. Pour ces analyses, nous avons utilisé les données issues des Etudes I et II de ma thèse, ainsi que les données issues du travail de Minier, Fagot, et Rey (2016). Nous y avons également inclus des données issues d’une tâche de mémoire de travail, qui fait partie d’un ensemble de tâches de fond proposées aux babouins lorsqu’aucune expérience particulière n’est en cours.

Introduction

Classiquement, l'extraction de régularités statistiques dans l'environnement a été envisagée comme reposant sur des mécanismes d'apprentissage *généraux*, par opposition à des mécanismes spécifiques au langage (e.g. Bates & Elman, 1996; Elman, 1998). Un certain nombre de données vont en ce sens. Ainsi, l'extraction de régularités distributionnelles est observée avec des stimuli non linguistiques (e.g. Fiser & Aslin, 2002; Saffran et al., 1999), et des liens sont observés entre les performances des participants dans ces tâches et le traitement de langues naturelles : les individus avec troubles du langage (*Specific Language Impairment*) obtiennent des performances plus faibles dans ces tâches que des individus au développement typique (chez individus avec troubles spécifiques du langage : (Evans et al., 2009; Hsu, Tomblin, & Christiansen, 2014b; Lum et al., 2014)), et des corrélations entre les performances d'apprentissage statistique et de traitement des langues naturelles sont observées au sein de cette dernière population (Misyak & Christiansen, 2012; Misyak, Christiansen, & Tomblin, 2010). De plus, certaines contraintes opèrent de manière similaire sur différents types de stimuli linguistiques et non linguistiques (e.g. similarité perceptive facilitant l'extraction de dépendances non-adjacentes, observée avec des classes de phonèmes, Newport & Aslin, 2004; Onnis et al., 2005 ; des sons musicaux, Creel et al., 2004; Endress, 2010 ; et non musicaux Gebhart et al., 2009).

Cependant, des théories récentes sur l'apprentissage statistique proposent que cet apprentissage ne soit pas sous-tendu par un mécanisme unique, mais plutôt par des principes computationnels communs, soumis à des contraintes différentes selon les régularités traitées (Frost, Armstrong, Siegelman, & Christiansen, 2015; Saffran & Thiessen, 2007; Siegelman, Bogaerts, Christiansen, & Frost, 2017). Ces contraintes seraient définies par un ensemble de facteurs perceptifs (non linguistiques), tels que la modalité de présentation (auditive, visuelle, etc), le type de stimuli, et le type de dépendance (adjacentes ou non-adjacentes). Concernant la modalité, un ensemble de données chez l'humain montre : (1) des différences quantitatives et qualitatives entre des tâches d'apprentissage statistique effectuées dans différentes modalités (e.g. auditive, visuelle, et tactile, Conway & Christiansen, 2005) ; (2) l'absence de transfert des régularités apprises d'une modalité à l'autre (e.g. Conway & Christiansen, 2006) ; et (3) l'absence de corrélation entre les performances individuelles obtenues dans différentes tâches d'apprentissage statistique (Frost et al., 2015). Des spécificités sont également retrouvées au niveau de caractéristiques plus fines des stimuli. Par exemple, dans une tâche de recherche

visuelle, les performances des participants dépendent de la localisation précise du stimulus dans leur champ visuel. Ainsi, lorsque l'élément cible est présenté à une nouvelle localisation, les performances des participants s'en trouvaient fortement dégradées (Sigman & Gilbert, 2000). Enfin, des études corrélationnelles ont montré que les performances de participants humains dans des tâches d'apprentissage de dépendances adjacentes et non-adjacentes ne sont pas corrélées (Misyak & Christiansen, 2012; Siegelman & Frost, 2015).

Cet ensemble de données suggère que l'apprentissage statistique est fortement contraint par les différentes caractéristiques des stimuli. Toutefois, le nombre d'études ayant examiné cette question reste très faible chez l'humain, et à notre connaissance, aucune étude n'a été menée chez une autre espèce. Nous avons donc effectué des analyses corrélationnelles sur les performances individuelles obtenues par des babouins dans différentes expériences d'apprentissage de régularités séquentielles, ainsi que dans une tâche de mémoire de travail, en vue de répondre aux trois questions suivantes :

- (1) Les performances individuelles des babouins pour un même test sont-elles stables ?
- (2) Ces performances corréleraient-elles entre les différentes tâches et les différents types de dépendances (adjacentes et non-adjacentes) ?
- (3) Ces performances corréleraient-elles avec des différences inter-individuelles de mémoire de travail ?

Méthodes

Procédure Générale

Nous avons examiné les performances individuelles des babouins dans trois études menées au sein de notre laboratoire, dont deux réalisées dans le cadre de ma thèse. Deux de ces études ont été menées avec une tâche visuo-spatiale de pointage sériel, et examinaient l'apprentissage de dépendances adjacentes (Minier, Fagot, & Rey, 2016), et non-adjacentes (Etude I, Expériences 1 et 2). La troisième étude testait la détection de répétitions adjacentes et non-adjacentes dans des séquences de stimuli visuels (Etude II, Expérience 1).

Pour répondre à notre première question, il est nécessaire d'examiner les performances intra-test des participants, i.e. dans deux expériences réalisées avec la même tâche et le même type de dépendance. L'Etude I, qui portait sur l'apprentissage de dépendances non-adjacentes

dans une tâche visuo-spatiale de pointage sériel, comporte deux expériences, menées à plusieurs mois d'intervalle. Les phases d'exposition étaient similaires dans leur principe général, mais les langages employés différaient d'une expérience à l'autre (i.e. les babouins devaient apprendre de nouvelles dépendances non-adjacentes dans la deuxième expérience). Nous avons donc sélectionné cette étude pour examiner la corrélation intra-tâche des performances.

Pour répondre à notre deuxième question, nous avons examiné les corrélations entre les performances individuelles des babouins dans les deux tâches (visuo-spatiale et visuelle) et pour les deux types de dépendances (adjacentes et non-adjacentes).

Enfin, nous avons examiné si nous observions des corrélations entre les performances des babouins dans ces tâches et leur mémoire de travail (Question 4). Comme mesure de cette dernière, nous avons choisi d'examiner leurs résultats dans une tâche de pointage auto-ordonnée (en anglais, *self-ordered pointing task*, ci-après SOPT).

Participants

La Table 1 présente le nombre de babouins ayant complété chacune des expériences. Nous avons sélectionné pour nos analyses corrélationnelles les expériences pour lesquelles nous avons un nombre maximal de participants en commun. L'examen de la corrélation intra-test a donc été effectué sur les données issues de notre Etude I, entre l'Expérience 1 et l'Expérience 2 (Langage 2). L'examen des corrélations inter-test a été réalisé en comparant l'Expérience 2 de l'Etude I (Langage 2) avec l'Etude II et l'étude de Minier et al. (2016).

Table 1**Nombre de babouins ayant complété chacune des expériences.**

		Nadj				Adj	
		VS			V	VS	V
		Etude I Exp. 1	Etude I Exp. 2 (L1)	Etude I Exp. 2 (L2)	Etude II	Minier et al. (2016)	Etude II
Nadj	VS	Etude I - Exp. 1					
		14					
		16	14				
V		14	12	17			
	VS	8	8	10	9		
Adj	V	14	12	17	17	9	
	SOPT	15	13	18	17	9	17

Types de dépendances : Nadj = non-adjacentes, Adj = adjacentes ; Tâches : VS = visuo-spatiale, V = visuelle. L1 = langage 1, L2 = langage 2. La valeur en bleu indique la comparaison sélectionnée pour l'analyse corrélationnelle intra-test, et les valeurs en vert les comparaisons sélectionnées pour les analyses corrélationnelles inter-tests.

Tâches

Tâches d'apprentissage de régularités séquentielles

La Figure 1 présente le principe général des deux tâches d'apprentissage statistique utilisées. Dans la tâche visuo-spatiale, le participant doit suivre une cible se déplaçant dans une matrice sur un écran tactile. En présentant des séquences non-aléatoires, on peut tester avec cette tâche la capacité des participants à prédire la *localisation* des cibles à venir.

Dans la tâche visuelle, le participant doit sélectionner à chaque étape de l'essai un stimulus signalé par un indice (encadré brièvement), entre deux alternatives. Avec cette tâche, on teste la capacité des participants à former des prédictions sur la *forme* du prochain stimulus à sélectionner, la localisation de ce stimulus (présenté soit dans la partie supérieure de l'écran, soit dans sa partie basse, cf. Figure 1) étant aléatoire. Pour plus de détails sur ces tâches, voir les sections *Méthodes* des Etudes I & II.

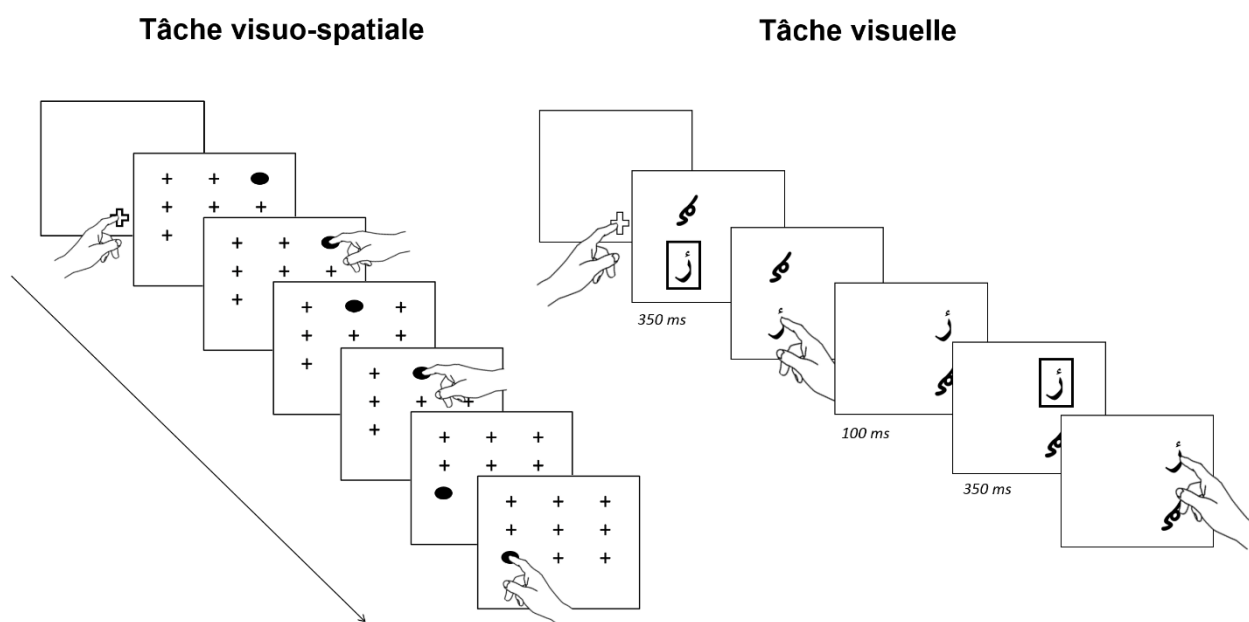


Figure 1 Illustration du déroulement d'un essai dans les tâches d'apprentissage statistique visuo-spatiale et visuelle.

Tâche SOPT

Dans la tâche SOPT, un certain nombre de stimuli visuels sont présentés simultanément sur un écran tactile. Les babouins doivent sélectionner chacun de ces stimuli, dans n'importe quel ordre mais sans sélectionner deux fois le même. Après chaque sélection, l'emplacement des stimuli est modifié. Effectuer cette tâche requière de maintenir en mémoire les stimuli déjà touchés.

Séquences

Dans l'étude de Minier et al. (2016), les babouins devaient suivre une séquence de neuf cibles à chaque essai. Après une phase contrôle dans laquelle ces cibles étaient présentées dans un ordre aléatoire, les babouins étaient exposés de manière répétée à trois triplets de la forme ABC, présentés dans un ordre pseudo-aléatoire à chaque essai.

Dans l'Etude I, des séquences de trois cibles étaient présentées aux participants. Elles étaient de la forme AXB, avec deux paires A-B, et cinq (Expérience 1) ou six (Expérience 2) éléments intermédiaires X. Ainsi, la localisation de la première cible permettait de prédire celle

de la dernière, tandis que la deuxième cible apparaissait à une localisation variable et non informative. Dans une condition contrôle, des séquences de la forme XXB étaient présentées.

Enfin dans l'Etude II, les babouins étaient exposés à des séquences comportant une répétition adjacente, de la forme AA, et à des séquences comportant une répétition non-adjacente, de la forme ABA (l'ordre de passation de ces deux expériences étant contrebalancé à travers les participants). Une petite proportion d'essais contrôle était également présentée, qui impliquaient des séquences non-congruentes de la forme AB et ABC, respectivement. Ces expériences impliquaient également une phase de test dans laquelle de nouvelles séquences, isomorphes ou non, étaient présentées aux participants. Ces données n'ont pas été incluses dans nos analyses, car elles testaient une autre question, la capacité des babouins à généraliser les régularités structurelles auxquelles ils avaient été exposés à de nouveaux éléments. Seules les données des phases d'exposition aux régularités ont donc été utilisées pour nos analyses corrélationnelles.

Analyse des Données

Tâches d'apprentissage de régularités séquentielles

Dans les tâches que nous avons utilisées, ce sont les temps de réponse qui constituent la mesure de l'apprentissage effectué par les participants. Toutefois, les valeurs absolues de ces temps de réponse sont susceptibles de varier fortement d'un individu à l'autre, indépendamment de l'apprentissage (e.g. différences d'âge, de taille, etc). Nous avons donc calculé pour chacune des expériences et chacun des participants la différence (en ms) entre les temps de réponse moyens dans la condition contrôle, et dans la condition expérimentale.

Pour les données de Minier et al. (2016), nous avons calculé les temps de réponse moyens obtenus au cours de phase d'exposition aux triplets, que nous avons soustraits aux temps de réponse moyens de la phase aléatoire. Nous avons fait le choix d'utiliser pour ce calcul seulement les temps de réponse sur le deuxième élément des triplets.

Pour les Etudes I et II, nous avons calculé les différences entre les temps de réponse dans la condition contrôle et dans la condition expérimentale obtenus au cours de la phase d'exposition, pour le dernier élément de la séquence (i.e. l'élément qui est prédictible dans la condition expérimentale).

Pour chaque analyse de corrélation, nous avons éliminé les participants pour lesquels cette différence était inférieure ou supérieure à 2 écarts-types de la moyenne calculée sur le sous-groupe de sujets en commun aux deux expériences considérées.

SOPT

Dans la tâche SOPT, ce sont les scores qui constituent la variable dépendante, ainsi que le niveau de difficulté atteint. En effet, le nombre d'items à sélectionner à chaque essai augmente de manière incrémentale en fonction des performances du participant. Ainsi, si le participant atteint le critère ($> 80\%$ d'essais corrects sur un bloc) à l'étape où il doit sélectionner trois items, dans le bloc suivant il aura à sélectionner quatre items ; et ainsi de suite. Nous avons donc, pour chaque participant, additionné les scores obtenus à chacune de ces étapes (cf. Annexe 1). Cette somme constitue notre indicateur de performance pour la tâche SOPT.

Résultats

Données Descriptives

La Table 2 présente les données descriptives pour chacune des expériences. Une forte variabilité inter-individuelle et inter-expériences est observée. Pour chacune des expériences, des individus ne présentent pas de différence de temps de réponse entre la condition contrôle et la condition expérimentale, tandis que d'autres présentent des différences de plusieurs dizaines de millisecondes. De plus, les différences moyennes sur le groupe sont plus importantes pour les expériences portant sur l'apprentissage de régularités adjacentes, que pour celles portant sur l'apprentissage de régularités non-adjacentes.

Table 2 **Données descriptives**

Sujet	Nadj				Adj		SOPT
	VS			V	VS	V	
	Etude I Exp. 1	Etude I Exp. 2 (L1)	Etude I Exp. 2 (L2)	Etude II Exp. 1	Minier et al. (2016)	Etude II Exp. 1	
ANG	5	-5	48				1.64
ARI	22		3	-1		29	2.27
ATM	-12	0	10		7		
BOB			2	18		55	2.26
BRI	1	8	9	8		-5	1.92
DAN	-1		0	21	63	24	2.54
DOR	4	3	4	5	86	20	3.34
DRE			13	16	75	29	3.13
EWI	8	4	-2	33	112	54	2.94
FAN	5	5	5	61	65	126	3.95
FEL	4	-2	7	30	55	-9	2.44
FEY	-2	12	7	40	96	13	3.00
FIL	6	14	7	10	91	34	2.86
FLU			-2	-17	139	67	2.43
HAR	-3	9	-4	-22		17	2.09
HER	9	12	9	20		36	2.37
KAL	9	15	28	-4		60	1.42
VIO	7	10	2	-21		19	2.13
VIV	0	0	-3	2		-11	1.62

Types de dépendances : Nadj = non-adjacentes, Adj = adjacentes ; Tâches : VS = visuo-spatiale, V = visuelle. L1 = langage 1, L2 = langage 2. Moy = moyenne, SD = écart-type, Min = valeur minimale, Max = valeur maximale. Pour les tâches d'apprentissage statistique, les valeurs correspondent à la différence (en ms) de temps de réponse entre la condition contrôle et la condition expérimentale. Les valeurs positives indiquent des temps de réponse inférieurs dans la condition expérimentale, où les régularités présentes dans les séquences permettent de prédire la localisation (tâche VS) ou la forme (tâche V) du dernier élément de la séquence, par rapport à la condition contrôle où cette prédiction n'est pas possible. Pour SOPT, les valeurs réfèrent à la somme des scores moyens obtenus pour les différents niveaux de difficulté de la tâche atteints par le participant (voir texte principal).

Stabilité des Performances Individuelles

La Figure 2 représente les performances des babouins dans l'Expérience 1 de l'Etude I en fonction de leurs performances dans l'Expérience 2 (Langage 2) de cette même étude. Une corrélation positive est observée, mais elle reste non-significative ($r = 0.49$, $n = 13$, $p = 0.09$, 95% CI = [-0.09,0.82]). Les babouins qui ont appris les dépendances non-adjacentes dans l'Expérience 1 tendent donc à être également ceux parvenant à apprendre les nouvelles dépendances non-adjacentes présentées dans l'Expérience 2. Ce résultat suggère que pour une

même tâche et un même type de dépendances, les performances des babouins seraient relativement stables d'un test à l'autre.

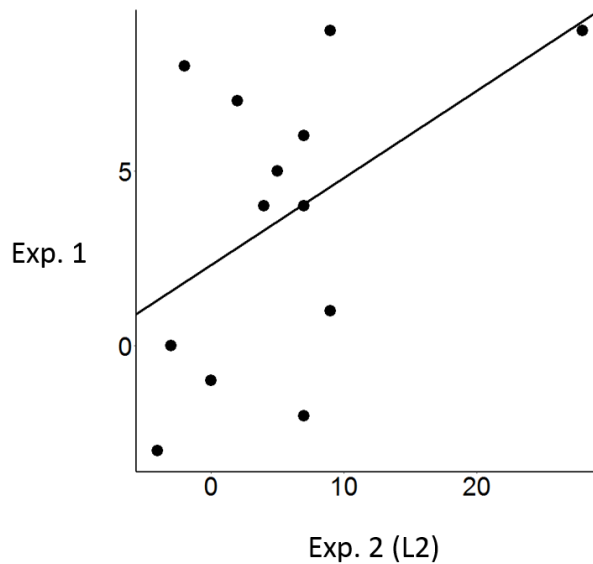


Figure 2 Performances des participants dans l'Expérience 1 de l'Etude I en fonction de leurs performances dans l'Expérience 2 (Langage 2) de cette même étude.

Effets de la Tâche et du Type de Dépendance

La Figure 3 représente les performances des babouins dans les différents tests. Les valeurs de ces corrélations sont rapportées dans la Table 3. Une corrélation positive restant non significative est observée entre les deux tâches dans lesquelles les babouins devaient apprendre des dépendances non-adjacentes ($r = 0.34, p = 0.11$). Une corrélation positive significative ($r = 0.85, p < 0.01$) est observée entre les deux tâches d'apprentissage de dépendances adjacentes. Les performances des babouins ne sont en revanche pas corrélées entre les expériences impliquant des types de dépendance différents. Des corrélations négatives tendent même à être observées entre les performances des babouins dans la tâche visuo-spatiale d'apprentissage de dépendances adjacentes, et celles obtenues dans les deux tâches d'apprentissage de dépendances non-adjacentes.

Les performances des babouins ne sont donc pas corrélées au sein d'une même tâche (i.e. visuo-spatiale ou visuelle) lorsque le type de dépendance diffère. Ainsi, les babouins qui

obtiennent les meilleures performances dans l'apprentissage de dépendances adjacentes ne sont pas ceux qui parviennent à apprendre les dépendances non-adjacentes.

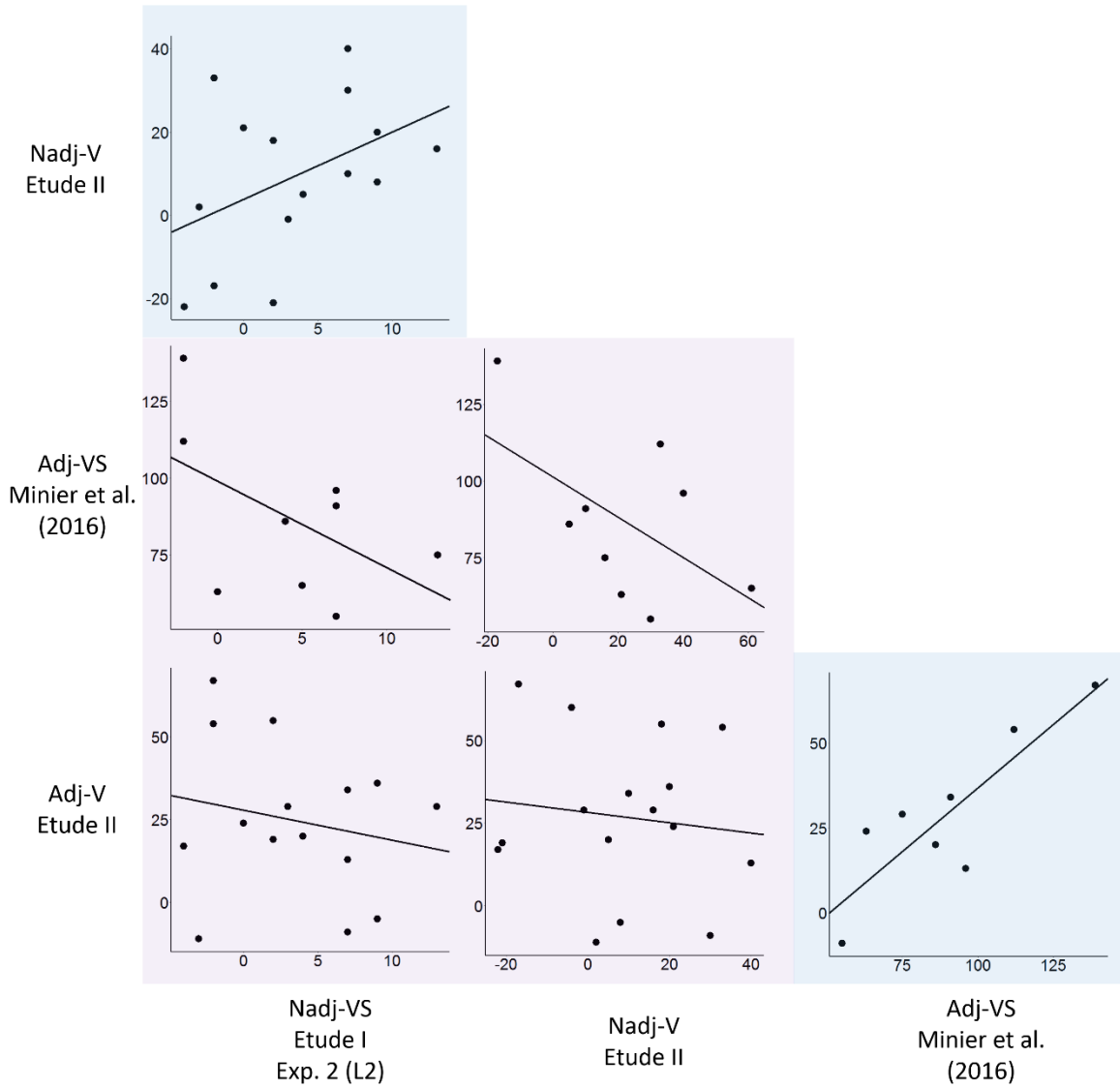


Figure 3 Performances individuelles en fonction des différentes expériences, pour un même type de dépendance (en bleu), et pour des dépendances différentes (en violet). Types de dépendances : Nadj = non-adjacentes, Adj = adjacentes ; Tâches : VS = visuo-spatiale, V = visuelle. L2 = langage 2.

Table 3 **Corrélations entre les différentes tâches d'apprentissage statistique**

			Nadj		Adj
			VS	V	VS
			Etude I - Exp. 2 (L2)	Etude II	Minier et al. (2016)
Nadj	V	Etude II	r = 0.34, n = 15, p = 0.11 [-0.10, 0.77]		
	VS	Minier et al. (2016)	r = -0.52, n = 9, p = 0.15 [-0.88, 0.21]	r = -0.55, n = 9, p = 0.12 [-0.89, 0.17]	
Adj	V	Etude II	r = -0.20, n = 15, p = 0.47 [-0.65, 0.35]	r = -0.12, n = 16, p = 0.65 [-0.58, 0.40]	r = .85, n = 8, p < 0.01 [0.38, 0.97]

Note : Les valeurs entre crochets indiquent les intervalles de confiance à 95%.

Apprentissage Statistique et Mémoire de Travail

L'apprentissage de dépendances non-adjacentes est susceptible d'impliquer davantage la mémoire de travail que l'apprentissage de dépendances adjacentes. En effet, cet apprentissage pourrait nécessiter le maintien temporaire en mémoire de travail du premier élément de la séquence, l'élément prédictif, pour détecter son association avec l'élément prédictible. Nous avons donc examiné si les performances d'apprentissage statistique des babouins, pour l'un et l'autre de ces types de dépendance, corrôlaient avec une mesure de leurs capacités de mémoire de travail. La Figure 4 et la Table 4 présentent les corrôlations entre les performances des babouins dans la tâche de SOPT et les différentes expériences d'apprentissage statistique. Une corrôlation positive proche du seuil de significativité ($r = 0.47$, $p = 0.06$) est observée entre la tâche SOPT et l'expérience dans laquelle les babouins devaient apprendre des dépendances non-adjacentes dans la tâche visuelle (Nadj-V, Etude II). Aucune autre corrôlation significative n'est observée.

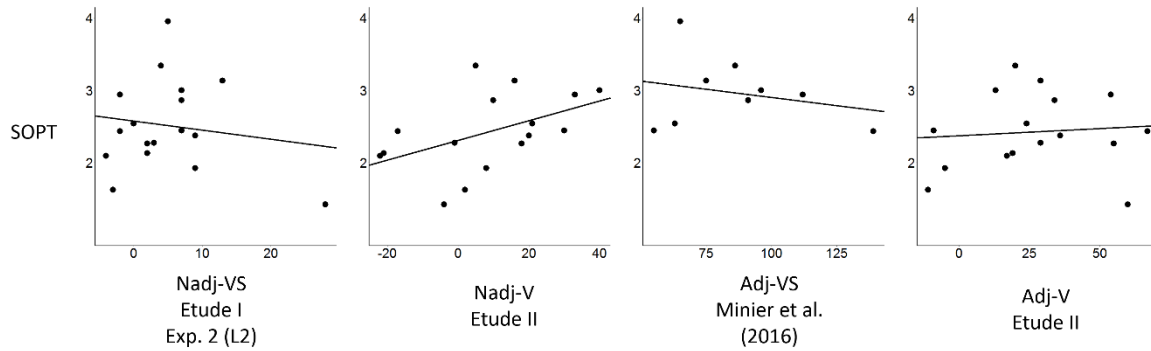


Figure 4 Performances des participants dans la tâche de SOPT, en fonction de leurs performances dans les différentes expériences d'apprentissage statistique.

Table 4 Corrélations entre les différentes tâches d'apprentissage statistique

	Nadj		Adj	
	VS	V	VS	V
	Etude I - Exp. 2 (L2)	Etude II	Minier et al. (2016)	Etude II
SOPT	$r = -0.15, n = 17, p = .56$ [-0.59, 0.36]	$r = 0.47, n = 16, p = .06$ [-0.03, 0.79]	$r = -0.25, n = 9, p = .52$ [-0.78, 0.50]	$r = 0.09, n = 16, p = .74$ [-0.43, 0.56]

Les valeurs entre crochets indiquent les intervalles de confiance à 95%.

Ces résultats confirment donc partiellement notre hypothèse : notre mesure de la mémoire de travail des babouins tend à être corrélée avec leurs performances dans l'une de nos expériences testant leur capacité à détecter des régularités non-adjacentes, tandis qu'elles ne sont pas significativement corrélées à leurs performances dans deux tâches d'apprentissage de dépendances adjacentes.

Discussion

Dans cette étude, nous avons réalisé des analyses corrélationnelles sur un ensemble d'expériences réalisées sur un même groupe de babouin. Afin de discuter les résultats obtenus au regard de la littérature, nous avons résumé dans la Table 5 les principaux résultats des deux seules études dont nous ayons connaissance ayant effectuées des analyses similaires chez l'humain. Misyak et Christiansen (2012) ont soumis un groupe de participants humains à deux tests successifs, l'un visant à évaluer leurs capacités d'apprentissage de dépendances

adjacentes, et l'autre de dépendances non-adjacentes. Ils ont dans les deux cas utilisé une tâche présentée dans la modalité auditive, impliquant la présentation de stimuli linguistiques (pseudo-mots). Siegelman et Frost (2015) ont employé une plus grande diversité de tâches d'apprentissage statistique. Dans ces deux études, les participants étaient également soumis à une batterie de tests évaluant d'autres fonctions cognitives (mémoire de travail, QI, traitement de langue naturelle, etc). Ci-dessous nous discutons nos résultats au regard des données de la littérature, pour chacune des questions examinées.

Table 5 Résultats de deux études examinant les corrélations entre les performances individuelles d'apprentissage statistique en fonction du type de dépendance et de la tâche chez l'humain.

		Misyak & Christiansen (2012)		Siegelman & Frost (2015)		
		Adj		Adj		Nadj
		Aud-V		Vis-NV	Aud-NV	Aud-V
Misyak & Christiansen (2012)	Adj	Aud-V	0.14			
	Nadj	Aud-V				
		Vis-NV		0.58		
Siegelman & Frost (2015)	Adj	Aud-NV		-0.29	0.23	
		Aud-V		0.11	0.10	0.63
	Nadj	Aud-V		-0.22	-0.25	-0.04
						0.31

Valeurs des corrélations (r) intra-tests (en bleu) et inter-tests (en noir). Types de dépendances : Adj = adjacentes, Nadj = non-adjacentes. Tâches : Aud-V = auditive verbale, Aud-NV = auditive non-verbale, Vis-NV = visuelle non-verbale.

Stabilité des Performances Individuelles

Nous avons observé une stabilité relative des performances des babouins au sein d'une même tâche et pour un même type de dépendance. A notre connaissance, seule l'étude de Siegelman et Frost (2015) a également évalué la stabilité intra-test des performances de leurs participants. Comme le montre la Table 5, ils ont observé une certaine répliquabilité intra-test, qui varie cependant d'une tâche à l'autre. Nos données suggèrent que les performances individuelles des babouins dans une tâche d'apprentissage statistique donnée permettent de prédire leurs performances ultérieures dans cette même tâche et pour le même type de dépendance, mais la conduite de répliquations additionnelles serait nécessaire pour confirmer la tendance observée.

Effets de la Tâche et du Type de Dépendance

Un examen des corrélations inter-tâches et inter-dépendances a révélé que les performances individuelles des babouins étaient influencées par le type de dépendance à apprendre (adjacente ou non-adjacente), tandis que la tâche ne semble pas avoir d'effet majeur. Les performances des babouins dans les expériences impliquant l'apprentissage de dépendances adjacentes d'une part, et celles impliquant l'apprentissage de dépendances non-adjacentes d'autre part, ne sont pas corrélés entre elles. Ces résultats sont similaires à ceux obtenus dans la littérature chez l'humain (cf. Table 5).

A la différence de Siegelman et Frost (2015) cependant, nos résultats indiquent que les performances des babouins dans une tâche corrélaient de manière significative ou tendaient à l'être, avec leurs performances dans l'autre tâche, pour un type de dépendance donné. Cette différence de résultats entre ces deux études pourrait toutefois découler des tâches que nous avons employées. En effet, bien que nos deux tâches présentent des différences importantes (la tâche visuo-spatiale impliquant la prédiction de la localisation d'un stimulus, tandis que la tâche visuelle implique la prédiction de sa forme), elles impliquent toutes deux la modalité visuelle. Nos résultats suggèrent donc *a minima* qu'au sein de cette modalité, pour un même type de dépendance, les performances individuelles des babouins tendent à être stables d'une tâche à l'autre.

Apprentissage Statistique et Mémoire de Travail

Nous avons fait l'hypothèse que l'extraction de régularités non-adjacentes serait susceptible d'impliquer la mémoire de travail de manière plus importante que celle de régularités adjacentes. Nous avons donc testé si des différences inter-individuelles de mémoire de travail, mesurée à l'aide d'une tâche de SOPT, corrélaient avec les performances des babouins dans les expériences d'apprentissage de dépendances non-adjacentes. Notre hypothèse est partiellement supportée par les analyses corrélationnelles effectuées. Les babouins qui parviennent à maintenir en mémoire de travail le plus grand nombre d'items dans la tâche de SOPT sont également ceux ayant obtenu les meilleures performances dans l'une de nos expériences impliquant des dépendances non-adjacentes. Chez l'humain, Misyak et

Christiansen (2012) rapportent une forte corrélation positive entre une mesure de la mémoire de travail (verbale) et les performances de leurs participants pour l'apprentissage de dépendances non-adjacentes ($r = 0.53, p < 0.01$). Toutefois, cette mesure est également corrélée avec l'apprentissage de dépendances adjacentes ($r = 0.46, p < 0.05$). De leur côté, Siegelman et Frost (2015) ne trouvent aucune corrélation entre les performances de leurs participants pour l'extraction de dépendances non-adjacentes et deux mesures de la mémoire de travail (mémoire de travail générale et verbale).

Conclusions

Les données issues de nos analyses corrélationnelles suggèrent que les babouins étant le plus sensibles aux dépendances non-adjacentes ne sont pas ceux présentant la plus forte sensibilité pour les dépendances adjacentes, et *vice-versa*. Ces résultats sont compatibles avec deux hypothèses de la littérature : (1) l'apprentissage de dépendances adjacentes et non-adjacentes est supporté par des mécanismes de nature différente (e.g. Peña, Bonatti, Nespor, & Mehler, 2002), et (2) ces deux apprentissages reposent sur les mêmes mécanismes d'apprentissage statistique, mais l'apprentissage de dépendances non-adjacentes est susceptible d'impliquer de manière plus importante la mémoire de travail (e.g. Wilson, Marslen-Wilson, & Petkov, 2017). Les analyses corrélationnelles que nous avons menées pour examiner les liens potentiels entre l'apprentissage de dépendances non-adjacentes et la mémoire de travail, ainsi que ceux de la littérature chez l'humain, ne soutiennent que partiellement cette deuxième hypothèse. Des recherches supplémentaires sur cette question seraient nécessaires pour apporter des conclusions plus solides, par exemple en examinant si les capacités d'apprentissage de dépendances non-adjacentes de différentes espèces corrélerent ou non avec leurs capacités de mémoire de travail.

Un dernier aspect de nos résultats retient notre attention : nous avons observé une tendance à des corrélations négatives entre les performances des babouins pour l'apprentissage de dépendances adjacentes et non-adjacentes, pour certaines de nos expériences. Une corrélation négative a également été observée par Siegelman et Frost (2015) entre deux de leurs tâches ($r = -0.25, p < 0.05$; cf. Table 5). Ces résultats suggèrent que les individus présentant la plus forte sensibilité aux dépendances adjacentes ne seraient pas seulement pas les plus sensibles aux dépendances non-adjacentes, mais qu'ils seraient moins sensibles à ces

dépendances, et *vice-versa*. Ces observations suggèrent une troisième hypothèse : une sensibilité importante aux régularités locales pourrait se faire au détriment de la sensibilité à des régularités non-adjacentes. À l'inverse, une focalisation moins importante sur les régularités locales pourrait permettre aux autres individus d'être sensibles aux régularités plus distantes.

Références

- Bates, E., & Elman, J. (1996). Learning rediscovered. *Science (New York, N.Y.)*, 274(5294), 1849-1850.
- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 31(1), 24-39. <https://doi.org/10.1037/0278-7393.31.1.24>
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: pitting abstract against stimulus-specific representations. *Psychological Science*, 17(10), 905-912. <https://doi.org/10.1111/j.1467-9280.2006.01801.x>
- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 30(5), 1119-1130. <https://doi.org/10.1037/0278-7393.30.5.1119>
- Elman, J. L. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. MIT Press.
- Endress, A. D. (2010). Learning melodies from non-adjacent tones. *Acta Psychologica*, 135(2), 182-190. <https://doi.org/10.1016/j.actpsy.2010.06.005>
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical Learning in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, 52(2), 321-335. [https://doi.org/10.1044/1092-4388\(2009/07-0189\)](https://doi.org/10.1044/1092-4388(2009/07-0189))
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 28(3), 458-467.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117-125. <https://doi.org/10.1016/j.tics.2014.12.010>
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review*, 16(3), 486-490. <https://doi.org/10.3758/PBR.16.3.486>

- Hsu, H. J., Tomblin, J. B., & Christiansen, M. H. (2014). Impaired statistical learning of non-adjacent dependencies in adolescents with specific language impairment. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00175>
- Lum, J. A. G., Conti-Ramsden, G., Morgan, A. T., & Ullman, M. T. (2014). Procedural learning deficits in specific language impairment (SLI): A meta-analysis of serial reaction time task performance. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 51(100), 1-10. <https://doi.org/10.1016/j.cortex.2013.10.011>
- Minier, L., Fagot, J., & Rey, A. (2016). The Temporal Dynamics of Regularity Extraction in Non-Human Primates. *Cognitive Science*, 40(4), 1019-1030. <https://doi.org/10.1111/cogs.12279>
- Misyak, J. B., & Christiansen, M. H. (2012). Statistical Learning and Language: An Individual Differences Study. *Language Learning*, 62(1), 302-331. <https://doi.org/10.1111/j.1467-9922.2010.00626.x>
- Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-Line Individual Differences in Statistical Learning Predict Language Processing. *Frontiers in Psychology*, 1. <https://doi.org/10.3389/fpsyg.2010.00031>
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127-162.
- Onnis, L., Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, 53(2), 225-237. <https://doi.org/10.1016/j.jml.2005.02.011>
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science (New York, N.Y.)*, 298(5593), 604-607. <https://doi.org/10.1126/science.1072901>
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27-52.
- Saffran, Jenny R., & Thiessen, E. D. (2007). Domain-General Learning Capacities. In E. Hoff & rilyn Shatz (Éd.), *Blackwell Handbook of Language Development* (p. 68-86). Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470757833.ch4>
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal*

Society of London. Series B, Biological Sciences, 372(1711).

<https://doi.org/10.1098/rstb.2016.0059>

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105-120.

<https://doi.org/10.1016/j.jml.2015.02.001>

Sigman, M., & Gilbert, C. D. (2000). Learning to find a shape. *Nature Neuroscience*, 3(3), nn0300_264. <https://doi.org/10.1038/72979>

Wilson, B., Marslen-Wilson, W. D., & Petkov, C. I. (2017). Conserved Sequence Processing in Primate Frontal Cortex. *Trends in Neurosciences*, 0(0).

<https://doi.org/10.1016/j.tins.2016.11.004>

Annexe 1 Mode de calcul de la somme des scores obtenus dans la tâche SOPT.

Sujet	2	3	4	5	6	Somme
ANG	0.96	0.68				1.64
ARI	0.95	0.79	0.53			2.27
BOB	0.96	0.79	0.51			2.26
BRI	0.97	0.66	0.29			1.92
DAN	0.98	0.91	0.64			2.54
DOR	0.98	0.93	0.80	0.63		3.34
DRE	0.98	0.93	0.73	0.49		3.13
EWI	0.99	0.88	0.66	0.41		2.94
FAN	0.99	0.96	0.85	0.68	0.47	3.95
FEL	0.96	0.86	0.62			2.44
FEY	0.98	0.89	0.70	0.44		3.00
FIL	0.98	0.86	0.64	0.37		2.86
FLU	0.97	0.86	0.61			2.43
HAR	0.96	0.71	0.41			2.09
HER	0.98	0.83	0.57			2.37
KAL	0.96	0.46				1.42
VIO	0.97	0.73	0.43			2.13
VIV	0.95	0.67				1.62

Dans les colonnes intitulées 2-6 sont indiqués les scores moyens obtenus par chaque sujet sur l'ensemble de l'année 2015, pour chaque niveau de difficulté de la tâche (de deux à six items à sélectionner).

V. ETUDE IV – Traitement de Grammaires Supra-Régulières chez le Babouin

Malassis, R., Fagot, J., & Dehaene, S. (*en préparation*).

Nous avons vu que les études réalisées sur l'apprentissage de grammaires supra-régulières n'offraient pas pour le moment de démonstration non-ambiguë de cette capacité chez une autre espèce que l'humain (cf. section II.2.3.3 du volet théorique). Toutefois, nous avons émis l'hypothèse que ces échecs puissent s'expliquer par un biais des singes vers le traitement d'informations locales (e.g. détection de répétitions immédiates, apprentissage de dépendances adjacentes), et qu'ils ne révèlent donc peut-être pas une réelle incapacité des différentes espèces NH testées à traiter ce type de grammaire.

Dans le travail présenté ci-dessous, nous avons introduit deux modifications par rapport à ces études précédentes. La première de ces modifications a consisté à s'intéresser aux grammaires centrée-emboîtée et croisée plutôt qu'à la grammaire A^nB^n . En effet nous avons vu que cette dernière avait été largement utilisées jusqu'ici pour la conduite d'expériences chez l'animal NH, sans succès. Les problèmes méthodologiques inhérents à l'utilisation de cette grammaire (présence d'indices locaux et de surface rendant non nécessaire le traitement de la grammaire sous-jacente) nous ont conduit à considérer que cette grammaire ne constitue pas un matériel à privilégier pour évaluer l'hypothèse de la supra-régularité. De plus, comparer l'apprentissage des grammaires centrée-emboîtées et croisées permet de situer plus finement les capacités syntaxiques d'une espèce donnée (Fitch & Friederici, 2012), puisqu'elles sont respectivement une grammaire non-contextuelle et une grammaire contextuelle (cf. Figure 3, p. 20). Nous avons donc testé la capacité des babouins à traiter d'une part des séquences comportant plusieurs dépendances organisées de manière centrée-emboîtée, en employant une grammaire en miroir (*Mirror grammar*) dans laquelle la seconde moitié de la séquence est l'image en miroir de la première moitié (Figure 28). Nous avons d'autre part testé leur capacité à traiter des séquences comportant plusieurs dépendances croisées, en utilisant une grammaire en copie (*Copy grammar*) dans laquelle la seconde moitié de la séquence est une copie de la première moitié.

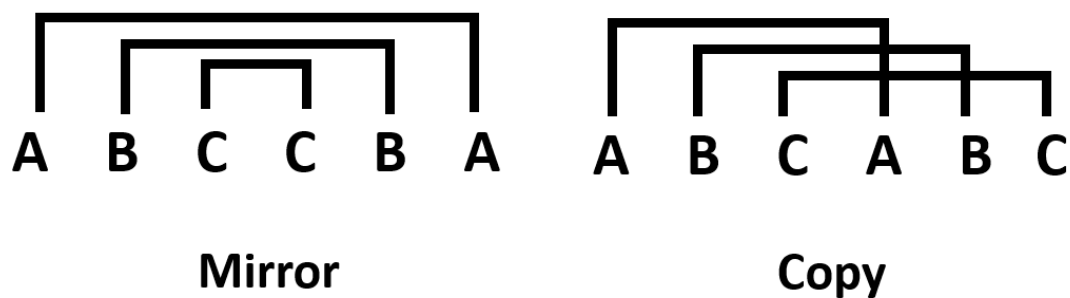


Figure 28 Illustration de l'organisation centrée-empoîtée des dépendances dans une séquence générée par une grammaire Mirror (à gauche), et de l'organisation croisée des dépendances dans une grammaire Copy (à droite).

La seconde nouveauté implémentée dans notre étude est que nous avons testé les babouins dans une tâche de prédiction plutôt que dans une tâche de classification. Ceci nous permet de prévenir l'exploitation par le babouin d'une seule régularité locale, les babouins devant répondre à chaque élément de la séquence. Cette méthode permet également de nous renseigner sur les prédictions des babouins sur chacun des éléments de la séquence.

Dans cette étude, les babouins ont d'abord été exposés à un ensemble de 64 séquences de la forme AB|BA (Expérience Mirror) ou AB|AB (Expérience Copy), générées à partir de 16 éléments. Nous avons choisi d'utiliser un nombre important d'exemplaires afin de favoriser une généralisation de la réponse à d'autres séquences ayant les mêmes propriétés structurelles (Katz, Wright, & Bodily, 2007; Maugard, Wasserman, Castro, & Fagot, 2014).

Les babouins étaient ensuite soumis à une série de tests, dans lesquels de nouvelles séquences de longueur étendue (i.e. comportant six éléments au lieu de quatre) leur étaient présentées. Certaines de ces séquences étaient congruentes avec la grammaire (Mirror : ABC|CBA ; Copy : ABC|ABC), et d'autres non-congruentes, leurs deux derniers éléments étant inversés (Mirror : ABC|CAB ; Copy : ABC|ACB). Les temps de réponse sur chaque élément de la seconde moitié de la séquence étaient mesurés. Nos prédictions étaient que si les babouins détectent la violation de la structure, alors leurs temps de réponse devraient être plus élevés dans les séquences non-congruentes que dans les séquences congruentes.

Les résultats révèlent que les babouins ont réagi aux violations de la grammaire Mirror, y compris pour les séquences test ne comprenant que des transitions adjacentes nouvelles. En

revanche, ils n'ont pas réagi aux violations de la grammaire Copy. En d'autres termes, nos résultats suggèrent que les babouins ont été sensibles dans cette tâche à une organisation centrée-emboîtée, mais pas à une organisation croisée des dépendances.

Pour éviter toute redondance avec l'introduction de cette thèse, je ne présente ici que les sections « méthodes » et « résultats » de ce travail. Je reviendrai sur l'interprétation de ces résultats dans le volet discussion de ce manuscrit.

Methods

General Procedure

This experiment used an adaptation of the serial response time task (Nissen & Bullemer, 1987) which is illustrated in Figure 1. In practice, each sequence was split in two halves. In each trial, the baboons firstly saw the first half of the sequence, corresponding to several targets appearing successively in specific locations of a 4 x 4 matrix on a touch screen. No response to this first part of the sequence was required. After a short delay, the second half of the sequence was presented on the touch screen, and the subjects had to touch the target, each time it appeared on the screen (Figure 1). Response times collection on those touches provides a measure of the participants expectations about the location of the incoming target. Requiring responses on the second half of the sequence only allowed us to compare response times on the same motor sequences across conditions and experiments. All baboons ($n = 14$) performed sequentially both mirror & copy experiments, order being counterbalanced across participants.

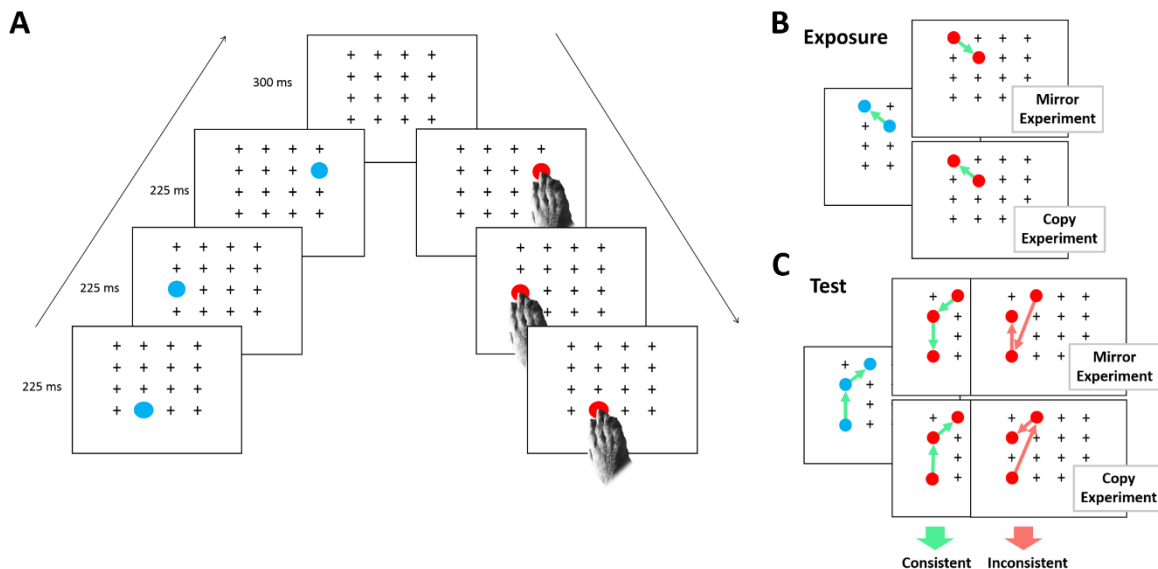


Figure 1 (A) Timeline of a trial (example of a 6-target mirror sequence). After seeing the first half of the sequence (blue targets) displayed on a touchscreen, baboons had to touch the (red) targets of the second half. (B, C) General principles of the experimental design. (B) Baboons were first exposed to 4-target sequences of the form AB | BA (mirror experiment) or AB | AB (copy experiment). (C) They were then tested on consistent and inconsistent 6-target sequences.

Table 1 present an overview of the experiments timeline. Baboons were first presented with sequences of four targets, following either a mirror (AB|BA) or a copy structure (AB|AB).

After this exposure phase, baboons received a series of four tests (Test 1, Test 2, Test 3, Control), in which sequences of six targets, instead of four, were presented to the baboons. These tests assessed whether the baboons were sensitive to the structure underlying the targets' displacements in novel sequences of extended length.

Table 1 Overview of the experiments timeline.

Phases	Description	Conditions	Prop. of Trials	Sequences structure		No. of blocks
				Mirror experiment	Copy experiment	
Exposure	4-target sequences	Baseline	100%	AB BA	AB AB	40
Test 1	6-target sequences	Baseline	4/6			4
		Consistent	1/6	ABC CBA	ABC ABC	
		Inconsistent	1/6	ABC <u>CAB</u>	ABC <u>ACB</u>	

Interruption (2 months)						

Control	Presentation of the different sequence structures at equal frequencies	Mirror-Consistent	1/4	ABC CBA		4
		Mirror-Inconsistent	1/4	ABC CAB		
		Repeat-Consistent	1/4	ABC ABC		
		Repeat-Inconsistent	1/4	ABC <u>ACB</u>		

Exposure	4-target sequences	Baseline	100%	AB BA	AB AB	20
Test 2	6-target sequences containing familiar 4-target sequences	Consistent	1/7			4
		Inconsistent	1/7			
		Baseline	3/7	= Test 1		
Test 3	6-target sequences made of novel targets combinations	Consistent	1/7			
		Inconsistent	1/7			

Note: All baboons (n = 14) performed sequentially both mirror & repeat experiment, order being counterbalanced across participants. Tests 2 and 3 were conducted concurrently.

In Test 1, baboons were presented with two types of test sequences presented at equal frequencies: consistent sequences (mirror: ABC|CBA; copy: ABC|ABC), and inconsistent ones, in which the last two targets were presented in a reversed order. These sequences were of the form ABC | CAB in the mirror experiment, and ABC|ACB in the copy experiment (violations underlined). In all those test sequences, the targets appeared at locations on the matrix which were never used before within sequences of six targets.

The structure of the consistent and inconsistent sequences was similar in Tests 2 and 3, in which we manipulated the novelty of the subsets of four targets composing the complete sequences of six targets. In Test 2, test sequences contained baseline sequences. For instance, the inconsistent mirror sequence 4 5 3 | 3 4 5 (these numbers refer to different locations on the matrix, see *Supplementary Information*) involved violations of the subsequences 4 5 | 5 4 and 5 3 | 3 5, which are both 4-target baseline sequences encountered a substantial number of times during the exposure phase. In Test 3, the sequences did not contain baseline sequences, as they were constructed with novel combinations of locations.

In the Control test, the four types of sequences (mirror consistent, mirror inconsistent, copy consistent, copy inconsistent) were presented at equal frequencies. This test was administered after a 2-month interruption of the experiment.

Predictions

This series of tests allowed precise predictions about the baboons' pattern of response times, depending on the regularities learned during the exposure phase. Our rationale was that detection of the violations should result in slower response times on inconsistent sequences compared to consistent sequences. If baboons relied on rote memorization of the baseline sequences, they should succeed in Tests 1 and 2, but should fail to notice the violations occurring in Test 3. By contrast, if baboons' behavior was triggered by idiosyncratic features of the different sequences types, rather than sensitivity to the underlying structure, we should observe in the Control phase a pattern of responses similar to Tests 1-3 phases. Finally, if baboons are sensitive to the mirror and/or copy structure, they should show slower response times on inconsistent trials in Tests 1-3, while no difference would be observed in the Control test.

Participants and apparatus

Participants were 14 Guinea baboons (*Papio papio*, 8 females, age range 4-21 years) from the CNRS primate facility (Rousset-sur-Arc, France). They live within a larger group of 26 individuals, within a 700 m² outdoor enclosure and had a permanent access to ten Automated

Learning Devices for Monkeys (ALDM, for a detailed description, see (Joël Fagot & Bonté, 2010b; Joël Fagot & Paleressompouille, 2009) equipped with a 19-inch touch screen and a food dispenser. The main feature of ALDM equipment is that a radio frequency identification reader (RFID) identifies each baboon via a microchip implanted in each arm. The baboons can therefore participate to the research at will, without being captured, as the test programs recognize them automatically. The experiment was controlled by the EPrime software (Version 2.0, Psychology Software Tools, Pittsburgh).

Task

Each trial began by the display of a fixation cross (120 x 120 pixels) presented at the bottom of the screen. Touching this stimulus triggered the display of a matrix of 16 (4 x 4) white crosses (60 x 60 pixels each), one of these crosses (hereafter, location) being replaced by a target stimulus (a blue circle of 80 x 80 pixels). This target was displayed during 225 ms, before it moved to another location on the screen for another period of 225 ms. There was no delay between the offset of the target and its next display. During this first phase of the trial, participants should not touch the screen. After disappearance of the second target, the screen remained devoid of targets (only the matrix of locations remained display) for an interval of 125 ms. This interval signaled the middle of the sequence. In the second part of the trial, participants had to touch the targets. Another target (red circle of 80 x 80 pixels) was used to facilitate the distinction between the two phases of the trial. A correct touch of this target immediately triggered the reappearance of the target at another location of the matrix. Trial ended when the participant touched this last target. An accurate completion of the second half of the sequence delivered grains of dry wheat inside the ALDM test unit. An incorrect response (i.e., selection of an incorrect location, or a touch on the screen during the display of the first half of the sequence or the interval signaling the middle of the sequence) stopped the trial and triggered a 5-sec timeout without food reward. Trials in which the participants failed to select a stimulus within 5-sec after the third or fourth target's appearance aborted and presented again in the next trial. This procedure was followed for sequences of four targets. The same procedure was used for sequences of six targets, with the exception that the interval signaling the middle of the sequence lasted 300 ms instead of 125 ms. The time elapsed between the appearance of each target and the baboon's touch on this target were recorded as the response times.

Experimental design

We describe below the procedure used for each phase of the experiments. Each experiment involved exposure phases, a pre-test phase, and four testing phases (Tests 1-3, Control). The lists of all the sequences used in each phase are available in *Supplementary Information*. Unless otherwise noted, for each phase, the sequences were generated with the following constraints: (1) the frequency of each location across all sequences and within each position was equated, (2) repetitions were avoided (i.e. no location could be used twice within a sequence half). Moreover, within all pre-test and test phases, sequences were equated for motor constraints across conditions and experiments (i.e. the second halves of the sequences were identical in the mirror consistent, mirror inconsistent, copy consistent and copy inconsistent conditions). In addition, to prevent learning of the novel sequences during the testing periods, the first halves of the sequences were also equated between conditions, rendering the second half unpredictable on that basis. Sequences from the different conditions therefore differed from each other only regarding the matching between their two halves. For instance, for a given combination of three locations {3, 4, 5}, the following sequences were generated: 453|**354**, 543|**354**, 453|345, and 543|345 (mirror experiment); 345|**354**, 354|**354**, 345|345, and 354|345 (copy experiment; bold and underlined characters denote motor sequences shared across conditions and experiments; note that first halves of the sequences are equated between conditions within each experiment). Custom-written Python codes (www.python.org) were used to generate the sequences.

Exposure phases

For the exposure phases, sixty-four 4-target baseline sequences were constructed. Each of the 16 locations of the matrix was presented four times in the first position of the sequence and combined with four different locations presented in second position of the sequence (i.e. $16 * 4 = 64$ sequences; transitional probabilities between the first and second locations = .25). The same combinations of locations were used to generate the mirror and copy baseline sequences. Before Test 1, as well as Tests 2-3, baboons performed an exposure phase of twenty 96-trial blocks during which these 64 4-target baseline sequences were presented. In addition, right before each test (including control), baseline sequences of six targets were presented to the baboons during one block to familiarize them with the introduction of additional targets.

For Tests 1-3, these sequences were the same baseline sequences as those presented during the corresponding tests. For the Control test, a different set of sequences was generated, having the same balanced structure ($1/4$ of each sequence type).

Pre-tests

Baboons received two pre-tests, aiming at verifying that they paid attention to the first half of the sequences. In the first pre-test, $1/6$ of the trials were consistent sequences consisting in novel combinations of locations, and $1/6$ of the trials were made of the same novel combinations but followed the reverse pattern (i.e. in the mirror experiment, consistent and inconsistent sequences were of the form AB | BA and AB | AB, respectively, and *vice-versa* in the copy experiment). The remaining $4/6$ of the trials were 4-target baseline sequences (the same as those presented during the exposure phases), randomly intermixed with the novel sequences. Baboons performed four 96-trials blocks. Results were analyzed using a repeated-measures ANOVA conducted on RTs for the third target and involving the experiment (Mirror, Copy) and the condition (Consistent, Inconsistent) as within-participant factors. This analysis revealed an Experiment * Condition interaction ($F(1, 13) = 70.92, p < .001, \eta_p^2 = .85$). Post-hoc Tukey's HSD tests revealed that baboons were faster to respond to the third target in sequences of the form AB | BA in both experiments (mirror experiment: Mean \pm SD = 307 ± 85 ms; copy experiment: 415 ± 69 ms) compared to sequences of the form AB | AB (mirror: 469 ± 43 ms; copy: 433 ± 36 ms). Results from this first pre-test confirm that baboons paid attention at least to the last displayed target. But they also indicated that immediate repetition of a target in the middle of the sequence triggered shorter response times compared with non-repeated target, regardless of the experiment. This observation is consistent with previous findings suggesting that immediate repetition of an item within a sequence is a particularly salient feature (e.g. (M. H. de Vries et al., 2008; J. Gervain et al., 2008); Malassis & Fagot, *in prep*). Following these observations, we decided to avoid applying violations to the first target occurring after the middle of the sequence.

In a second pre-test, baboons were therefore presented with inconsistent sequences of the form AB | BC (mirror experiment) and AB | AC (copy experiment), the last target appearing at a location that differed from those of the first half of the sequence. Experimental settings were otherwise identical to Pre-test 1. A repeated-measures ANOVA was conducted on RTs on the fourth target and involving the experiment (Mirror, Copy) and the condition (Consistent,

Inconsistent) as within-participant factors. It showed a main effect of condition, $F(1, 13) = 10.62, p < .01, \eta_p^2 = .45$. Baboons' response times were overall slower in inconsistent (Mean \pm SD = 466 ± 85 ms) compared to consistent sequences (443 ± 99 ms). No other main effect or interaction were found ($ps > .38$). These results indicate that baboons detected the introduction of an incorrect location at the end of the sequence in both experiments, confirming that they paid attention to the first half of the sequences.

Test 1

In Test 1, $1/6$ of the trials were consistent sequences (mirror: ABC | CBA, copy: ABC | ABC), and another $1/6$ of the trials were inconsistent sequences in which the last two targets were presented in a reversed order (mirror: ABC | CAB, copy: ABC | ACB). We chose to apply violations to the last two targets (instead to, e.g., reversing the order of the fourth and fifth targets) because results from our first pre-test revealed that the presence/absence of a repetition of a location in the middle of the sequence was particularly salient for the baboons. In the remaining $4/6$ of the trials, a novel set of six baseline sequences made of six instead of four targets. These sequences were randomly intermixed with the consistent and inconsistent sequences.

To construct the three types of sequences presented during Test 1, we randomly split the 16 possible locations in two sets. A first set of six locations (locations 1, 2, 3, 10, 13 and 15) was used to generate six baseline sequences for each experiment. The remaining 10 locations were used to generate 80 test sequences (20 sequences per condition and experiment). Consistent and inconsistent sequences of Test 1 were therefore made of targets' locations which never occurred in sequences of six targets before.

Test 1 involved four 100-trial blocks (60 baseline, 20 consistent and 20 inconsistent sequences per block). As noted previously, the baboons received a re-exposure phase of twenty 96-trial blocks before to perform Test 1, during which they were exposed to the set of 64 4-target baseline sequences. In addition, they were exposed right before Test 1 to the set of six 6-target baseline sequences during one block, to familiarize them with the introduction of additional targets.

Control Test

The control test was conducted between Test 1 and Tests 2-3. To limit potential influences of previous exposure on the Control phase, this test was administered after a 2-month of break, during which the baboons were exposed to different unrelated tasks. During this test, four types of 6-target sequences (mirror consistent, mirror inconsistent, copy consistent, copy inconsistent) were presented at equal frequencies, randomly intermixed. The control test was conducted with the same set of sequences as in Test 3, to favor comparison across tests types. Control phase involved four 64-trial blocks (16 sequences of each structure per block).

Tests 2-3

Tests 2 and 3 were conducted concurrently. In that phase, $\frac{1}{7}$ of the trials were consistent 6-target sequences made of familiar 4-target baseline sequences, $\frac{1}{7}$ of the trials were inconsistent sequences made of the same baseline sequences, $\frac{1}{7}$ of the trials were consistent sequences made of novel targets combinations, and $\frac{1}{7}$ of the trials were inconsistent sequences made of the same novel targets combinations. During the remaining $\frac{3}{7}$ of the trials, baboons were presented with a novel set of 6-target baseline sequences, randomly intermixed with the consistent and inconsistent sequences.

A novel set of forty-eight 6-target baseline sequences was generated for each experiment, using the 16 locations of the matrix. Thirty-two (16 consistent, 16 inconsistent) test sequences were generated for Test 2 from combinations of 4-target baseline sequences. The same amount of sequences was generated for Test 3, using combinations of locations which were neither presented in any of the 4-target baseline sequence, nor in any of the 6-target baseline sequences generated for Tests 2-3. However, some of these combinations have already been used in other phases (pre-tests and Test 1). Our rationale was that no learning could have taken place during this previous exposure, as the combinations of locations were presented an equal number of time within consistent and inconsistent sequences. As it was impossible to both equate for the frequency of each location within each position and the familiarity *versus* novelty of the combinations of locations, the former variable was not controlled in these last phases (Tests 2 and 3, and Control). Tests 2-3 involved four 112-trial blocks (48 baseline sequences, 16 consistent and 16 inconsistent sequences from each test per block).

Data analyses

Incorrect trials were removed from the data set (3.35 %), as well as RTs greater than three standard deviations from the mean (2.43 % of the remaining trials; computed for each participant, target and phase, after removal of trials with any abnormally long RT, i.e. > 3 seconds). All the analyses used as dependent variable the response times (RTs) on the different targets (i.e., the time elapsed between the apparition of each target and the participant's touch). For the analyses conducted on Tests 1-3, we conducted only planned pair-wise comparisons in relation to our predictions (i.e. assessment of the difference in RTs between consistent and inconsistent sequences), and therefore use two-tailed paired t-tests with Bonferroni-Holm correction (alpha level = .05). Other post-hoc analyses were conducted using Tukey's HSD tests, as we had no predictions regarding other differences in RTs. Statistical analyses were done with R (V3.1.1, www.R-project.org).

Results

Main results

Analyses conducted on the response times (RTs) in Tests 1-3 for each experiment (Figure 2), were performed using a repeated-measures ANOVA involving the condition (Consistent, Inconsistent), experiment (Mirror, Copy), test (1-3), and target (4-6) as within-participant factors. It revealed a significant interaction among condition, experiment and target ($F(2, 26) = 7.43, p < .01, \eta_p^2 = .37$). Pairwise comparisons between RTs in the consistent and inconsistent conditions showed a reliable difference in the mirror experiment for Target 5. This difference corresponded to slower response times on the first violation occurring in mirror sequences. No other significant difference was found between these two conditions. Importantly, no interaction between these factors (condition, experiment, target) and test was found ($p = .44$), indicating that the observed slowdown did not differ among tests. This result confirms that baboons reacted to violations of the mirror structure occurring in test sequences containing familiar baseline sequences, as well as to those occurring in sequences made of entirely novel combinations of locations. Consequently, baboons' results in this mirror experiment cannot be accounted for by rote learning of the baseline sequences, but instead reveal correct generalization to novel sequences following a mirror pattern. By contrast, baboons' response times remained unaffected by the violations occurring in copy sequences.

Taken together, these results suggest that baboons are sensitive to the mirror structure, while they fail to notice, or exploit, the copy structure in the current task.

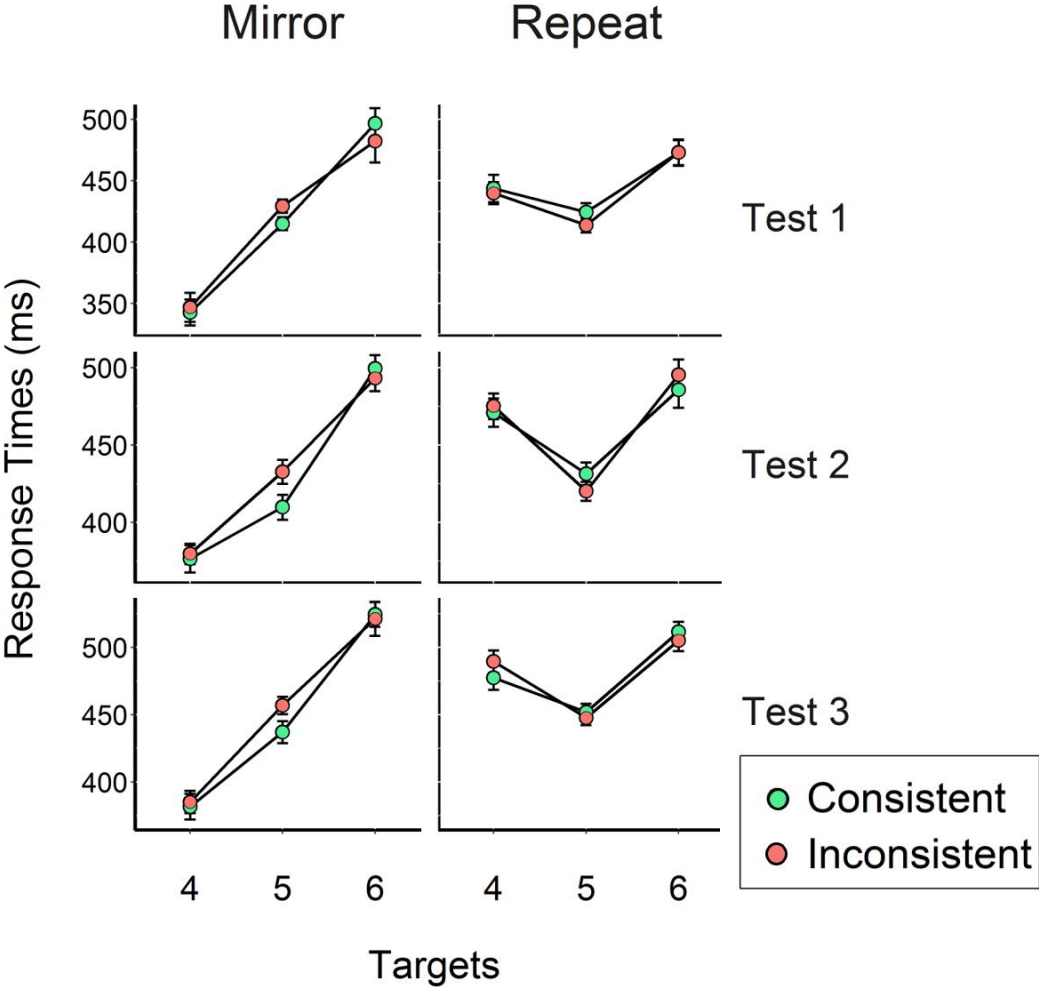


Figure 2 Mean response times on each target and for each condition, in Tests 1-3 of the mirror and copy experiments. Error bars represent standard errors after subtraction of the individuals mean response times.

To tease apart sensitivity to the mirror structure after repeated exposure and potential effects of idiosyncratic features of the different sequences types, data from the Control phase were analyzed. As a consistency effect was only observed for Target 5 in Tests 1-3, we examined RTs on that target as a function of sequence type. We found no difference between mirror consistent (Mean \pm SD = 458 \pm 53 ms), mirror inconsistent (462 \pm 59 ms), copy consistent (456 \pm 41 ms) and copy inconsistent sequences (449 \pm 58 ms), $F(3, 39) = 0.77, p = .52$ in the Control test. This result confirms that when they are presented at equal frequencies, the different sequence types did not elicit differential response times.

Additional results

The ANOVA conducted on data from Tests 1-3 (Figure 2), also revealed a significant interaction between experiment and target, $F(2, 26) = 43.18, p < .001, \eta_p^2 = .77$. Post-hoc analyses showed that RTs differed among the three targets, for both experiments (Mirror: Target 1 < Target 2 < Target 3; Copy: Target 1 > Target 2 < Target 3). RTs were moreover significantly faster in the mirror experiment compared to the copy experiment for Target 1, and longer for Target 3. The ANOVA revealed in addition a main effect of experiment, $F(1, 13) = 26.76, p < .001, \eta_p^2 = .67$, corresponding to overall faster RTs in the mirror (Mean \pm SD = 433 \pm 199 ms) compared to the copy experiment (461 \pm 193 ms). This effect appears however to be attributable to the especially fast RTs on the first target in the mirror experiment. The ANOVA revealed a main effect of target as well, $F(2, 26) = 27.63, p < .001, \eta_p^2 = .68$, corresponding to significantly slower RTs for Target 3 (496 \pm 233 ms) compared to both Targets 1 (416 \pm 195 ms) and 2 (430 \pm 144 ms). Finally, a main effect of test was found, $F(2, 26) = 10.70, p < .001, \eta_p^2 = .45$, RTs being overall greater in Test 3 (466 \pm 208 ms) compared to both Tests 1 (432 \pm 190 ms) and 2 (447 \pm 193 ms). This result suggest that baboons were slower to perform the novel motor transitions involved in Test 3.

DISCUSSION GENERALE



I. Résumé et Discussion des Travaux Expérimentaux

Dans cette thèse, nous avons étudié le traitement de régularités séquentielles et grammaticales dans une perspective comparée. Nous nous sommes plus particulièrement intéressés au traitement des grammaires supra-régulières. Une revue de la littérature sur l'apprentissage de grammaires artificielles a montré qu'aucune étude ne permettait de conclure de manière non-ambiguë quant à la capacité d'une autre espèce que l'humain à traiter ce type de grammaire. L'objectif de cette thèse était d'examiner l'hypothèse selon laquelle ces échecs soient imputables à des difficultés que rencontrent les animaux NH pour le traitement des dépendances non-adjacentes et l'extraction de propriétés structurelles.

Cette hypothèse a été testée à l'aide d'une série de quatre études menées chez le babouin de Guinée, visant d'une part à vérifier que l'extraction de régularités à longue distance pose des difficultés particulières à cette espèce (Etudes I-III), et d'autre part à tester si les babouins s'avéraient sensibles à des grammaires supra-régulières dans des conditions expérimentales prévenant l'exploitation exclusive de régularités locales (Etude IV).

Table 4 Récapitulatif des études réalisées, en fonction de la question étudiée, du type de régularité et du matériel employé.

Question	Etude	Régularités	Matériel
Effet de distance	Etude I	Associations	A_1XB_1, A_2XB_2
	Etude II	Répétitions	AA, ABA
	Etude III	Associations & Répétitions	A_1XB_1, A_2XB_2 (Etude I), AA, ABA (Etude II), $A_1B_1C_1, A_2B_2C_2, A_3B_3C_3$ (Minier et al. 2016)
Supra-régularité	Etude IV	Répétitions	ABCCBA, ABCABC

I.1 Effet de la distance

I.1.1 Etude I

Dans notre première étude, nous avons examiné le traitement d'associations non-adjacentes (A_1XB_1 , A_2XB_2) chez le babouin et l'humain. Dans une série de deux expériences, nous avons mis en évidence des continuités, mais également des discontinuités entre ces deux espèces. Dans l'Expérience 1, nous avons montré que les babouins étaient capables d'apprendre des dépendances non-adjacentes dans une tâche de pointage sériel. La deuxième expérience montre que le babouin partage avec l'humain (e.g. Frost & Monaghan, 2016) la capacité à généraliser les associations non-adjacentes apprises à de nouveaux éléments intermédiaires. Ces données confirment que les babouins ont bien appris les dépendances entretenues entre les éléments non-adjacents, plutôt qu'une collection de triplets. Les performances de généralisation des babouins sont toutefois influencées par des facteurs perceptifs (ici, spatiaux), ce qui fait écho à des résultats obtenus chez l'humain pour l'extraction de régularités dans la modalité auditive (e.g. Creel, Newport, & Aslin, 2004; Onnis, Monaghan, Richmond, & Chater, 2005).

Mais cette étude montre également que l'apprentissage de dépendances non-adjacentes nécessite une quantité d'exposition considérablement plus importante chez les babouins que chez des humains testés dans des conditions expérimentales similaires. En effet, on observe cet apprentissage après 400 présentations de chacune des dépendances non-adjacentes chez le babouin, tandis qu'il est observé dès les 50 premières présentations chez l'humain. Ces données sont à comparer avec celles obtenues dans une autre étude menée au sein de notre laboratoire (Rey, Minier, Malassis, Bogaerts, & Fagot, 2018, en annexe p. 324). Dans cette étude, babouins et humains devaient apprendre des dépendances adjacentes dans une tâche de pointage sériel. Dans ce cas, et pour ce type de dépendances, il n'y avait pas de différence dans les dynamiques d'apprentissage entre ces deux espèces. Les résultats rapportés dans notre l'Etude I ne semblent donc pas refléter une difficulté d'ordre générale pour le babouin à extraire des régularités dans la tâche de pointage sériel. Ces résultats sont particulièrement intéressants au regard de la première hypothèse de travail de cette thèse : l'extraction de régularités à longue distance, qui constitue un prérequis cognitif pour le traitement de grammaires supra-régulières, semble bien présenter une difficulté particulière pour le babouin par rapport à l'humain.

1.1.2 Etude II

Dans notre seconde étude, nous avons examiné l'hypothèse selon laquelle les babouins sont plus sensibles à des régularités structurelles locales (pattern comportant une répétition immédiate d'un élément) qu'à des régularités plus distantes (pattern comportant une répétition non-adjacente). La première expérience montre que les babouins sont capables d'apprendre et de généraliser ces deux types de patterns. Ils réagissent toutefois davantage aux violations de répétitions immédiates qu'à des violations des dépendances non-adjacentes. Dans une deuxième expérience, nous avons examiné si les babouins s'avéraient sensibles aux répétitions même lorsqu'elles n'étaient pas prédictibles, en leur présentant à fréquence égale des séquences de la forme ABB, ABA et ABC. Les données obtenues révèlent que les babouins sont sensibles uniquement aux répétitions adjacentes dans ces conditions. Ils présentent une tendance à répondre plus rapidement à celles-ci dès les tout premiers essais, ainsi qu'une accélération progressive au cours de l'expérience, tandis que leurs temps de réponse restent stables pour les deux autres types de séquences. Nous avons vu dans la section V.1 du volet théorique que la sensibilité aux répétitions adjacentes et non-adjacentes suit chez l'humain un calendrier développemental différent, une sensibilité à ces premières étant observée plus précocement (Gervain, Macagno, Cogoi, Pena, & Mehler, 2008; Gervain & Werker, 2013; Johnson et al., 2009). Nos résultats rejoignent ces données, en montrant que les répétitions immédiates sont particulièrement saillantes pour le babouin. Des expériences comparatives supplémentaires seront nécessaires pour vérifier si ce phénomène est partagé avec d'autres espèces, notamment par des espèces animales qui ne font pas partie de l'ordre des primates. Toutefois, notre étude permet plusieurs conclusions sur les plans méthodologique et théorique.

Du point de vue des méthodes, les résultats de notre étude rendent discutable l'utilisation de procédures comparant des patterns à répétition immédiate (AAB/ABB) à d'autres patterns comportant une répétition non-adjacente (ABA) pour l'étude de l'apprentissage de régularités structurelles (Chen, van Rossum, & ten Cate, 2015; Hauser & Glynn, 2009; Murphy, Mondragón, & Murphy, 2008; Ravignani & Sonnweber, 2017; Spierings & ten Cate, 2016; van Heijningen, Chen, van Laatum, van der Hulst, & ten Cate, 2013). En effet, la plus grande difficulté à apprendre les patterns ne comportant aucune répétition adjacente (ABA) que met en évidence notre étude ne permet pas d'exclure la possibilité que les animaux testés se basent exclusivement sur la détection de la présence d'une répétition immédiate dans ces tâches (ten

Cate & Okanoya, 2012). Pour éviter ce problème d'interprétation, nous préconisons donc d'employer la méthodologie de notre première expérience (sur le modèle de Gervain et al. 2008; Gervain & Werker, 2013), dans laquelle nous avons comparé chacun de ces deux patterns à des séquences ne comportant pas de répétition (ABC). D'autre part, nos résultats tendent à confirmer que la grammaire A^nB^n n'est pas adaptée pour tester l'hypothèse de la supra-régularité. En effet, les séquences générées par cette grammaire (qui consistent en un nombre n d'éléments appartenant à une catégorie A suivis du même nombre n d'éléments d'une catégorie B, les catégories A et B étant marquées perceptivement) comportent nécessairement des répétitions immédiates d'éléments appartenant à une même catégorie perceptive. Les résultats des deux études ayant examiné les stratégies individuelles employées par les animaux (van Heijningen et al., 2009; Ravignani et al., 2015) avaient déjà mis en évidence que la détection de répétitions immédiates était une stratégie privilégiée par une majorité de participants. La présente étude confirme que ces répétitions constituent un indice extrêmement saillant qui nécessite peu d'exposition pour être détecté. Etudier le traitement de grammaires supra-régulières doit donc passer par l'utilisation de grammaires ne permettant pas aux participants d'utiliser cette stratégie (grammaires centrées-emboîtées et croisées).

D'un point de vue théorique, ces données interrogent la nature des mécanismes sous-tendant la détection des répétitions. Nous avons vu dans la section V.3.3 du volet théorique que certains auteurs proposent l'existence de mécanismes spécialisés pour cette opération (Endress, Dehaene-Lambertz, & Mehler, 2007; Endress, Nespors, & Mehler, 2009). Les résultats de notre seconde expérience sont compatibles avec cette hypothèse, puisque d'un point de vue statistique l'occurrence de ces répétitions était équiprobable avec les deux autres types de structure (ABA et ABC). Notre étude suggère donc que si ces mécanismes sont à l'origine la détection des répétitions, ils sont largement partagés au sein de la lignée des primates et opèrent dans différentes modalités sensorielles (ici, dans la modalité visuelle). Ils seraient toutefois limités à la détection de répétitions immédiates chez le babouin. Une explication alternative de la saillance des répétitions immédiates est qu'elles cumulent deux facteurs facilitant l'extraction de régularités statistiques : la similarité perceptive (Creel et al., 2004; Onnis et al., 2005), et leur caractère adjacent (Gebhart, Newport, & Aslin, 2009; Newport & Aslin, 2004; Wilson, Smith, & Petkov, 2015). Cette hypothèse diffère de la précédente en ce qu'elle ne postule pas l'existence de mécanismes *spécialisés* pour la détection de répétitions, mais que cette opération reposerait sur des mécanismes d'apprentissage statistique. Enfin, les résultats de notre seconde

étude illustrent de nouveau un effet majeur de la distance entre les éléments entretenant des relations prédictives sur les performances des babouins.

1.1.3 Etude III

Dans notre troisième étude, nous avons examiné les performances individuelles des babouins dans différentes études (nos Etudes 1 et 2, ainsi que celle de Minier, Fagot, & Rey, 2016). Ces analyses ont révélé que pour un même type de dépendances (adjacentes d'une part, et non-adjacentes d'autre part), les performances sont fortement corrélées d'une expérience à l'autre. En revanche, les babouins qui parviennent à extraire les dépendances non-adjacentes ne sont pas ceux obtenant les meilleures performances dans les tâches d'apprentissage de régularités adjacentes. Ces résultats rejoignent ceux obtenus dans des études employant des méthodes similaires chez l'humain (Misyak & Christiansen, 2012; Siegelman & Frost, 2015). Les études de ce type sont encore peu nombreuses, mais l'étude des différences inter-individuelles offre une porte d'entrée vers une meilleure compréhension des mécanismes d'apprentissage sous-jacents (Frost, Armstrong, Siegelman, & Christiansen, 2015; Siegelman, Bogaerts, Christiansen, & Frost, 2017). Nos résultats montrent que ces questions peuvent être étudiées chez l'animal NH comme elles le sont chez l'humain, en administrant à un même groupe d'individus différentes tâches d'apprentissage de régularités séquentielles.

L'absence de corrélation statistiquement significatives entre performances individuelles dans des tâches d'extraction de régularités adjacentes et non-adjacentes suggère que des contraintes différentes opèrent sur ces deux apprentissages. Notre hypothèse selon laquelle les capacités d'apprentissage de régularités non-adjacentes seraient corrélées avec les capacités de la mémoire de travail de l'individu n'est que partiellement validée par nos analyses. Des recherches supplémentaires sont donc nécessaires pour tester cette hypothèse. Selon une explication complémentaire, cette absence de corrélation pourrait également s'expliquer par des différences inter-individuelles dans les stratégies employées. Pacton, Sobaco, et Perruchet (2015; voir également Pacton & Perruchet, 2008) ont en effet proposé que ces deux types de dépendances sont traitées par les mêmes mécanismes d'apprentissage statistiques, mais que les processus attentionnels contraignent le type de dépendances apprises (voir section V.3.2 du volet théorique). Selon cette théorie, c'est le traitement attentionnel conjoint de deux éléments qui permet d'apprendre les relations qu'ils entretiennent, indépendamment de leur distance. Les

événements proches dans le temps ayant une plus grande probabilité de se retrouver simultanément - ou presque - dans le focus attentionnel, ceci expliquerait que les relations adjacentes soient de plus souvent celles traitées « par défaut » (e.g. Gómez, 2002). Du moins, c'est le cas lorsque d'autres facteurs n'entrent pas en conflit avec cette proximité temporelle. Or, dans nos expériences, les régularités non-adjacentes impliquent systématiquement les éléments situés aux deux extrémités des séquences (e.g. A_1XB_1 , ABA). Les extrémités des séquences sont connues pour être particulièrement saillantes³³. Ainsi, certains babouins pourraient focaliser leur attention sur les événements proches dans le temps (i.e. adjacents), tandis que d'autres se focaliseraient davantage sur les événements situés en début et fin de séquence, ce qui pourrait expliquer au moins en partie les différences inter-individuelles observées.

Enfin, le fait que les performances individuelles des babouins tendent à être corrélées d'une tâche à l'autre pour un même type de dépendance est particulièrement intéressant. La corrélation positive significative observée entre les performances individuelles des babouins dans une tâche d'apprentissage de probabilités transitionnelles entre éléments adjacents (Minier et al., 2016), et dans une tâche de détection de répétitions immédiates (Etude II), supporte l'hypothèse selon laquelle ces deux opérations reposeraient sur les mêmes mécanismes d'apprentissage statistique, plutôt que l'hypothèse selon laquelle la détection de répétitions reposerait sur un mécanisme spécialisé. Il est toutefois impossible de conclure sur la base d'analyses de corrélations, et d'autres méthodes devront être employées pour infirmer cette hypothèse ou la valider.

1.1.4 Conclusions sur les Etudes I-III

Prises dans leur ensemble, ces trois premières études tendent à valider notre première hypothèse de travail, qui était que les babouins rencontrent des difficultés particulières pour l'extraction de dépendances non-adjacentes. Rappelons que chez l'humain, une tendance similaire est observée : les régularités locales tendent à être exploitées en priorité (e.g. de Vries, Monaghan, Knecht, & Zwitserlood, 2008; Wilson et al., 2015). Notre hypothèse n'est donc pas

³³ Du fait des effets classiques de primauté et de récence, ou, selon certains auteurs, de l'existence de mécanismes spécialisés dans l'extraction de régularités présentes en bordure de séquence (e.g. Endress et al., 2009).

qu'il existe une discontinuité entre humains et NH dans la priorité donnée aux informations locales et à plus longue distance. Mais que ce biais vers l'exploitation d'informations locales serait plus fort chez l'animal NH, du fait d'une asymétrie plus importante dans le degré de difficulté que présente l'apprentissage de ces deux types de régularités.

Une limite de cette thèse est qu'elle avait pour objectif de mettre en évidence ce phénomène, et non d'en offrir une explication. Des travaux devront être consacrés à cette question³⁴. En définitive, les résultats de nos trois premières études tendent à confirmer que les échecs des différentes espèces testées sur l'apprentissage de grammaires supra-régulières (e.g. van Heijningen, Visser, Zuidema, & ten Cate, 2009; Ravignani, Westphal-Fitch, Aust, Schlumpp, & Fitch, 2015) puissent être imputés -au moins en partie- à des limitations cognitives pour le traitement de régularités non-adjacentes.

I.2 Traitement de grammaires supra-régulières

I.2.1 Etude IV

Dans notre quatrième et dernière étude, nous avons testé la capacité des babouins à traiter des grammaires supra-régulières: une grammaire en miroir, qui implique des dépendances centrées-emboîtées, et une grammaire en copie, qui implique des dépendances croisées (Fitch & Friederici, 2012). Nous avons proposé en conclusion de la section II.2 (volet théorique) qu'un matériel expérimental permettant l'exploitation d'indices locaux saillants n'était pas adapté pour tester la capacité d'une espèce à traiter une grammaire donnée. Nous avons donc, en complément de notre choix de ne pas utiliser la grammaire A^nB^n , choisi d'employer une tâche d'un type nouveau. Dans une adaptation de la tâche de pointage sériel, les babouins devaient visionner la première moitié d'une séquence de déplacement d'une cible,

³⁴ Les différences de performances pour l'extraction de régularités non-adjacentes entre l'humain et les autres espèces pourraient notamment découler de différences interspécifiques de capacité de stockage en mémoire de travail (Etude III ; Joël Fagot & De Lillo, 2011), de compression de l'information (e.g. Chekaf, Cowan, & Mathy, 2016; Chekaf, Gauvrit, Guida, & Mathy, 2018), ou encore de formation de représentations explicites et le recodage verbal (e.g. Rey et al., 2018).

avant de suivre la cible dans la deuxième moitié de la séquence. Cette tâche implique donc que le participant émette une réponse sur chaque élément prédictible plutôt que d'effectuer un jugement sur la séquence (jugement pouvant être effectué comme nous l'avons vu en exploitant exclusivement une sous-partie de la séquence, van Heijningen et al., 2009; Ravignani et al., 2015). Dans notre tâche, les babouins étaient donc encouragés à traiter l'ensemble des régularités présentes dans les séquences. L'utilisation de cette procédure nous a permis de mettre en évidence des résultats suggérant pour la première fois chez une autre espèce que l'humain la capacité à traiter la grammaire en miroir. En revanche, les babouins se sont révélés insensibles à la grammaire en copie dans cette tâche.

Ces données suggèrent que cette espèce de primates NH possède des capacités computationnelles supra-régulières. Cependant, ces capacités semblent restreintes aux grammaires non-contextuelles comme la grammaire en miroir, puisque les babouins échouent à traiter la grammaire en copie, qui est un exemple de grammaire contextuelle. Ces données sont compatibles avec la hiérarchie proposée par Chomsky, les grammaires non-contextuelles constituent l'ensemble des grammaires supra-régulières ayant la puissance générative la plus restreinte (1956, 1957 ; voir Figure 3, p. 20). Les capacités computationnelles des babouins resteraient donc plus limitées que celles des humains, ces derniers parvenant à traiter des dépendances croisées, dans les langues naturelles (Bresnan, Kaplan, Peters, & Zaenen, 1982; Culy, 1985; Jäger & Rogers, 2012), et dans des paradigmes d'apprentissage de langages artificiels (Uddén, Ingvar, Hagoort, & Petersson, 2012).

Nous proposons cependant une explication alternative de nos résultats. Nos données suggèrent en effet l'existence d'un effet de récence important dans notre étude. L'effet de récence correspond à une meilleure rétention des derniers éléments de la première moitié de la séquence. Cet effet prédit des temps de réponse plus rapides pour les premiers éléments de la seconde moitié de la séquence. Or ceci est bien observé dans l'expérience testant leur capacité à traiter la grammaire en miroir (Figure 29). Un tel effet de récence pourrait avoir facilité l'apprentissage de cette grammaire (Rey, Perruchet, & Fagot, 2012). En suivant la même logique, l'apprentissage de la grammaire en copie aurait pu bénéficier d'un effet de primauté (meilleure rétention des premiers éléments, ce qui prédit des temps de réponse plus courts sur la première cible par rapport à la seconde cible, et sur la seconde par rapport à la troisième). Mais nous n'observons pas cet effet : la Figure 29 montre les temps de réponse des babouins pour la grammaire en copie ne sont pas plus rapides pour le premier élément par rapport au

second. En l'absence de cet effet, cette facilitation potentielle n'est pas disponible pour la grammaire en copie.

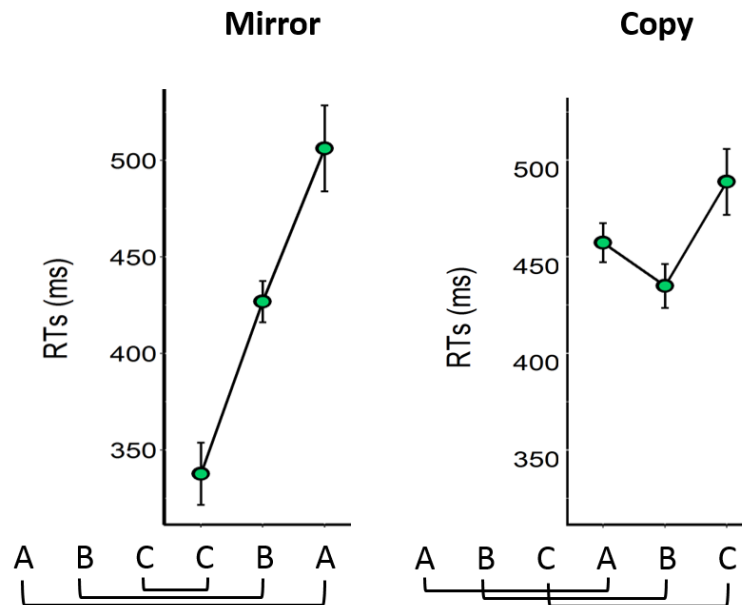


Figure 29 Temps de réponse (RTs) des babouins dans l'Etude IV, moyennées sur les Tests 1-3, pour les séquences congruentes dans les expériences testant l'apprentissage de la grammaire en miroir (à gauche), et de la grammaire en copie (à droite).

Des expériences ayant étudié la mémorisation de séquences de stimuli chez le primate NH ont montré que seul un effet de récence était observé lorsque le délai entre la présentation de cette séquence et le test est court (Wright, Rivera, Katz, & Bachevalier, 2003), comme c'est le cas dans la tâche que nous avons utilisée (avec un délai de 350ms entre les deux moitiés de la séquence). Pour des délais plus longs, en revanche (à partir de deux secondes chez le macaque rhésus, et de dix secondes chez le capucin), on observe également un effet de primauté (Wright et al., 2003). Il serait donc essentiel de tester la capacité des primates NH à traiter des grammaires en copie avec un délai plus important entre les deux moitiés de séquences.

L'idée ici n'est pas de réduire le traitement de grammaires supra-régulières à une tâche de mémorisation de séquence. En effet, l'extraction de régularités grammaticales implique un traitement des propriétés structurelles d'un ensemble de séquences isomorphes, et non pas

seulement la rétention en mémoire de travail d'une séquence³⁵. Mais nous proposons que les propriétés de la mémoire de travail soient susceptibles de jouer un rôle clé dans la définition du type de structure que pourra traiter un système cognitif.

Les résultats de l'Étude IV sont corroborés par un travail mené en parallèle par Wang & Dehaene (soumis) chez le macaque. Dans cette étude, l'apprentissage de grammaires en miroir et en copie a été testé à l'aide d'une tâche de production de séquences. Les séquences étaient séparées en deux moitiés, et pour la première moitié les singes devaient seulement regarder la cible se déplacer, comme dans notre expérience. Mais les singes devaient ensuite non pas suivre les déplacements de la cible, mais produire la seconde moitié de la séquence (i.e. les réponses n'étaient pas indicées). Cette recherche a montré que, dans ce dispositif, les singes étaient capables, après un entraînement conséquent, de produire des séquences en miroir et en copie. De plus, ils ont pu généraliser la grammaire en miroir à de nouvelles séquences et de nouvelles configurations, ce qui confirme que les singes n'ont pas seulement appris un ensemble de séquences, mais bien extrait la règle abstraite sous-jacente. Malheureusement, cette étude ne comporte pas de test de généralisation pour la grammaire en copie, ce qui interdit toute conclusion quant à la capacité du macaque à traiter une grammaire contextuelle.

1.2.2 Conclusions sur l'Étude IV

Notre étude suggère donc que le babouin est capables de maîtriser une grammaire en miroir, et possèdent donc a minima une mémoire « à pile » (Chomsky, 1956, 1957; Fitch, 2014 ; voir la section I.2 du volet théorique). Ce travail révèle donc la capacité d'une espèce animale NH à traiter une grammaire supra-régulière, capacité qui avait été proposée comme étant potentiellement restreinte à l'espèce humaine (Fitch & Hauser, 2004). Nos résultats vont dans le sens de la seconde hypothèse de travail de cette thèse, i.e. que la mise en évidence de la capacité d'une espèce à traiter une grammaire supra-régulière nécessite de prévenir

³⁵ Les données de notre étude montrent bien que les temps de réponse des babouins ne reflètent pas simplement un effet de récence. En effet, dans notre expérience contrôle, dans laquelle nous avons présenté à fréquence égale des séquences congruentes et non-congruentes, les babouins ne se comportent pas différemment en réponse à ces deux types de séquences. En revanche, après une exposition répétée à des séquences générées par la grammaire en miroir, ils réagissent à la violation survenant dans les séquences non-congruentes avec cette grammaire. C'est donc bien une extraction de la structure en miroir que nous observons dans notre étude, mais nous proposons que cet apprentissage ait pu être favorisé par un effet de récence.

l'exploitation exclusive de régularités locales pour effectuer la tâche. Toutefois, le succès du babouin à traiter une grammaire en miroir comme celui du macaque (Wang & Dehaene, soumis), mis au regard des échecs des différentes espèces précédemment testées (singe tamarin, Fitch & Hauser, 2004; diamant mandarin, van Heijningen et al., 2009; kéra et pigeon, Ravignani et al., 2015), peuvent également refléter des différences interspécifiques dans la capacité à traiter des grammaires supra-régulières. La conduite d'études comparatives chez ces différentes espèces, en manipulant de manière systématique le type de grammaire et de tâche utilisés permettra de valider l'une ou l'autre de ces hypothèses.

I.3 Conclusions sur les procédures expérimentales

La conduite d'expériences contrôlés en laboratoire permet, en complément de l'observation des comportements et productions naturelles, d'étudier finement les mécanismes entrant en jeu dans le traitement de régularités grammaticales. Mais cette méthode rencontre fréquemment des critiques, qui portent notamment sur le fait que ces expériences impliquent fréquemment un entraînement conséquent des animaux, comme c'est le cas des expériences que nous avons mené dans cette thèse.

Watumull, Hauser, Roberts, & Hornstein (2014) ont ainsi proposé que la conduite d'expériences comparatives reposant sur un paradigme d'habituation-déshabituaiton³⁶ soit à privilégier par rapport à l'emploi du conditionnement opérant, qui peut impliquer un entraînement sur des centaines, voire milliers d'essais. Ceci repose toutefois sur une sous-estimation de la quantité d'exposition typiquement dispensée dans les tâches d'habituation-déshabituaiton. Dans une expérience testant l'apprentissage de dépendances non-adjacentes chez le tamarin par exemple, Newport, Hauser, Spaepen, et Aslin (2004) exposent les participants à ces régularités durant 21 minutes. Au cours de ces 21 minutes, chaque dépendance

³⁶ Pour rappel, dans les paradigmes d'habituation-déshabituaiton, les participants sont typiquement exposés de manière passive à des régularités (e.g. auditives) durant quelques minutes. Puis ce qu'ils ont appris est évalué en mesurant les réponses d'orientation (latence, durée) du participant lors de la diffusion de séquences congruentes et non-congruentes avec ces régularités. Ces expériences n'impliquent pas de renforcement des réponses des participants. Dans un paradigme de conditionnement opérant, les participants sont renforcés différenciellement (i.e. ils reçoivent une récompense en cas de réponse correcte, tandis qu'ils ne reçoivent pas de récompense et se voient éventuellement infligés une pénalité en cas de réponse incorrecte).

est diffusée 672 fois au total. A titre de comparaison, la première expérience de notre Etude I, dans laquelle nous examinons le décours temporel de cet apprentissage chez le babouin et l'humain, impliquait 500 présentations de chaque dépendance non-adjacente. La quantité d'exposition dispensée dans ces deux études, qui impliquent pourtant des paradigmes expérimentaux bien différents, est donc sensiblement similaire.

Les tâches que nous avons utilisées dans cette thèse correspondaient à un troisième type de paradigme expérimental, reposant sur un renforcement non-différentiel. Les babouins étaient récompensés à chaque fin d'essai (i.e. à chaque complétion correcte de séquence), qu'ils aient extrait ou non les régularités sous-jacentes. Cette méthode a pour avantage de (1) maintenir la motivation des participants, (2) s'assurer qu'ils traitent les séquences, (3) permettre d'avoir une mesure continue de l'apprentissage. Cette procédure offre donc la possibilité de collecter des données comparatives de qualité, puisqu'elle met les animaux dans des conditions de simple exposition aux régularités proches de celles de la plupart des tâches d'apprentissage de langage artificiel conduites chez l'humain.

De manière plus générale, à l'instar de la plupart des expériences menées en laboratoire, les données de cette thèse ont été acquises chez un groupe de singes bien particuliers. En effet, ces individus peuvent être considérés comme étant « surentraînés » à effectuer des tâches cognitives implémentées dans des expériences informatisées. Les données obtenues sont donc susceptibles de ne pas mettre en lumière les performances typiques de cette espèce, mais leurs capacités « potentielles » (Lyn, 2017). En ce sens, la démarche poursuivie dans cette thèse n'était pas d'étudier seulement les types de régularités privilégiés par le babouin, mais d'examiner si dans des conditions expérimentales appropriées, ils parviennent à effectuer des apprentissages plus complexes.

II. Implications pour les Théories sur l'Evolution de la Syntaxe

Nos recherches ouvrent la porte à la conduite de travaux chez le babouin visant à déterminer si elles sont également capables d'enchâssement récursif (cf. sections I.3.2 du volet théorique), la possession de capacités computationnelles supra-régulières constituant un

prérequis à cette opération (e.g. Fitch & Hauser, 2004). Il reste également à déterminer si les babouins sont capables de traiter une grammaire supra-régulière implémentée dans des séquences organisées hiérarchiquement (section I.5).

Mais nos recherches suggèrent également que les contraintes mnésiques jouent un rôle déterminant dans la définition des régularités syntaxiques que peuvent extraire les babouins. Ces contraintes mnésiques sont également observées chez l'humain : un nombre conséquent d'études montre que notre capacité à traiter des phrases suivant une structure centrée-emboîtée est limité à un faible nombre d'emboîtements (Bach, Brown, & Marslen-wilson, 1986; Hakes, Evans, & Brannon, 1976; Karlsson, 2007, 2010; Marks, 1968; George A. Miller & Isard, 1964; M. D. Wang, 1970). Une analyse de corpus effectuée par Karlsson (2007) sur plusieurs langues montre par exemple que les phrases comportant plus de deux niveaux d'emboîtement sont quasiment absentes du langage parlé. Ces données rendent la simple distinction entre grammaires régulières et supra-régulières, basée sur une notion d'infinité virtuelle (voir section I.2 du volet théorique), peu pertinente, comme le soulignent de Vries, Christiansen, et Petersson (2011): « *A theory of psychological processing typically focuses on actual and measurable performance.* » (p. 11). Etant donné les limites perceptives et mnésiques intrinsèques aux systèmes cognitifs biologiques (Petersson, Folia, & Hagoort, 2012), les théories sur l'évolution de la syntaxe prenant en compte ces contraintes pourraient s'avérer plus pertinentes (Christiansen & Chater, 2008, 2015). Wilson, Marslen-Wilson, et Petkov (2017) proposent ainsi de replacer les différents types de régularités séquentielles dans un continuum de complexité croissante, caractérisé par deux dimensions quantitatives : la distance séparant les éléments entretenant des dépendances (i.e. en termes de nombre d'éléments intermédiaires), et le nombre de dépendances à traiter, comme illustré par la Figure 30.

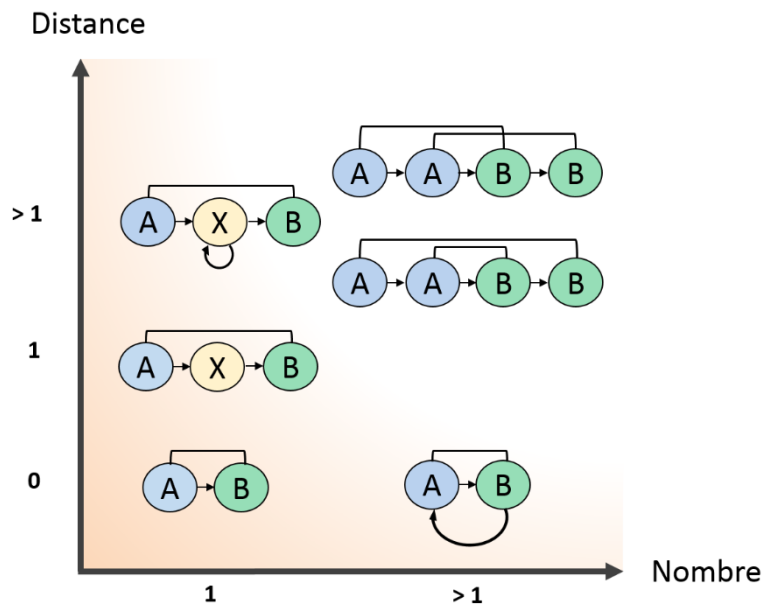


Figure 30 Taxonomie de différents types de régularités séquentielles, replacées au sein d'un continuum de complexité croissante en fonction de la distance séparant les éléments dépendants (axe vertical), et le nombre de ces dépendances (axe horizontal). Modifié d'après Wilson et al. (2017).

De manière plus générale, la perspective adoptée dans cette thèse a été de décomposer le traitement de grammaires supra-régulières en prérequis cognitifs. En poursuivant cette démarche, nous nous sommes inspirés de théories sur l'évolution de la syntaxe qui proposent que la complexité et la productivité des langues naturelles, par rapport aux systèmes de communication des autres espèces, ne s'expliqueraient pas par une apparition saltatoire de certaines capacités spécifiques au cours de l'évolution récente de l'espèce humaine (e.g. la capacité d'effectuer des computations récursives, Hauser et al., 2002 ; voir également Chomsky, 2017), mais par des modifications graduelles des mécanismes cognitifs supportant le traitement des régularités séquentielles (e.g. Christiansen & Chater, 2015; Seidenberg, MacDonald, & Saffran, 2002). Dans cette perspective, les recherches comparatives menées ont pour objectif d'identifier les éléments de continuité et de discontinuité dans le traitement de régularités séquentielles, qui pourraient expliquer les différences de capacités syntaxiques observées entre l'humain et les autres espèces.

Un nombre croissant de recherches comparatives adoptent cette démarche. Par exemple, des travaux de Ghirlanda, Lind, et Enquist (2017) suggèrent des différences importantes entre

les humains et les autres espèces dans le degré de résolution des représentations de séquences. Ces auteurs proposent que ces différences aient des conséquences majeures sur la capacité de ces espèces à discriminer des séquences, et donc à extraire des régularités syntaxiques. Santolin et Saffran (2018) proposent de leur côté de privilégier la conduite de recherches visant à mieux comprendre les propriétés et contraintes des mécanismes d'apprentissage statistique chez les différentes espèces (i.e. qui ne soient pas centrées sur la mise en évidence de capacités supposément spécifiques à l'espèce humaine). Ces recherches pourront nous informer en retour sur les liens que ces mécanismes entretiennent avec le traitement de la syntaxe. Toujours dans cette perspective, nous proposons que l'étude des régularités séquentielles exploitée préférentiellement par les animaux NH permette de mieux comprendre pourquoi ceux-ci se focalisent sur certaines propriétés de la séquence plutôt que d'autres, souvent au détriment de l'extraction de la grammaire sous-jacente (Chen et al., 2015; van Heijningen et al., 2009; Ravignani et al., 2015; Spierings & ten Cate, 2016; van Heijningen et al., 2013). Par exemple, nous avons mené une étude qui suggère que l'information positionnelle (e.g. A apparaît en première position, B en second) est favorisée lorsqu'elle entre en conflit avec l'information relationnelle (A est associé à B), même dans le cas de simples associations adjacentes (Fagot, Malassis, & Medam, 2018, en annexe p. 310). Ce phénomène ne semble donc ne pas être restreint aux régularités syntaxiques complexes.

Dans le même sens, nous proposons d'élargir cette perspective à l'étude des fonctions cognitives qui ne relèvent pas spécifiquement du traitement des régularités séquentielles. Par exemple, des recherches menées chez l'animal NH sur la perception visuelle montrent que le biais vers l'exploitation préférentielle de régularités locales n'est pas seulement observé pour des séquences. On l'observe également dans des tâches requérant de traiter des figures visuelles complexes, impliquant l'intégration de plusieurs sous-éléments (pour une revue, voir Fagot & Barbet, 2006). Ce type de données est susceptible de révéler des contraintes générales opérant sur la cognition NH, qui pourraient expliquer certaines spécificités de la cognition humaine pour des fonctions cognitives de plus haut niveau telles que le langage. Dans la section suivante, nous discutons plus en détail de ces aspects.

III. Perspectives : Etudier l'Evolution du Langage à travers l'Etude Comparative des Fonctions Cognitives Générales

Nous présentons ici en guise d'ouverture un chapitre théorique issu de l'ouvrage "*Origins of human language: continuities and splits with nonhuman primates*" (Eds. L-J Boë, J Fagot, P Perrier, & J-L Schwartz). Ce chapitre est le fruit d'une collaboration au sein de l'équipe. Nous y passons en revue des données comparatives issues de domaines d'études divers, tel le traitement de régularités séquentielles, mais également la perception visuelle ou encore la catégorisation, et recensons un certain nombre de différences se dessinant entre l'humain, d'une part, et les primates NH, d'autre part. Nous argumentons que l'étude comparative de certaines fonctions cognitives générales (i.e. non directement liées au langage) soit fondamentale pour une meilleure compréhension de l'évolution du langage humain.

Joël Fagot, Raphaëlle Malassis,
Tiphaine Medam and Marie Montant
*Laboratory of Cognitive Psychology, Institute of Language,
Communication and the Brain, Aix-Marseille University, and CNRS*

Comparing Human and Nonhuman Animal Performance on Domain-General Functions: Towards a Multiple Bottleneck Scenario of Language Evolution

Abstract: How come that *Homo sapiens* is the sole animal species communicating with a language (i.e., a human language)? Theorists of language evolution have mostly adopted a human-centered approach to address this question. This chapter discusses the limits of this approach and proposes an alternative that consists in studying the domain general functions that serve language comprehension and production from a comparative and evolutionary perspective. Special attention is given to domain general processes which allow humans and animals to integrate information in space and time, and thus develop perceptual and more conceptual abstract categories. This chapter presents illustrative studies that reveal the various aspects in which these integration processes differ in human and nonhuman animals. Finally, we discuss the source of these species differences and their potential implications for our understanding of language evolution.

Keywords: Animal communication, integration processes, categorization, language evolution, baboons

1. Introduction

1.1 Limits of strictly human-centered approaches

In the various scientific disciplines interested in language evolution – anatomy, physiology, paleoanthropology, linguistics, psychology, neurosciences, computer sciences – the precisely defined architecture of human language is classically used as the basic reference, the *canon*, to which all nonhuman communication systems are compared. Said differently, the communication system of *Homo sapiens* -an approximately 200,000 year-old isolated species (there are no other living species in the genus *Homo*) with almost

1.5kg brain tissue, living in complex societies with a long history of cultural evolution -is used as the unique reference system in most comparative studies on language evolution (for a review, Hauser et al., 2002). Obviously, this makes sense given that the purpose of these comparative studies was not to build a descriptive catalogue of other animal communication systems but to better understand human language evolution. However, when pushed too far, such an anthropomorphic approach can be misleading because, by definition, nonhuman animals cannot equal human animals in performance when it comes to human language. A current caveat in human-centered comparisons consists in assuming non-explicitly that nonhuman cognitive architectures must *resemble* human cognitive architecture, in parts or as a whole. However, such an assumption can hold only if: 1) human and nonhuman cognitive architectures had followed similar evolutionary paths and were adapted to comparable environmental, social and biological constraints, and 2) if the cognitive architecture of each species was a construction made of independent (non-interacting) cognitive components that are not sensitive to developmental and phylogenetic factors. Given that every species has a unique cognitive architecture, it seems like a vain enterprise to search for strictly identical components in humans and nonhuman animals.

Consider one example: syntax. Different types of syntaxes have been formalized in Chomsky's hierarchy, from simple (finite state) grammars to complex (supra-regular) grammars. In an attempt to better understand human cognitive "uniqueness", major efforts have been put in the investigation of nonhuman species' ability to process supra-regular grammars (e.g., Fitch and Hauser, 2004; Gentner et al, 2006; Abe and Watanabe, 2011). Unsurprisingly, nonhuman animals do not equal *Homo sapiens* in that particular type of linguistic computation. The very attempt to search for strict human-like syntax in nonhuman animals implies that there is only one way to compute information in a complex communication system: the human way. This view does not consider the possibility that each species, even phylogenetically close to *Homo sapiens*, might have developed its own original and complex –possibly multimodal- cognitive architecture, which does not include human-like syntax. It also neglects the existence of interactions between the various components of a cognitive architecture whose effects increase over phylogenetic and developmental time scales (syntax might not exist as an independent computational subsystem, not even in

humans, see Seidenberg and MacDonald, 1999). These interactions make it very unlikely that complex integrated levels of computation (syntax-like processes) in the cognitive architectures of two different species resemble each other.

We must acknowledge, however, that more and more studies are looking for simpler forms of syntax in nonhuman animals. This approach has recently gained interest in the field of ethology and to some extent in linguistics (for a review, Schlenker et al., 2016; see also Petkov and Wilson, 2012). It refers to basic learning principles that allow the extraction of combinatorial semantics, statistical regularities, or adjacent and nonadjacent dependencies according to which qualitatively comparable basic learning principles are supposed to hold for human as well as nonhuman communication sequence learning (e.g., Seidenberg et al., 2002).

The narrow anthropomorphic approach of language evolution that we question here is not only found in linguistics but has dominated other domains as well, such as comparative anatomy. Rapid progress of imaging techniques over the past 20 years has spurred a frantic search for the anatomical landmark of language in *Homo sapiens*. After a couple of unsuccessful attempts to teach human vocal language to nonhuman primates (e.g., Hayes, 1951), it has become clear that our closest cousins – chimpanzees – are not able to pronounce human phonemes. The first proposed explanation for the incapacity to produce speech was that their larynx was too high (e.g., Lieberman, 1968, 1975). The existence of a physical limitation for speech in chimps has been admitted in the research community for many years. However, recent studies suggest that vocal tract anatomy cannot suffice to explain the absence of speech in nonhuman mammals (for a review, Fitch 2010). Fitch et al. (2016) inferred from X-ray videos and a modeling approach that macaca's tract has the potential to produce a broad range of speech sounds. This was confirmed in one of our recent study in which we recorded the spontaneous vocalization of baboons (*Papio papio*, Boë et al., 2017). We found that the baboons produce sounds that share the acoustic F1/F2 formant properties of human [i æ a o u] vowels, and those baboon sounds were produced by movements of the tongue in a human-like articulatory space defined by two axes (anterior-posterior and superior-inferior, see Boë et al., 2017 for more details). Therefore, the inability of nonhuman primates to produce human phonemes is more likely

due differences in the neural circuits that command oro-facial muscles, or to other cognitive differences, than to differences in the anatomy of the vocal tracts.

Other anatomical factors have been proposed to explain the uniqueness of language in the human species. Most of them concern human-centered brain features. Neuroscientists have been looking for the homologue structures of Broca's area and Wernicke's area in the brain of various apes. Broca and Wernicke are important regions of the human brain whose injury provokes severe language disorders, the so-called Broca's and Wernicke's aphasias. Because of interspecific anatomical differences, direct comparisons of brain regions across species are difficult to make. Instead, researchers have compared brain asymmetries across species. In the human brain, the language function is mostly hosted in the left hemisphere. Broca's area, Wernicke's area and the *planum temporale* are bigger in the left than in the right hemisphere (Geschwind and Levitsky, 1968). Those asymmetries, that were first thought to be unique to humans, have been observed as well in nonhuman primates (e.g., Cantalupo and Hopkins, 2001; Cantalupo et al., 2003). In a recent paper, Gomez-Robles and collaborators (2013) propose that, at a whole brain scale, there is continuity in asymmetric variation between humans and chimpanzees: similar brain asymmetries exist in both species, even though the human brain tends to be a little more asymmetric and more sensitive to developmental constraints.

More recently, tremendous efforts have been put into the identification of precise brain regions or neural structures that might be unique to humans and explain the emergence of language in our species. For example, Leroy et al. (2014) have proposed that the superior temporal sulcus (STS) critically differs in human and nonhuman species: human STS is deeper in the right hemisphere compared to the left, and this depth asymmetry is not found in chimps. Given that STS is central in the perisylvian language region, this particular landmark could be a promising "human-only" candidate. However, as the authors acknowledge, the link between this anatomical feature and the language ability in humans remains loose: MRI anatomical measurements made on various groups of human subjects show that the STS depth asymmetry is bigger in men than in women, it persists in children with impaired language development, and is unchanged in adults with reversed language lateralization (*situs inversus*). To drive the point home, an even

more recent study shows that the very same STS brain asymmetry exists in baboons (Meguerditchian et al., 2016). Further investigations on the relationship between variation in human brain anatomy (at a brain network scale) and variation in human language function are needed before one can make between-species comparisons and draw convincing conclusions about the role of precise human brain features in the evolution of human language.

The literature on human brain landmarks for language also raises a crucial question: what can gross anatomy tell us about fine-grained and complex interacting cognitive functions, in particular in a comparative perspective? Without a well-defined theoretical model of the anatomy-to-function relationship in each species that would make the comparison possible, the explanatory power of anatomy brain differences remains very limited. This would not be the case if we knew precisely the functional significance of those particular brain regions in their neural networks, in both species. For example, recent studies have shown that the homologue of Broca's area in chimps is involved in communicative behaviors (Tagliabue et al., 2008; 2011), however little is known about the type of computation it makes, or about the way it deals with a combination of communicative gestures and vocalizations, or about the way this region interacts with deeper structures involved in emotional vocalizations (Jürgens, 1979). Just like the linguistic/cognitive human-centered components, the anatomical landmarks of language -when strictly human-inspired (and functionally underspecified)- seem of very little explanatory value to understand language evolution in the human lineage, at least in the current state of our scientific knowledge.

The search for a unique (cognitive or anatomical) key factor at the origin of human language will inevitably lead to a stalemate. Brains, just like the cognitive function they host, are shaped by species-specific phylogenetical and developmental trajectories. They are complex systems and the differences between human and nonhuman brains and functions cannot be easily reduced to single "keys" components, especially when a strict human-centered definition of these components is applied.

1.2 An alternative ...

What could be the alternative to this narrow human-centered approach? In this chapter, we propose to step backward to enlarge our view of language evolution. Rather than focusing on the presence or absence of strictly human-inspired language features in nonhuman animals, we believe that an alternative is to examine the background of the language function, namely the inherited domain-general elements of “the machinery required to master human language” (Saffran and Thiessen, 2008), that might be used in non-human species for communicative and/or other purposes. Domain-general mechanisms correspond to learning devices that apply to a variety of different cognitive functions, as opposed to domain-specific mechanisms that are dedicated to specific cognitive functions. The opposition between domain-general *versus* domain-specific mechanisms was first proposed in a cognitive development framework where Skinner’s view (1957) was opposed to that of Chomsky (1959). In the present chapter, we use the notion of domain-general mechanism in a more evolutionary perspective. This idea we defend here is that the human language-device is mainly made of domain-general elements, some of which are shared with other species (either because they were present in a common ancestor or they result from convergence processes). These inherited cognitive components are very likely to take different forms in different species, as a function of the cognitive domain they are involved in, because different domains show different regularities and constraints that shape these components. However, we expect that the close comparison between human and nonhuman performance will reveal what aspects of those domain-general cognitive components are shared across species. The underlying hypothesis we uphold here is that complex and phylogenetically recent cognitive functions, including language, are probably the product of intense re-use and re-combination of subsets of inherited anatomical, cognitive, behavioral components (Anderson 2010). Phylogenetically close species might share some (but not all) of these components, as a support of communication and/or other cognitive functions. For example, the serial organization and structuration of elements that we find in the processing of syntax might as well serve the planning of complex motor sequences in humans (Koechlin and Jubault, 2006), in other primates or in birds, including the sequences of bird’s songs (Suzuki et al., 2016).

In the remaining part of this chapter, we will firstly describe some of the multilevel integrative processes that are critically involved in human language, which likely evolved from domain-general inherited functions. We will then consider and discuss main findings of the literature on comparative cognition regarding these domain general mechanisms. In this context, we will pay special attention to the mechanisms by which animals integrate stimulus information at various levels, from low level perceptual grouping mechanisms that lead to global percepts, to more cognitively complex processing mechanisms that make it possible to associate meanings to objects or categories for example. Importantly, this chapter is not aimed at making an exhaustive review of the literature regarding these cognitive processes. Our goal is to document potential *species* differences and similarities concerning these processes, and to illustrate these differences and similarities by a selection of suggestive findings. In a final section of this paper, we will discuss the potential impact of these results on our understanding of language evolution.

2. Integration processes in nonhuman animals

2.1 Integration in time and space

Language production and comprehension involve the processing of a continuous flow of information, and a temporal integration of the different linguistic elements. There is an obvious connection between sequential learning and language, because these two cognitive processes require the extraction and further handling of elements occurring in temporal sequences. However, because most behaviors have a temporal structure, the capacity to relate events in the time dimension might derive from a domain-general function that is involved in many different contexts, beyond language. We will consider below four different aspects of sequence processing, which are the ability to (1) remember sequence of events, and to process that sequence as temporally ordered, (2) to learn the transitional probabilities of items occurring in sequence, (3) to learn and process nonadjacent dependencies, and therefore to know that event A is followed by event B with intervening events between them, and (4) to gain information regarding the general structure of the sequence (i.e., syntax as concerns human language). For all these problems, we document below which aspects of sequence processing seem shared by animals and humans, and which aspects seem more restricted to humans.

2.1.1 *Serial list learning*

List learning corresponds to the capacity to both remember items in a list, but also to remember their temporal ordering. Nonhuman animals can learn and remember lists of items, (e.g. Terrace et al., 2003). Comparative investigations of serial list learning in animals, especially pigeons and monkeys, suggest that the memory of lists of items is not qualitatively different from what is found in humans. When taught that $A < B$, $B < C$, $C < D$, $D < E$, various nonhuman species have been able to properly order previously un-trained test pairs such as $C < E$. Pigeons and monkeys show serial position effects (Wright et al., 1985) sharing similarities with those of humans. Monkeys also show the symbolic distance effects found in humans: their latency to compare two items is shorter when the items are far apart in the list (e.g., AC vs CD ; Colombo and Frost, 2001). All these discoveries show that animals are capable of developing representations of series of items based on their ordinal position. Although the succession of words in a sentence is not strictly equivalent to a succession of unrelated items, serial list learning capacity is probably necessary in both cases. These results encourage the view that the underlying features of the learned representations are shared by human and nonhuman animals, suggesting quite old evolutionary mechanisms for item list learning.

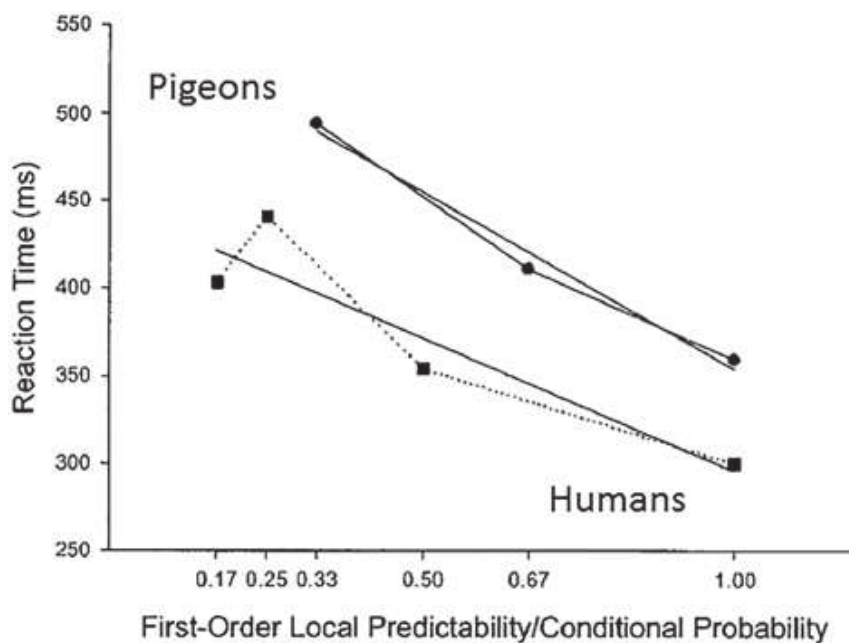
2.1.2 *Processing of adjacent dependencies and chunking*

Adjacent dependencies refer to a predictive relationship between one event and the event immediately following it in the sequence. Consider two three-item sequences, the sequences $A-B-C$ and $A-B-D$, which are presented an equal number of times. In this very small corpus composed of two sequences, A is always followed by B and B is followed half time by C and half time by D . The transitional probability between A and B is thus equal to 1, while the transitional probability between either B and C or B and D is equal to .5. Consideration of the transitional probabilities is one of the mechanisms promoting the learning of auditory and visual sequences in humans (e.g., Hunt and Aslin, 2001). It is one of the mechanisms by which children learn word boundaries and segment speech streams into words: in a sentence, the transitional probability between the last syllable of a word

and the first syllable of the following word is lower than the probability between two successive syllables within one word (Saffran et al., 1996).

Comparative studies have shown that a high transitional probability between two AB items facilitates the processing of the second B item in pigeons and monkeys. Froehlich et al. (2004), for instance, tested pigeons in a serial response time task requiring to peck a stimulus appearing sequentially at three possible locations and in a predefined order. The transitional probabilities between the stimulus locations were controlled in this task, and the authors report that response time to peck is a direct function of these probabilities. Thus, high transitional probabilities gave rise to short response times at the second location of the considered pair, while low transitional probabilities gave rise to longer response times. These authors directly compared their results on pigeons to Hunt and Aslin's study conducted on humans (2001), and report that although slower, pigeons processed information at roughly the same rate as humans, as reflected in similar overall regression slopes (Figure 1).

Figure 1: Average response times of pigeons and humans depending on transitional predictability of the items in a sequence. Results on pigeons are from Froehlich et al. (2004), those on humans are from Hunt and Aslin (2001). Also shown are linear regression lines for the two sets of data. Pigeons responded slower but processed information at roughly the same rate as reflected in similar overall slopes. Figure adapted from Froehlich et al. (2004).



In humans, the processing of transitional probabilities between adjacent elements comes with another type of processing known as chunking that consists in clustering information within sequences (e.g., the syllables in a word). Chunking in humans occurs during the memorization of verbal items, but chunking is not restricted to verbal material, and can be found as well in the visual domain (e.g., Orbán et al., 2008). Experimental evidence suggests that the capacity to organize sequences in chunks is also in the scope of numerous animal species (rats: Fountain, 1990; pigeons: Terrace, 1991; tamarins: Hauser et al., 2001; baboons: Minier et al., 2016). In a recent experiment involving a serial response time task, it was found for example that baboons organized 9-items sequences in three chunks of three items each, and that these chunks precisely included the items sharing the highest transitional probabilities (Minier, Rey and Fagot, 2016). Therefore, it seems that human and nonhuman animals are prone to statistical learning making use of transitional probabilities to both segment streams of (visual or auditory) information and organize the elements composing these streams into chunks.

2.1.3 Nonadjacent dependencies

Equally important for language is the capacity to detect and learn nonadjacent dependencies. Consider the following sequence structure A-X-B: A is followed by a variable item X, and item X is systematically followed by B. Given this structure, there is a nonadjacent transitional probability of 1 between A and B. Learning nonadjacent dependencies is important for language. For instance, a listener has to detect the relation between the subject and the verb in a sentence despite the presence of intervening words such as adverbs. This capacity can also be useful in very different nonlinguistic contexts, for example when we have to detect a systematic relation between two events separated in time (e.g., the ring of the doorbell signaling that someone is coming), and irrespective of the intervening other events. Experiments have shown that humans and animals can both process nonadjacent dependencies in temporal sequences of events, although nonadjacent probabilities are more difficult to detect and learn than adjacent dependencies (humans: Newport and Aslin, 2004; tamarins: Newport et al., 2004; rats: Fountain and Benson, 2006). Moreover, there are similarities

between species regarding the factors that affect the learning of nonadjacent dependencies (e.g., facilitatory effect of perceptual similarity of the nonadjacent elements in both humans and monkeys; humans: Creel et al., 2004; Gebhart et al., 2009; squirrel monkeys: Ravignani et al., 2013).

However, animals may have more difficulties when presented with complex sequences. Wilson et al. (2015) tested two monkeys and 33 humans using an auditory artificial grammar containing both adjacent and nonadjacent (long-distance) relationships. After an initial exposure to the sequences, the subjects from the two species were exposed to sequences containing violations of either the adjacent or both adjacent and nonadjacent relationships. Both species showed sensitivity to adjacent transitions, but only humans, and even roughly half of them, indicated significant sensitivity to nonadjacent dependencies. Wilson et al. (2015) concluded that in some conditions, nonadjacent probabilities are less salient in macaques than in humans. Although replications and extensions are required, this study suggests that, compared to monkeys, humans have a greater facility to deal with several dependencies of different types (i.e. both adjacent and nonadjacent) at the same time.

2.1.4 Learning of sequence structure

Learning the structure of a sequence requires the extraction of the relationships between the constitutive elements of that sequence. This kind of learning probably supports, among others, the encoding of grammatical and syntactic linguistic regularities (e.g., in most German sentences, the verb is at the end). However, such structural regularities also exist in many other (non-linguistic) domains, such as the motor domain. For instance, Byrne et al. (2001) reported that the preparation of food items requiring complex manipulations (thistle leaves) in wild gorillas follows a hierarchical sequential organization.

Marcus et al. (1999) have shown that very young 7-month-old infants can quickly learn that sequences of auditory stimuli follow an ABB or an ABA structure. The processing of such structures was also studied in nonhuman animals, but the results were quite inconclusive: zebra finches (Heijningen et al., 2012), rats (Toro and Trobalón, 2005) and even rhesus macaques (Procyk et al., 2000) do not seem to catch the difference between

the ABA vs ABB structures. One study compared zebra finches and humans using the same experimental procedure and stimuli (Chen et al., 2015): zebra finches did not learn these two types of structures, while humans learned them readily. There is to our knowledge only one study in which an animal species could successfully learn sequence patterns of the ABA/ABB type. This study from Spierings and Ten Cate (2016) compared two avian species, budgerigars and zebra finches, and obtained positive results in the former. From this set of experiments, we can conclude that learning the structure of sequences might very well be in the scope of some nonhuman species, but this ability is clearly not as developed as it is in humans.

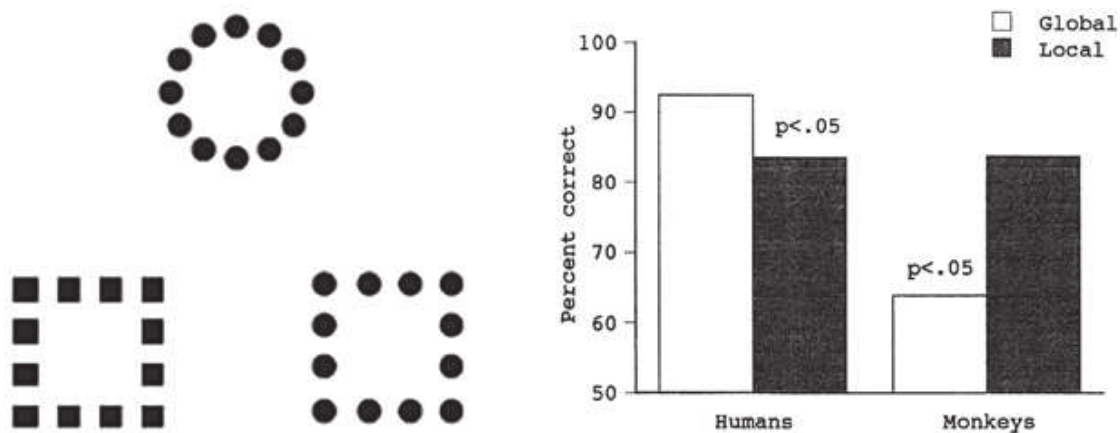
2.1.4 Integration in space

We documented above that human language function makes great use of temporal information, but spatial information is also crucially important for language. Infants at very young age learn that words and communicative gestures refer to entities in their immediate surrounding space. Words also make it possible to refer to objects that are spatially (or temporally) absent (the displacement feature proposed by the linguist Charles Hockett in 1960 to characterize human language). Finally, nonverbal forms of language, such as sign language or writing gestures strongly rely on an encoding of spatial information. No doubt that being capable to integrate and combine information in the spatial domain is also a domain general function important for a complex use of language.

In the late seventies, Navon (1977) has shown that human subjects tend to process the global shapes of visual object before they process their constitutive (local) details. This effect has been named the “global precedence effect” and is often considered to be an attentional phenomenon. Global precedence in humans was demonstrated in experimental research using large letters (global shape) made of smaller letters (local features). The degree to which animals perceive the global properties of the visual input in comparison to more featural ones has been an issue in animal cognition for some time. In our laboratory, we explored this effect in baboons (Deruelle and Fagot, 1998; Fagot and Deruelle 1997) using large shapes (square, circle, cross) made of smaller shapes (again square, circle, and cross) as stimuli. In our tests, the baboons were re-

quired to either match (Fagot and Deruelle, 1997) or identify (Deruelle and Fagot, 1998) these stimuli considering either their global or local structure. For comparative purposes, humans were also tested in the same experimental conditions.

Figure 2: Processing of the global / local stimulus structure in humans and baboons. Left: Illustration of the stimuli used with humans and baboons in Fagot and Deruelle (1997). This experiment required to match hierarchical stimuli considering their global or local structure (in this local trial, they have to match the circle made of circles, with the square made of circles, considering their common local features). Right: percentage correct obtained in humans and baboons in global and local trials. Humans showed an advantage to process the global structure of the stimuli, while baboons showed a local advantage. This local advantage in baboons is accounted for by a general difficulty to “group” the local elements of the hierarchical global/local stimuli into a global whole. Humans are much less sensitive than monkeys to the spatial distance separating the local elements (e.g., Deruelle and Fagot, 1998). Figure adapted from Fagot and Deruelle (1997).



These experiments revealed a striking species differences in global–local processing (Figure 2). Human participants exhibited the global advantage already found by Navon (1977), whereas baboons demonstrated their best performance and fastest response times in the local condition. Several experiments were conducted to understand the cause of this human-baboon difference, which suggested that the performance of the baboons strongly depended on the distance separating the local elements (Deruelle and Fagot, 1998): when the distance was enlarged, the strength of the local bias increased and this effect was amplified in baboons compared to humans. This

effect has been replicated many times in pigeons, capuchin monkeys and chimpanzees (for reviews, see Fagot and Barbet, 2006; Fagot and Parron, 2012).

At first glance, one might consider local precedence in animals as a purely perceptual/attentional phenomenon, but we propose that it is more than this. One of the main properties of language is that it can convey information about things that are not immediately present (spatially or temporally; Hockett, 1960). This displacement feature is crucial in the comparison between human language and other forms of primate communication. We have previously proposed (Fagot and Barbet, 2006) that a strong local bias limits the processing of the relation between and among objects. This effect is for instance demonstrated in Fagot and Parron (2010), showing that an increase in the separation of the distance between two bars of either identical or different colors limits the classification of this stimulus on a same/different relational basis. We suggest that a strong bias in favor of a local processing mode (in either the spatial or temporal domain) in nonhuman animals might place important constraints on their communicative systems. It reduces the possibility to make non (temporally or spatially) adjacent relations between or among the communicative signals, and between or among the communicative signals and the objects in the real world to which they refer, especially when they are far or absent.

2.2 Integration of stimulus dimensions and sensory modalities.

Animals, including humans, live in a rich world of information, and dealing with this complexity is probably critical for the survival of every species. The processing of this complexity may be achieved by a variety of cognitive mechanisms, which include, among others, the integration of the different perceptual dimensions into single entities (i.e., integration of stimulus dimensions), and the grouping of various exemplars of a given object into categories (e.g., object categorization; see a discussion of this issue below). Interestingly, these functions, which are of general adaptive value, are all critical for a multidimensional/multimodal system, such as language. For instance, speech comprehension requires that multiple prosodic (e.g., intonation, stress) and phonemic (e.g., voice onset time, place of articulation) dimensions present in the acoustic signal are processed and integrated.

Language comprehension is also achieved using a multimodal/multisensory mode of processing when various auditory and visual information are simultaneously integrated (e.g., for the mapping of the lip movements with the auditory signals).

2.2.1 *Combining multiple stimulus dimensions*

Discrimination tasks are often proposed to animals as a mean to assess their perceptual abilities. Typically, stimuli are presented to the subjects, and the subjects' behavioral responses to some perceptual dimensions or combinations of dimensions are reinforced, while responses to the non-relevant dimensions are not. Evidence suggests that animals can discriminate stimuli along various dimensions: color, shape, luminance, orientation of the visual objects, or the pitch of auditory stimuli. Animals can as well base their behavioral responses on combination of two or more stimulus dimensions. For instance, Cook (2001) showed that pigeons can learn to select the computer screen area where horizontal green lines are presented, while avoiding the screen areas showing non-green lines and green lines in a non-horizontal orientation.

In their review article, Lea and Wills (2008) comment on three main trends emerging from the literature on learned discrimination in nonhuman animals. The first one is that unidimensional discrimination is easier to learn than multidimensional discrimination based on combinations of features or conjunctions. Smith et al. (2012) have for instance trained monkeys to sort sine wave gratings depending on their orientations or both the orientation and spatial frequency considered in conjunction. Learning was much faster in the unidimensional than in the bi-dimensional test condition. Another example of this effect comes from research on conceptual discrimination by monkeys (e.g., D'Amato and van Sant, 1988) and the demonstration that discrimination performance relies strongly on an analysis of features, such as color, rather than on configurations of features. Lea and Wills' (2008) second conclusion is that when the stimuli are made of multiple relevant dimensions, nonhuman animals express a tendency to focus their attention on one dimension only, mostly when this dimension has sufficient discriminative values. There are also multiple examples of this trend in the literature. In our laboratory, we found that baboons discriminated

computerized human faces considering exclusively the contour of the face (Martin-Malivel and Fagot, 2001) or pixel luminance information (Martin-Malivel et al., 2006), instead of the multiple levels of information (e.g., configural information, identity, etc.) that the facial stimuli may provide, and that humans process. Finally, Lea and Wills' (2008) report than even with ingenious experimental designs, attempts to force nonhuman animals to process multiple aspects of the stimuli mostly lead to failures. This can be nicely illustrated by Dépy et al. (1997). The baboons in this study were initially trained to discriminate between two categories of stimuli defined by the possession of any combination of two out of three possible binary features. Baboons could sort these two classes of stimuli to a good accuracy level, albeit after a long training process of several thousands of trials, but remained unable to take the three discriminative features into consideration to achieve this performance, two of the three features taking a leading role in the task.

Wang et al. (2015) recently recorded the brain activity (IRMf) of rhesus monkeys and humans in two test phases. In the first phase, subjects from the two species perceived passively sequences of four tones, the last one being either of a lower or higher pitch than the first three. After this habituation procedure, the same subjects perceived sequences violating the general structure used during the habituation phase. Thus, some test sequences contained a number of tones different from the habituation sequences (number deviant), some other sequences contained four tone units with identical pitch (sequence deviant), and a last set of sequences differed from the habituation sequences regarding both the number of items and pitch. In both species, homologous brain areas were particularly responsive to violations in number (intraparietal and dorso premotor areas), and sequence (ventral prefrontal and basal ganglia), but humans were the only primates showing a joint sensitivity to both factors in the perisylvian language region (bilateral inferior frontal and superior temporal gyri). One limitation of this study is that an absence of evidence in baboons is not the evidence of an absence. Although this study does not address directly the relationship between brain and behavior, its results suggest that the perisylvian region is involved in humans only in the integration of various stimulus dimensions contained in auditory sequences of stimuli.

2.2.2 *Combining multiple sensory modalities*

The question of multimodal integration warrants a discussion in this section, due to its relevance to the origin of language. Evidence suggests that the ability to integrate information across sensory modalities is not at all restricted to humans. Cross-modal integration was demonstrated for instance using a task requiring the processing of multi-sensory stimuli (Lanz et al., 2013), cross-modal interference tasks in baboons (Martin-Malivel and Fagot, 2001), and cross-modal matching tasks in chimpanzees (e.g., Davenport and Rogers, 1970). Unfortunately, we are aware of only one study in which the performance of humans and nonhuman animals were directly compared in cross-modal tasks using the same stimulus material. This experiment from Fagot et al. (2000) requested the subjects to categorize pictures of humans and baboons in one condition, and human and baboon vocalizations in another one. In this experiment, the subjects of the two species perceived a prime prior to the presentation of the stimulus to be categorized. Depending on the condition, the prime could be a picture or a vocalization of baboons or humans and three conditions were tested: intra-modal visual-visual priming, intermodal auditory-visual priming, and intramodal visual-auditory priming. Three subjects out of four in each species demonstrated intra-modal priming. Inter-modal priming was demonstrated in the three out of four human subjects in the auditory-visual condition, and all four in the visual-auditory condition, but it was only found in one baboon out of four in each intermodal condition, suggesting that inter-modal integration is more difficult in baboons than in humans. Given the small number of subjects involved in this study, a replication is warranted before drawing any firm conclusion on the evolution of intra-modal integration.

2.3 **Categorization and conceptual integration**

A central aspect of human cognition is our ability to form categories of various kinds of objects or mental entities. Categorization implies that the exemplars of each category are grouped into classes considering physical or more abstract properties. Categorization is a domain-general ability that is fundamental for a variety of more specialized functions (e.g., inference or decision making), including the language function considered in this book.

Historically, Herrnstein and Loveland (1964) were the first to demonstrate categorical abilities in nonhuman animals. They showed that pigeons could sort pictures of humans and pictures devoid of humans in two open-ended categories. Since this study, numerous papers have confirmed that various animal species can efficiently categorize stimuli considering low level perceptual stimulus dimensions, such as the pitch for auditory stimuli, and the color, shape, size, motion, orientation, luminance for visual stimuli (see for instance Berg and Grace, 2011, where pigeons were trained to categorize sine-wave disks considering their spatial frequency and orientation). However, the ability of nonhuman animals to apply categorical processes to more abstract – human-like – stimulus dimensions remains a matter of debate.

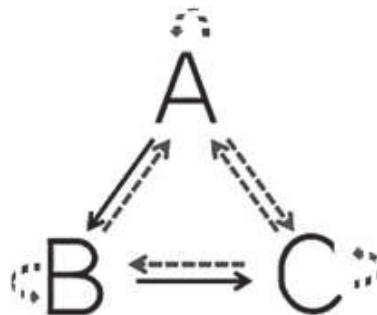
2.3.1 Equivalence classes: grouping arbitrary items within the same category

An important aspect of language is its arbitrariness. Arbitrariness corresponds to the fact that nothing in the physical form (acoustic properties) of most words refers to the objects they designate. For example, the word “car” does not “sound” like the vehicle it refers to. Therefore, words can refer to things in the real world, and things can refer to words, although there is no natural or necessary connection between them. In that case, words and objects are linked by a relation of equivalence, and the many exemplars of a given category of object (e.g., many different tables that vary in shape and color) can be categorized under a unique word label. The ability to form arbitrary connections between words and objects during ontogeny probably comes from the many co-occurrences of words-objects pairs that infants encounter during development. We presume that this ability relies on a more domain-general capacity to make associative (arbitrary) connections between items, or categories of items.

Experiments on “stimulus equivalence” have directly addressed the capacity of human and nonhuman species to group arbitrary items into categories, on the basis of their associative history (Sidman and Tailby, 1982). The prototypical design of stimulus equivalence experiments is shown in Figure 3. In stimulus equivalence experiments, the subject first learns a network of associations (shown in black in Figure 3) with repeated expo-

asures to these associations. Then, probe trials (shown by the red arrows in Figure 3) test if the subject (1) can associate each stimulus to itself (e.g., associate A to A, reflexivity relation), (2) can revert the trained relations (e.g., associate B to A, symmetry), and can associate the stimuli that have a common associate (associate A to C, transitivity). According to Sidman and Tailby (1982), stimulus equivalence is fully shown if the subject demonstrates, without further training, the relations of reflexivity, symmetry and transitivity in post training trials. One may easily imagine the serious limits of a cognitive system that would fail in this task. The formation of equivalence classes has the power to permit to use stimuli interchangeably, and is probably the corner stone of complex symbolic thought.

Figure 3: Typical paradigm for experiments on stimulus equivalence. In this experimental design, the black arrows illustrate the trained associations. The dotted arrows illustrate the untrained associations that emerge in humans after an initial training phase (e.g., Sidman and Tailby, 1982).



Equivalence relations emerge early in human infancy (23-month old, Lipkens et al., 1993). By contrast, the formation of equivalence classes seems especially difficult for nonhuman animals such as pigeons or monkeys, and among the three relations sustaining stimulus equivalence, the relation of symmetry seems the most difficult one to acquire. Lionello-DeNolf (2009) reviewed a total of 24 articles on symmetry testing, and found that the vast majority of these articles reported negative results. Moreover, the handful of articles reporting more positive findings all used a very small number of subjects who received special training procedures, such as long training combining symmetry and reflexivity trials (Frank and Wasserman, 2005) or forms of symbolic training (Pepperberg, 2006).

A recent paper from our research group nicely illustrates that difficulty (Medam et al., 2016). We trained baboons to associate pictures of bears and

pictures of cars to two different shapes which served as category labels. The baboons demonstrated category learning in this task, but failed to respond correctly when asked to reverse the trained relation and to associate the category labels to the pictures of cars and bears. The immediate conclusion from this study and others (see Lionello-DeNolf, 2009) is that the preferred mode of processing in nonhuman animals is not favorable to the spontaneous emergence of symmetrical relations, which is fundamental for the emergence of equivalences classes. Arbitrariness, as defined by linguists, requires arbitrary, bidirectional associations between the words and their referent (e.g., the written “CAT” refers to the cat animal, as the cat animal refers to the “CAT”). Evidence suggests that nonhuman animals too can form arbitrary associations between various items with no logical connections between them, but they apparently have great difficulties to process these associations as bi-directional.

2.3.2 *First order relations*

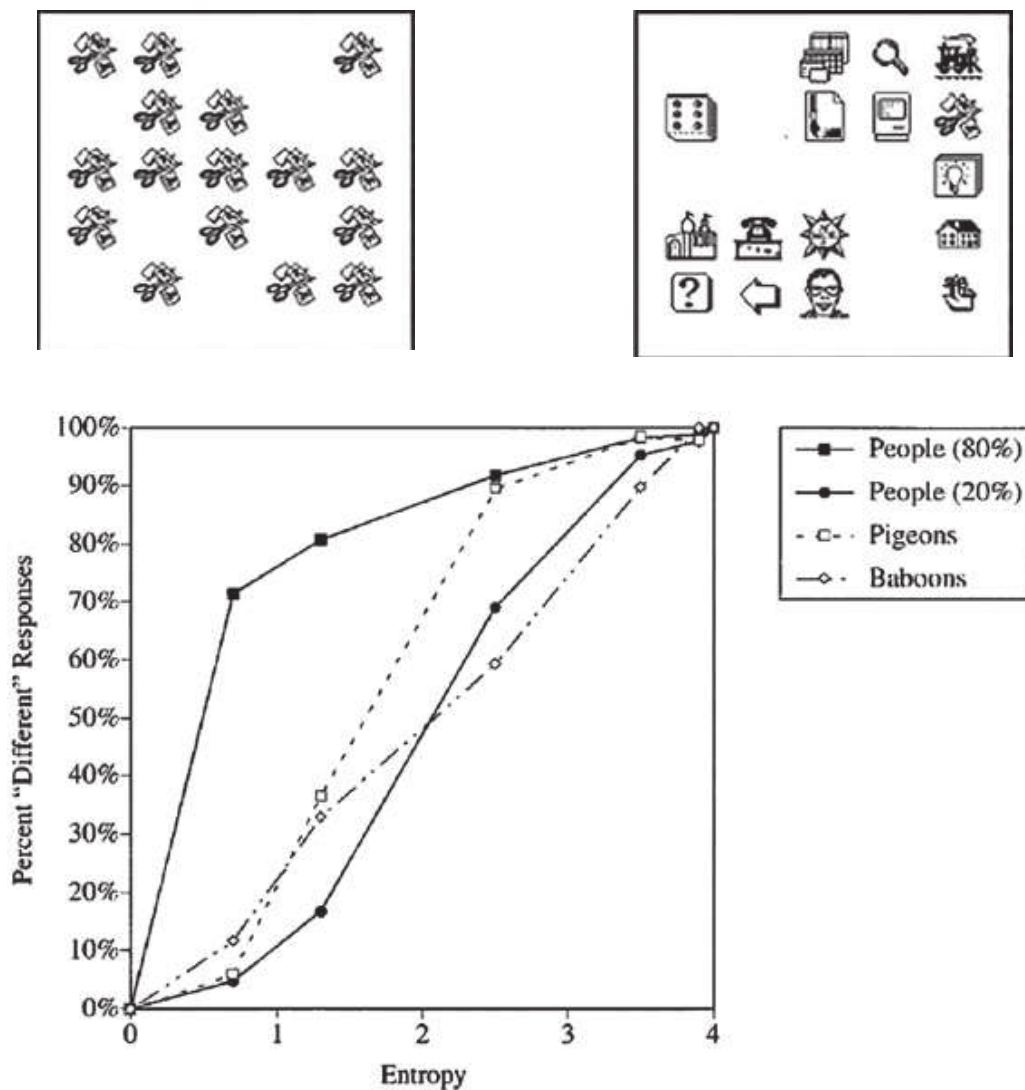
Nonhuman animals can master a broad range of discrimination tasks, and some of them involve the processing of first-order relations. First-order relations refer to spatial or more abstract relations among objects, such as the fact that an object is above or below another one, or that two objects have the same functions. Particularly important for our linguistic system is the first-order relation of sameness/differentness. Our language makes great use of categories, and the abstract concept of sameness is essential for the development of verbal categories. The concept of sameness can provide the basis for the most complex cognitive operations, such as the conservation of volumes or areas, or analogical reasoning (see below). Without this concept, we would be unable to understand sentences such as “This is a cat!”, and to get from this sentence the idea that the animal we see belongs to the cat category. We would as well be unable to understand sentences such as “It is warm again”, suggesting a similarity between the current and past weather. Comparative psychologists have shown that a variety of nonhuman animal species, such as the chimpanzee (Premack, 1983), the baboon (Wasserman et al., 2001) and the rat (Wasserman et al., 2012), succeed in learning same-different relational tasks. However, the nature of the mechanisms supporting this competence in nonhuman animals, and their similarity with humans, remain unclear.

Wright and Katz (2006) asked the following question: how much training nonhuman animals need to form the same/different concept? To answer this question, they tested pigeons, capuchin monkeys and rhesus monkeys with the same test design. Animals from the three species saw two pictures in succession on a touch screen, and two kinds of trials were distinguished. In the “same” trials, the second picture was identical to the first one, while it was different from the first one in the “different” trials. In both kinds of trials, a white key was always displayed on the right of the second picture. When the first and second pictures were identical (identical trials), then the subject was asked to touch/peck the second picture to obtain a reward. If the second picture was different from the first one (different trials), then a touch/peck at the white response key was considered correct. The monkeys needed much less items (about 32) to develop the concept of sameness, than did the pigeons (256). The number of trials children would need in this task is not known, but studies have shown that children can categorize cats as different from dogs with only 12 training exemplars (Quinn et al., 1993), and by 10 months of age, they can form categories with only 7 or 8 training exemplars (Younger and Cohen, 1986). The data therefore suggest an evolutionary trend in this ability: humans would require exposure to a smaller number of items than the other animals, to form categories and develop same/different concepts.

The concept of sameness can be applied to a broad range of attributes, from the most perceptual to the most abstract ones, and another interesting issue in the comparative literature is to know if animals use the same kind of information as humans, when solving similar tasks requiring an abstract concept of sameness/differentness. This kind of questions has been addressed extensively by Wasserman and collaborators (see review in Wasserman et al., in press; Wasserman et al., 2004). These authors trained pigeons, baboons and humans to categorize displays resembling those of the top of Figure 4. They consisted in arrays of 16 icons which were either all same (same relation) or all different (different relation). After they received this category training, the subjects were tested with arrays containing mixtures of icons, in which some icons were duplicated a number of times in the array. The authors reasoned that if the subjects have formed the concept of sameness, then they should classify the arrays containing at least one item different from the others as “different arrays”, irrespective of the fact that some icons are repeated. Figure 4 illustrates the most substantial findings of this set of experiments. This

figure indicates the percentage of different responses with the mixtures, as a function of the entropy of the stimulus. Entropy in this experiment should be understood as a quantification of the perceptual variability of the array: the all-same arrays have an entropy of 0, and the all different-arrays have the highest possible entropy value of 4 (mixtures have intermediate entropy values).

Figure 4: Use of perceptual (entropy) cues by pigeons, monkeys and humans, in a same/different discrimination task. The top panel shows the kind of displays employed in Wasserman and collaborators' experiments (Wasserman et al., 2004, left: same array, right: different array). The bottom figure shows that the same-different response of monkeys and pigeons is controlled in this task by the entropy of the arrays while 80% of the humans humans used more abstract cues in this task. Figure adapted from Wasserman et al. (2004).



The bottom part of Figure 4 shows that the behaviour of the baboons and the pigeons was controlled by the entropy of the arrays, which is a perceptual cue. Response from a subset of humans (about 20% of the group) was also largely controlled by the entropy of the arrays, but this constraint was released in most of the subjects (80%) who treated the arrays containing at least one item different from the others as illustrations of the “different” concept. It can therefore be concluded that humans expressed more abstract judgments than pigeons and baboons in this task. We will not present in this chapter the full series of experiments using this kind of stimuli with animals (for a recent review, see Wasserman et al., in press), but the reader should be aware that clear demonstrations also exist that monkeys can also base their same/different responses on abstract cues independently of the entropy of the stimuli (Flemmig et al., 2013). Nevertheless, although several animal species seem capable of abstract same/different judgements, humans, more readily than other animals, apply qualitative, rule-based frameworks on the Same-Different discrimination task.

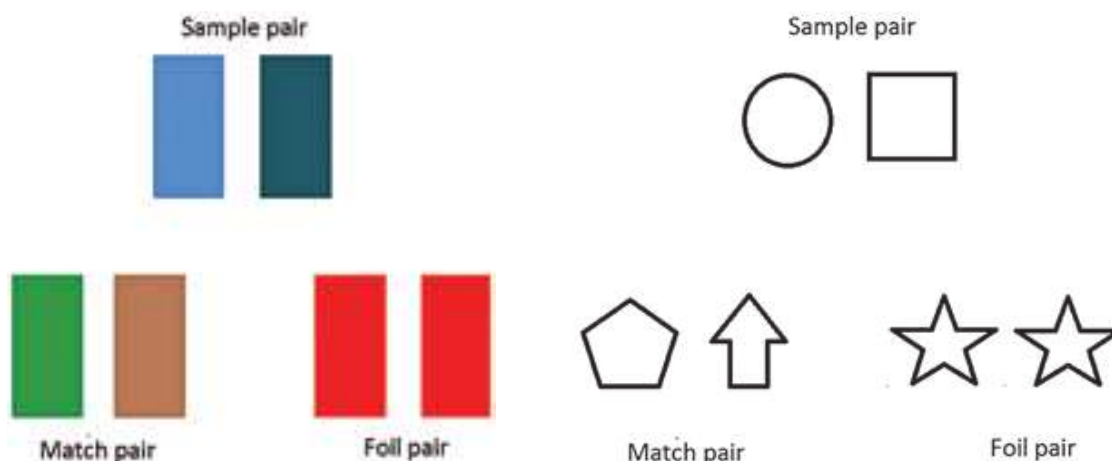
2.3.3 *Analogical (second-order) relational processing*

Our linguistic systems make great use of analogies, and our capacity to produce and understand analogies is considered by many as the “the Fuel and Fire of Thinking” (Hofstater and Sander, 2013). Developmental studies have shown that analogical reasoning is facilitated in children by the capacity to represent abstract relations in symbolic terms via linguistic labels (Christie and Gentner, 2014).

Most research on analogical reasoning in animals has used the Relational Matching-to-Sample task (RMTS: e.g., Fagot and Thompson, 2011) illustrated in Figure 5. In this task, the subject first perceives one pair of objects which are either identical or different. Two comparison pairs are then presented, and the subject must indicate the stimulus pair exemplifying the same (same or different) relation as the sample pair. In other words, the task can be conceptualized as “if AA then BB, and if AB then CD”. Researchers in the domain of comparative cognition have tested several animal species, including pigeons, monkeys and apes using the RMTS task (for a review, see Wasserman, Castro and Fagot, in press). Most of these attempts failed (Thompson and Oden, 2000), but a handful of studies also provide more

positive results, in particular in tests involving thousands of training trials (e.g., Fagot and Thompson, 2011).

Figure 5: Illustration of the relational matching task used, in baboons, with color (left part of the figure) and shape (right) stimuli.



In our laboratory, we could demonstrate that the baboons can successfully solve the RMTS task with pairs of color patches as stimuli (Fagot and Parron, 2010, see Figure 5). In a different study (Fagot and Thompson, 2011), we could further demonstrate that baboons can also solve this RMTS task considering the shape of the items (Figure 5). Again, this cognitive feat also required an extensive training period (from 17 to 30 000 trials per subject). In both studies, the different generalization tests confirmed the real abstract nature of the processes at work in these two tasks. For instance, the baboons could continue to solve the task with a high level of performance when we used novel colors (Fagot and Parron, 2010), and novel shapes (Fagot and Thompson, 2011) as stimuli. However, although cognitive flexibility is suggested by these findings, the data also suggest limits in this processing. In Fagot and Parron's (2010) study, color cues were in fact proposed to 6 baboons, and 4 of these 6 subjects eventually learned the task. In Fagot and Thompson's (2011) study, the same task was given one year later to a larger group of subjects, including the 6 already tested in Fagot and Parron (2010). Six out of 29 baboons learned the RMTS task with shapes, but importantly, none of the baboons who had initially learned the task with colors could also learn it with shapes. In other words, learning the RMTS task with color

cues did not help at all the subjects to learn. Generalization across domains is central to analogical reasoning in humans, and is probably what makes human reasoning so flexible. Generalization across domains allows us, for instance, to understand the meaning of a sentence like “atoms are like tiny solar systems” or “life is a gift, a chocolate box”. Baboons – and probably other nonhuman primates as well – are quite flexible to process items and their relations within given domains, those they have been trained with (e.g., color), but are clearly not as skilled as humans to generalize across domains. This, we believe, is another factor that may greatly affect nonhuman animal’s potential for developing elaborated forms of language.

3. Summary and Conclusions

The main goal of the chapter was to examine the origin of human language from the standpoint of comparative psychology. Language in its various forms (e.g., gesture, writing, speech) is a multi-level integrative process that requires, at the perceptual stage, the segmentation and grouping of perceptual information to extract the general meanings of the communicative signals. We have argued above that many of the integrative processes involved in the language function are in fact domain-general processes that can also be found in non-linguistic functions.

Considering that language uses a multitude of domain-general functions, examination of these functions in animals and especially in nonhuman primates should provide important information on the cognitive background that made it possible for language to emerge in our evolutionary history. Following this reasoning, we comparatively examined in this chapter a number of domain-general cognitive functions, which imply various forms of integration of perceptual/conceptual information. Among the considered functions, we examined the ability of animals to integrate information in time and space, to combine stimulus dimensions, to group objects into categories, and to develop conceptual/relational processes (first- and second-order concepts). Of course, this list is not exhaustive but we believe that it represents a significant selection of basic cognitive domain-general processes that serve language perception and comprehension.

The present overview of the literature allowed us to reach two main conclusions. First, we have identified clear-cut demonstrations that non-

human animals are capable of grouping information in time and space, can combine stimulus dimensions, and can form categories at different levels of abstractness. Thus, we suggest that these functions are shared by humans and nonhumans, at least to some extent, and that they are not language specific (i.e., they have a long phylogenetic history). Second, the literature also reveals important differences in performance among nonhuman animals, and between nonhuman animals and humans. For example, evidence suggest that, compared to humans, in nonhuman animals the integration of information in time and space is more “local”, and the capacity to integrate the information on a larger scale is more restricted. When we come to consider how animals integrate various stimulus dimensions, experimental evidence suggest that they tend to focus on some particular physical dimensions of the stimuli more than humans do, and they hardly combine information from different stimulus dimensions. When we come to consider more general categorization processes, it appears that nonhuman primates form categories but their categories seem to be more strongly tied to the perceptual input than those of humans, and abstract processes, when they emerge, need many more trials to develop, and/or do not generalize to untrained dimensions as readily as for humans. In other words, for most of the integrative functions we have considered in this chapter, nonhuman animals show behaviors that differ at least quantitatively, if not qualitatively from human behavior, and we propose that these differences might be the bottlenecks for the evolution of language. Obviously, evolution has no direction and mastering some human-like language is not and has never been an issue in nonhuman animals. However, it might be that a particular pattern of development of these domain-general functions was a prerequisite to the emergence of human language, and that favorable conjunctions occurred only once in the phylogeny of the primate group, giving rise to the human language. The idea that some domain-general nonlinguistic functions form the bottlenecks of language evolution will be further developed below.

To account for the evolution of language, many theorists have focused their attention on language specific functions, which were considered as key factors for the evolution of human language. For example, many traditional theories claimed that only humans have a low larynx (which was disconfirmed since), or that only humans have the ability to understand

and produce recursive structures allowing an infinite variability in language production. Here, we do not want to discount these explanations, but we think that they do not take the problem at its roots. At this point in the scientific endeavor, we think that it is now necessary to step back a bit to enlarge our view of the problem. Doing so, we can imagine two different scenarios on the origin of language.

The "language-first" scenario would be that our primate/prehominid ancestors had rather limited cognitive resources and it was the appearance of language that boosted their general cognitive capacities. Although there is no doubt that language is a booster for cognitive functions, we think that this scenario does not hold. To illustrate our rebuttal of this theory, imagine an animal cognitive system with a limited working memory span of $N=1$ item. How could a language system develop with such limited memory resources? A clear expansion of this memory system would be required before the animal would have a chance to develop a proto-language system. This very extreme example shows that language evolution requires a cognitive background to support it, and it is only when such a cognitive background has evolved – potentially as a response to ecological pressures – that a form of proto-language can have a chance to emerge.

The second possible scenario, called "prerequisites-first", is that language only emerged in our evolutionary history once critical domain-general functions had gained in cognitive power in our animal/pre-hominid ancestors. This scenario raises one major question: which domain general function(s), or combination of functions, must have evolved at first for the emergence of language? We have no clear answer to this question but can provide several lines of thinking on this issue. First, we note that studies focusing on the so called low level perceptual mechanisms showed important differences between the platyrrhines and catarrhines species, but very little differences are observed in perceptual functions in monkeys, apes and human species (see for instance Fobes and King, 1982 for a review of visual perception). This mere fact suggests that the evolution of these perceptual functions is probably not the factor that made the difference and triggered the evolution of speech in humans. Secondly, comparative experiments suggest more pervasive differences between humans and the other primates in two domains at least. The first domain of importance is the domain of working memory. Working memory in humans seems to depart from that of other animal

species in several important aspects, for instance regarding the ability to process large amount of information in parallel (e.g., Fagot and De Lillo, 2011), or to process long-distance dependencies (e.g., Wilson et al., 2015). Although not discussed in the context of this chapter, working memory in humans may also qualitatively depart from that of the other animals in its use of a phonological loop facilitating memorization in the short term. The second domain for which strong differences emerge between humans and the other animals is the domain of attention. Nonhuman animals seem to focus on single stimulus dimensions more than humans do, and tend to have a more local mode of processing of the perceptual input than humans. There remains a debate on whether the increase in cognitive functions followed the phylogenetic order, from the remote prosimian species to the ape species phylogenetically closest to humans (e.g., Reader et al., 2011), or whether variations in cognitive power among the different primate species have occurred at multiple times in the course of evolution, in independent unrelated primates groups, for instance under the pressure of social factors such as the complexity of the social network (Dunbar, 1998). Discussing these hypotheses is out of the scope of the current chapter. However, whatever the source of this increment in cognitive power is, we propose that the language ability appeared in the evolution of primates at a point in time where domain-general cognitive capacities, especially those pertaining to attention and working memory converged and were sufficiently developed to permit its evolution. This idea is in line with recent usage-based theories suggesting that that language could be acquired in humans by means of domain-general – evolutionary old – processes (Bybee, 2010; Tomasello, 2005).

From a more practical standpoint, we conclude from this chapter that the comparative investigation of non-linguistic, domain general functions should be considered of central importance in the debate on the evolution of language. Unfortunately, real comparative studies, in which humans and other species are tested on the same problems using the same tasks, are relatively rare in this literature, and most of them only concern a very limited number of species. Such studies will become mandatory to further test the hypothesis that the expansion of domain-general functions in nonhuman primates served as a basis for the evolution of language.

Acknowledgements

This project was supported by ANR-16-CONV-0002 (ILCB), ANR-11-LA-BX-0036 (BLRI) and ANR-11-IDEX-0001-02 (A*MIDEX).

References

- Abe, K., and Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, 14, 1067–1074.
- Anderson M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behavioural Brain Science*, 33, 245–266.
- Berg, M.E., and Grace, R.C. (2011). Categorization of multidimensional stimuli by pigeons. *Journal of Experimental Analysis of Behavior*. 95, 305–326.
- Boë, L.J., Berthommier, F., Legou, T., Captier, G., Kemp C., Sawallis, T.R., Becker, Y., Rey, A., and Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PLoS ONE*, 12, e0169321.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Byrne, R.W., Corp, N., and Byrne, J.M. (2001). Estimating the complexity of animal behaviour: How mountain gorillas eat thistles. *Behaviour*, 138, 525–557.
- Cantalupo, C., and Hopkins, W. D. (2001). Asymmetrical Broca’s area in great apes. *Nature*, 414, 505.
- Cantalupo, C., Pilcher, D., and Hopkins, W. D. (2003). Are planum temporale and sylvian fissure asymmetries directly related? A MRI study in great apes. *Neuropsychologia*, 41, 1975–1981.
- Chen, J., van Rossum, D., and ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition*, 18, 151–164.
- Chomsky, N. (1959). A review of B.F. Skinner’s “Verbal Behavior.” *Language*, 35, 26–58.
- Christie, S., and Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science*, 38, 383–397.

- Colombo, M., and Frost, N. (2001). Representation of serial order in humans: A comparison to the findings with monkeys (*Cebus apella*). *Psychonomic Bulletin and Review*, 8, 262–269.
- Cook R. G. (2001). Hierarchical stimulus processing in pigeons. In R. G. Cook (Ed.), *Avian visual cognition* [On-line]. Available: www.pigeon.psy.tufts.edu/avc/cook/
- Creel, S. C., Newport, E. L., and Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1119–1130.
- D'Amato, M. R., and Van Sant, P. (1988). The person concept in monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 43–55.
- Davenport, R.K., and Rogers, C.M. (1970). Intermodal equivalence of stimuli in apes. *Science*, 1968, 279–280.
- Dépy, D., Fagot, J. and Vauclair, J. (1997). Categorization of three-dimensional stimuli by humans and baboons (*Papio papio*): Search for prototype effects. *Behavioral Processes*, 37, 299–306.
- Deruelle, C., and Fagot, J. (1998). Visual search for global/local stimulus features in humans and baboons. *Psychonomic Bulletin and Review*, 5, 476–481.
- Dunbar, R.I.M (1998). The social brain hypothesis. *Evolutionary Anthropology*, 6, 178–190.
- Fagot, J., and Barbet, I. (2006). Grouping and segmentation of visual objects by baboons (*Papio papio*) and humans (*Homo sapiens*). In E. A. Wasserman and T. R. Zentall (Eds.), *Comparative cognition: experimental explorations of animal intelligence* (pp. 11–24). New York: Oxford University Press.
- Fagot, J., and De Lillo, C. (2011). A comparative study of working memory: immediate serial spatial recall in baboons (*Papio papio*) and humans. *Neuropsychologia*, 49, 3870–3880.
- Fagot, J., and Deruelle, C. (1997). Processing of global and local visual information and hemispheric specialization in humans (*Homo sapiens*) and baboons (*Papio papio*). *Journal of Experimental Psychology: Human Perception and Performance*, 23, 429–442.

- Fagot, J., Martin-Malivel, J., and Dépy, D. (2000). What is the evidence for an equivalence between objects and pictures in birds and nonhuman primates? In J. Fagot (Ed.), *Picture Perception in Animals* (pp 295–320). East Sussex: Psychology Press.
- Fagot, J. and Parron, C. (2010). Relational matching in baboons (*Papio papio*) with reduced grouping requirements. *Journal of Experimental Psychology: Animal Behavior Processes*, 36, 184–193.
- Fagot, J., and Parron, C. (2012). Visual cognition in nonhuman primates. In O. Lazaverra, T. Shimizu and E.A. Wasserman (Eds.), *How animals see the world* (pp. 371–389). Oxford: Psychology Press.
- Fagot, J., and Thompson, R. K. R. (2011). Generalized relational matching by Guinea baboons (*Papio papio*) in two by two-item analogy problems. *Psychological Science*, 22, 1304–1309.
- Fitch, W.T. (2010) *The evolution of language*. Cambridge University Press.
- Fitch, WT, de Boer, B., Mathur, N., and Ghazanfar, A.A. (2016). Monkey vocal tracts are speech-ready. *Science advances*, 2:e1600723.
- Fitch, W. T., and Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303 (5656), 377–380.
- Flemming, T. M., Thompson, R. K. R., and Fagot, J. (2013). Baboons, like humans, solve analogy by categorical abstraction of relations. *Animal Cognition*, 16, 519–524.
- Fobes, J.L., and King, J.E. (1982). Vision: The dominant primate modality. In J.L. Fobes and J.E. King (Eds.), *Primate behavior* (pp. 219–243). New York: Academic Press.
- Fountain, S. B. (1990). Rule abstraction, item memory, and chunking in rat serial-pattern tracking. *Journal of Experimental Psychology. Animal Behavior Processes*, 16, 96–105.
- Fountain, S. B., and Benson Jr., D. M. (2006). Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning. *Learning and Motivation*, 37, 95–112.
- Frank, A.J. and Wasserman, E.A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *Journal of Experimental Analysis of Behavior*, 84, 147–165.
- Froehlich, A. L., Herbranson, W. T., Loper, J. D., Wood, D. M., and Shimp, C. P. (2004). Anticipating by pigeons depends on local statistical infor-

- mation in a serial response time task. *Journal of Experimental Psychology: General*, 133, 31–45.
- Gebhart, A. L., Newport, E. L., and Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin and Review*, 16, 486–490.
- Gentner, T. Q., Fenn, K. M., Margoliash, D., and Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440, 1204–1207.
- Geschwind, N., and Levitsky, W. (1968). Human brain: Left-right asymmetries in temporal speech region. *Science*, 161, 186–187.
- Gomez-Robles, A., Hopkins, W.D., and Sherwood, C.C. (2013). Increased morphological asymmetry, evolvability and plasticity in human brain evolution. *Proceedings of the Royal Society: B*, 280, 20130575.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how does it evolve? *Science*, 298, 1569–1579.
- Hauser, M. D., Newport, E. L., and Aslin, R. N. (2001). Segmentation of the speech stream in a nonhuman primate: statistical learning in cotton-top tamarins. *Cognition*, 78, B53–B64.
- Hayes, C. (1951), *The Ape in Our House*. New York: Harper.
- Heijningen, C. A. A. van, Chen, J., Laatum, I. van, Hulst, B. van der, and Cate, C. ten. (2012). Rule learning by zebra finches in an artificial grammar learning task: which rule? *Animal Cognition*, 16, 165–175.
- Herrnstein, R.J., and Loveland, D.H. (1964). Complex visual concept in the pigeon. *Science*, 23, 549–551.
- Hockett, C. (1960). The origin of speech. *Scientific American*, 203, 88–111. Reprinted in Wang, William S-Y. (1982) *Human Communication: language and its psychobiological bases*, Scientific American pp. 4–12.
- Hofstater, D., and Sanders, E. (2013). *Surfaces and essences: analogy as the fuel and fire of thinking*. New-York: Basic Books.
- Hunt, R. H., and Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, 130, 658–680.

- Jürgens, U. (1979). Neural control of vocalization in nonhuman primates. In H.D. Steklis and M. J. Raleigh (Eds.), *Neurobiology of social communication in primates* (pp. 199–206). New-York: Academic Press.
- Koechlin, E, and Jubault, T (2006). Broca's area and the hierarchical organization of human behavior. *Neuron*, 50, 963–74.
- Lanz, F., Moret, V., Rouiller, E.M., and Loquet, G. (2013). multisensory integration in nonhuman primates during a sensory-motor task. *Frontiers in Human Neuroscience*, 7,799.
- Lea, S. G. E., and Wills, A. J. (2008). Use of multiple dimensions in learned discriminations. *Comparative Cognition and Behavior Reviews*, 3, 115–133.
- Leroy, F. et al. (2014). New human-specific brain landmark: The depth asymmetry of superior temporal sulcus. *PNAS*, 112, 1208–1213.
- Lieberman, P. (1968). Primate vocalizations and human linguistic ability. *Journal of the Acoustic Society of America*, 44, 1574–84.
- Lieberman, P. (1975). *On the Origins of Language: an Introduction to the Evolution of Human Speech*. New York: Macmillan.
- Lionello-DeNolf, K.M. (2009). The search for symmetry: 25 years in review. *Learning and Behavior*, 37, 188–203.
- Lipkens, G., Hayes, S. C., and Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, 56, 201–239.
- Marcus, G. F., Vijayan, S., Rao, S. B., and Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Malivel, M. and Fagot, J. (2001). Cross-modal interference and conceptual categorization in baboons. *Behavioural Brain Research*, 122, 209–213.
- Martin-Malivel, J. and Fagot, J. (2001). Perception of pictorial human faces by baboons (*Papio papio*): effects of stimulus orientation on discrimination performance. *Animal Learning and Behaviour*, 29, 10–20.
- Martin-Malivel, J. Mangini, M., Fagot, J. and Biederman, I (2006). Do humans and baboons use the same information when categorizing human and baboon faces ? *Psychological Science*, 17, 599–607.
- Medam, T., Marzouki, Y., Montant, M., and Fagot, J. (2016). Categorization does not promote symmetry in Guinea baboons (*Papio papio*). *Animal Cognition*, 19, 987–99.

- Meguerditchian, A., Marie, D., Love, S. A., Margiotoudi, K., Bertello, A., Lacoste, R., Roth, M., Nazarian, B, Anton, J.-L., and Coulon, O. (2016). Human-like brain specialization in baboons: an in vivo anatomical MRI study of language areas homologs in 96 subjects. In S.G. Roberts et al. (Eds.), *The evolution of language: Proceedings of the 11th International Conference (EVLANG11)*. New Orleans, LA, USA.
- Minier, L., Fagot, J., and Rey, A. (2016). The temporal dynamics of regularity extraction in nonhuman primates. *Cognitive Science*, 40, 1019–1030.
- Navon, D. (1977). Forest before the tree: The precedence of global feature in visual perception. *Cognitive Psychology*, 9, 353–383.
- Newport, E. L., and Aslin, R. N. (2004). Learning at a distance I. Statistical learning of nonadjacent dependencies. *Cognitive Psychology*, 48, 127–162.
- Newport, E. L., Hauser, M. D., Spaepen, G., and Aslin, R. N. (2004). Learning at a distance II. Statistical learning of nonadjacent dependencies in a nonhuman primate. *Cognitive Psychology*, 49, 85–117.
- Orbán, G., Fiser, J., Aslin, R.N., and Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences USA*, 105, 2745–50.
- Pepperberg, I.M. (2006). Ordinality and inferential abilities of a grey parrot (*Psittacus erithacus*). *Journal of Comparative Psychology*, 120, 205–216.
- Premack, D. (1983). The codes of man and beast. *Behavioral and Brain Sciences*, 6, 125–137.
- Petkov, C. I., and Wilson, B. (2012). On the pursuit of the brain network for proto-syntactic learning in nonhuman primates: conceptual issues and neurobiological hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2077–2088.
- Procyk, E., Ford Dominey, P., Amiez, C., and Joseph, J. P. (2000). The effects of sequence structure and reward schedule on serial reaction time learning in the monkey. *Brain Research. Cognitive Brain Research*, 9, 239–248.
- Quinn, P.C., Eimas P.D., and Rosenkrantz, S.L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception*, 22, 463–475.

- Ravignani, A., Sonnweber, R.-S., Stobbe, N., and Fitch, W. T. (2013). Action at a distance: dependency sensitivity in a New World primate. *Biology Letters*, 9, 20130852.
- Reader, S.M., Hager, Y. and Laland, K.N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society, B*, 366, 1077–1027.
- Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Saffran, J.R., and Thiessen, E.D. (2008). Domain-general learning capacities. In E. Hoff and M. Shatz (Eds.), *Blackwell Handbook of Language Development* (pp. 68–86). New-York: Blackwell Publishing Ltd.
- Schlenker, P., Chemla, E., and Zuberbühler, K. (2016). What do monkey calls mean? *Trends in Cognitive Sciences*, 20, 894–904.
- Seidenberg M., and MacDonald M. (1999). A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23, 569–588.
- Seidenberg, M., MacDonald, M., and Saffran, J. (2002) Does grammar start where statistics stop? *Science*, 295, 553–554.
- Sidman, M., and Tailby, W. (1982). Conditional discrimination vs. matching to sample: an expansion of the testing paradigm. *Journal of Experimental Analysis of Behavior*, 37, 5–22.
- Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Smith, J. D., Berg, M.E., Cook, R.G., Murphy, M.S., Crossley, M.J., Boomer, J., Spiering, B., Beran, M.J., Church, B., Ashby, G., and Grace, R.C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience and Biobehavioral Reviews*, 36, 2355–2369.
- Spierings, M. J., and Ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E3977-3984.
- Suzuki, T. N., Wheatcroft, D., and Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communication*, 7, 10986.

- Taglialatela, J.P., Russell, J.L., Schaeffer, J.A., and Hopkins, W.D. (2008). Communicative signaling activates “Broca’s” homologue in chimpanzees. *Current Biology*, 18, 343–348.
- Taglialatela, J.P., Russell, J.L., Schaeffer, J.A., and Hopkins, W.D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS ONE*, 6, e18852.
- Terrace, H. S. (1991). Chunking during serial learning by a pigeon: I. Basic evidence. *Journal of Experimental Psychology: Animal Behavior Processes*, 17, 81–93.
- Terrace, H. S., Son, L. K., and Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, 14, 66–73.
- Thompson, R. K. R., and Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science*, 24, 363–396.
- Tomasello, M. (2005). *Constructing a language*. Harvard: Harvard University Press.
- Toro, J. M., and Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception and Psychophysics*, 67, 867–875.
- Wang, L., Uhrig, L., Jarraya, B., and Dehaene, S. (2015). Representation of numerical and sequential patterns in macaque and human brains. *Current Biology*, 25, 1966–1974.
- Wasserman, E.A., Castro, L. and Fagot, J. (in press). Relational Thinking in Animals and Humans. In Call J.P. (Ed). *From Percepts to Concepts”: APA Handbook of Comparative Psychology*. New-York: American Psychological Association.
- Wasserman, E. A., Castro, L., and Freeman, J. H. (2012). Same-different categorization in rats. *Learning and Memory*, 19, 142–145.
- Wasserman, E. A., Fagot, J., and Young, M. E. (2001). Same-different conceptualization by baboons (*Papio papio*): The role of entropy. *Journal of Comparative Psychology*, 115, 42–52.
- Wasserman, E.A., Young, M.E., and Cook, R.G. (2004). Variability discrimination in humans and animals: implications for adaptive action. *American Psychologist*, 59, 879–890.
- Wilson, B., Smith, K., and Petkov, C. I. (2015). Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *The European Journal of Neuroscience*, 41, 568–578.

- Wright, A. A., and Katz, J. S. (2006). Mechanisms of same/different concept learning in primates and avians. *Behavioural Processes*, 72, 234–254.
- Wright, A. A., Santiago, H. C., Sands, S. F., Kendrick, D. F., and Cook, R. G. (1985). Memory processing of serial lists by pigeons, monkeys, and people. *Science*, 229, 287–289.
- Younger, B.A., and Cohen, L.B. (1986). Developmental change in infants' perception of correlations among attributes. *Child Development*, 5, 803–81.

IV. Conclusions

Il a été proposé que l'espèce humaine se distinguerait des autres espèces par sa capacité à traiter des grammaires supra-régulières (Fitch & Hauser, 2004), prérequis pour la formation des représentations récursivement enchâssées (Fitch, 2014; Hauser et al., 2002). L'étude comparative des capacités syntaxiques en laboratoire s'est par conséquent essentiellement concentré sur cette question cette dernière décennie (Petkov & Wilson, 2012). Notre revue détaillée de cette littérature montre que de nombreux problèmes méthodologiques empêchent de conclure quant à la capacité ou non d'une autre espèce que l'humain à traiter des grammaires supra-régulières. La principale limite que nous avons identifiée est que le matériel employé dans ces études permettait aux animaux testés de baser leur comportement sur diverses stratégies alternatives, rendant non nécessaire le traitement de la grammaire sous-jacente.

Les travaux que nous avons mené au cours de cette thèse ont confirmé que les prérequis cognitifs pour l'extraction de grammaires supra-régulières, et notamment l'apprentissage de régularités non-adjacentes, posent des difficultés particulières aux babouins. Mais ils suggèrent également que dans des conditions expérimentales prévenant l'exploitation unique de régularités locales ils se montrent sensibles à une structure en miroir, et semblent donc posséder des capacités computationnelles supra-régulières.

Cet ensemble de données pose la question de la pertinence de la simple distinction entre grammaires régulières et supra-régulières pour étudier l'histoire évolutive des capacités syntaxiques. La recherche de discontinuités évolutives permettant d'expliquer la productivité des langues naturelles par rapport aux systèmes de communication des autres espèces pourrait donc se focaliser à l'avenir sur d'autres pistes. Nous pensons notamment que le caractère séquentiel du langage implique de manière particulière la mémoire, et que cette fonction cognitive pourrait constituer l'un des facteurs cruciaux orientant le type de régularités qu'une espèce pourra exploiter. Ces contraintes s'appliquent toutefois également, bien que dans une autre mesure, à l'espèce humaine. Le texte de Perce rapporté au tout début de ce manuscrit l'illustre très bien. Il s'agit d'un palindrome de lettres (i.e. suivant une structure en miroir, la première lettre étant la même à la dernière, la seconde la même que l'avant-dernière, et ainsi de suite) :

« Trace l'inégal palindrome. [...] e mord ni la plage ni l'écart. »

Le texte dans son intégralité comprend 1247 mots. On raconte que lors d'une expérience réalisée en 1989, des enseignants et étudiants ayant lu ce texte en ont conclu qu'il devait avoir été produit sous l'influence « soit de de l'écriture automatique, soit d'un esprit gravement paranoïaque, soit de stupéfiants puissants. » (Audin & Fournel, 2014). Que penser de cette structure, qui peut être produite par un oulipien chevronné, mais ne peut être perçue par le lecteur non averti ? Etudier le traitement de la syntaxe en partant de l'usage qui en est fait, et en prenant en compte les fonctions cognitives dites *domain-general*, nous semble ainsi être l'approche la plus féconde pour comprendre l'évolution du langage.

Glossaire

Note : un lexique des concepts principaux de quelques concepts de linguistique générative peut être trouvé dans l'Encadré 1, p. 16.

ALL : *Artificial Language Learning* – Apprentissage de langages artificiels

AGL : *Artificial Grammar Learning* – Apprentissage de grammaires artificielles

BA : *Brodman Areas* – Aires de Brodmann

FLT : *Formal Language Theory* – Théorie des langages formels

NH : non-humain

Dendrophilie : hypothèse formulée par Fitch (2014), selon laquelle l'humain manifesterait une propension particulière à inférer des structures hiérarchiques à partir de données séquentielles.

Grammaire A^nB^n : grammaire supra-régulière générant des expressions constituées d'un nombre n d'éléments appartenant à une catégorie A suivis du même nombre n d'éléments d'une catégorie B, e.g. aabb, aaabbb.

Grammaire centrée-emboîtée : grammaire supra-régulière non-contextuelles générant des expressions constituées de plusieurs dépendances centrées-emboîtées, i.e. le dernier élément de la première moitié de la séquence est associé au premier élément de la seconde moitié, l'avant-dernier élément de la première moitié est associé au second élément de la deuxième moitié, et ainsi de suite. e.g. $A_1A_2B_2B_1$, $A_1A_2A_3B_3B_2B_1$.

Grammaire croisée : grammaire supra-régulière générant des expressions constituées de plusieurs dépendances croisées ; i.e. le premier élément de la première moitié de la séquence est associé au premier élément de la seconde moitié, le second élément de la première moitié est associé au second élément de la deuxième moitié, et ainsi de suite. e.g. $A_1A_2B_1B_2$, $A_1A_2A_3B_1B_2B_3$.

Grammaire en miroir : constitue un type particulier de grammaire centrée-empoîtée. Génère des expressions consistant en une série X, suivie par l'image en miroir de X. e.g. abba, abccba.

Grammaire en copie : constitue un type particulier de grammaire croisée. Génère des expressions consistant en une série X, suivie par la répétition de cette même série X. e.g. abba, abcabc.

Grammaires supra-régulières : grammaires générant des expressions nécessitant un automate à états-finis doté d'une mémoire additive pour être reconnues. On distingue les grammaires supra-régulières non contextuelles, requérant une mémoire « à pile » (en anglais, *push-down stack*), et les grammaires supra-régulières contextuelles, requérant une mémoire « à file » (*queue*).

Grammaires régulières : grammaires générant des expressions nécessitant un simple automate à états-finis pour être reconnues.

Grammaires supra-régulières indexées : grammaires dans lesquelles les différents éléments A et B sont appariés. e.g. $A_1A_2B_2B_1$, les index indiquant les appariements : A_1 est apparié avec B_1 , mais pas avec B_2 , etc. Les grammaires centrée-empoîtée et croisée sont des grammaires supra-régulières indexées, contrairement à la grammaire A^nB^n .

Pattern : structure sous-jacente à un ensemble de séquence isomorphes, e.g. ABB, ABA.

Récurtivité : une fonction récursive est une fonction pouvant être réappliquée à un résultat (*output*) issu d'une application antérieure de cette même fonction. En linguistique générative, on parle d'enchâssement récursif des syntagmes.

Références

- Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, 23(3), 275–290.
<https://doi.org/10.1515/TLR.2006.011>
- Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, 14(8), 1067–1074.
<https://doi.org/10.1038/nn.2869>
- Altmann, G. T. M. (2002). Learning and development in neural networks—the importance of prior experience. *Cognition*, 85(2), B43–50.
- Altmann, G. T. M., & Dienes, Z. (1999). Rule Learning by Seven-Month-Old Infants and Neural Networks. *Science*, 284(5416), 875–875.
<https://doi.org/10.1126/science.284.5416.875a>
- Ambrus, G. G., Janacsek, K., Trimborn, A. B. C., Kovacs, G., & Nemeth, D. (2017). When Less is More: Enhanced Statistical Learning After Disruption of Bilateral DLPFC. *BioRxiv*, 198515. <https://doi.org/10.1101/198515>
- Amiez, C., & Petrides, M. (2007). Selective involvement of the mid-dorsolateral prefrontal cortex in the coding of the serial order of visual stimuli in working memory. *Proceedings of the National Academy of Sciences*, 104(34), 13786–13791.
<https://doi.org/10.1073/pnas.0706220104>
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of Conditional Probability Statistics by 8-Month-Old Infants. *Psychological Science*, 9(4), 321–324.
- Audin, M., & Fournel, P. (Éd.). (2014). *OULIPO L'Abécédaire provisoirement définitif*. Larousse.
- Bach, E., Brown, C., & Marslen-wilson, W. (1986). Crossed and nested dependencies in German and Dutch: A psycholinguistic study. *Language and Cognitive Processes*, 1(4), 249–262. <https://doi.org/10.1080/01690968608404677>
- Bahlmann, J., Schubotz, R. I., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*, 42(2), 525–534.
<https://doi.org/10.1016/j.neuroimage.2008.04.249>

- Bahlmann, J., Schubotz, R. I., Mueller, J. L., Koester, D., & Friederici, A. D. (2009). Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*, *1298*, 161-170. <https://doi.org/10.1016/j.brainres.2009.08.017>
- Bates, E., & Elman, J. (1996). Learning rediscovered. *Science (New York, N.Y.)*, *274*(5294), 1849-1850.
- Beckers, G. J. L., Berwick, R. C., Okanoya, K., & Bolhuis, J. J. (2016). What do animals learn in artificial grammar studies? *Neuroscience & Biobehavioral Reviews*. <https://doi.org/10.1016/j.neubiorev.2016.12.021>
- Beckers, G. J. L., Bolhuis, J. J., Okanoya, K., & Berwick, R. C. (2012). Birdsong neurolinguistics: songbird context-free grammar claim is premature. *Neuroreport*, *23*(3), 139-145. <https://doi.org/10.1097/WNR.0b013e32834f1765>
- Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672-1677. <https://doi.org/10.1073/pnas.0809667106>
- Bergman, T. J., Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical Classification by Rank and Kinship in Baboons. *Science*, *302*(5648), 1234-1236. <https://doi.org/10.1126/science.1087513>
- Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences*, *15*(3), 113-121. <https://doi.org/10.1016/j.tics.2011.01.002>
- Bresnan, J., Kaplan, R. M., Peters, S., & Zaenen, A. (1982). Cross-Serial Dependencies in Dutch. *Linguistic Inquiry*, *13*(4), 613-635. <https://doi.org/10.2307/4178298>
- Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology. General*, *120*(3), 316-323. <http://dx.doi.org/10.1037/0096-3445.120.3.316>
- Bybee, J. (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

- Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition*, 20(4), 755-769. <https://doi.org/10.1007/s10071-017-1096-4>
- Byrne, R. W., & Russon, A. E. (1998). Learning by imitation: a hierarchical approach. *The Behavioral and Brain Sciences*, 21(5), 667-684; discussion 684-721.
- Byrne, Richard W., Corp, N., & Byrne, J. M. E. (2001). Estimating the Complexity of Animal Behaviour: How Mountain Gorillas Eat Thistles. *Behaviour*, 138(4), 525-557.
- Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Context-related call combinations in female Diana monkeys. *Animal Cognition*, 15(3), 327-339. <https://doi.org/10.1007/s10071-011-0456-8>
- Cantalupo, C., & Hopkins, W. D. (2001). Asymmetric Broca's area in great apes. *Nature*, 414(6863), 505. <https://doi.org/10.1038/35107134>
- Cäsar, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. *Biology Letters*, 9(5), 20130535. <https://doi.org/10.1098/rsbl.2013.0535>
- Chaplin, T. A., Yu, H.-H., Soares, J. G. M., Gattass, R., & Rosa, M. G. P. (2013). A conserved pattern of differential expansion of cortical areas in simian primates. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(38), 15120-15125. <https://doi.org/10.1523/JNEUROSCI.2909-13.2013>
- Chekaf, M., Cowan, N., & Mathy, F. (2016). Chunk formation in immediate memory and how it relates to data compression. *Cognition*, 155, 96-107. <https://doi.org/10.1016/j.cognition.2016.05.024>
- Chekaf, M., Gauvrit, N., Guida, A., & Mathy, F. (2018). Compression in Working Memory and Its Relationship With Fluid Intelligence. *Cognitive Science*. <https://doi.org/10.1111/cogs.12601>
- Chen, J., & ten Cate, C. (2015). Zebra finches can use positional and transitional cues to distinguish vocal element strings. *Behavioural Processes*, 117, 29-34. <https://doi.org/10.1016/j.beproc.2014.09.004>
- Chen, J., & ten Cate, C. (2017). Bridging the gap: Learning of acoustic nonadjacent dependencies by a songbird. *Journal of Experimental Psychology: Animal Learning and Cognition*, 43(3), 295-302. <https://doi.org/10.1037/xan0000145>

- Chen, J., van Rossum, D., & ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition*, *18*(1), 151-164.
<https://doi.org/10.1007/s10071-014-0786-4>
- Cheney, D. L., & Seyfarth, R. M. (2007). *Baboon Metaphysics*. The University of Chicago Press.
- Chomsky, N. (1956). Three models for the description of language. *IRE Trans Inf Theory*, *IT-2*, 113-124.
- Chomsky, N. (1957). *Syntactic Structures*. Walter de Gruyter.
- Chomsky, N. (1980). Rules and representations. *Behavioral and Brain Sciences*, *3*(1), 1-15.
<https://doi.org/10.1017/S0140525X00001515>
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press.
- Chomsky, N. (2017). The language capacity: architecture and evolution. *Psychonomic Bulletin & Review*, *24*(1), 200-203. <https://doi.org/10.3758/s13423-016-1078-6>
- Christiansen, M. H., & Curtin, S. (1999). Transfer of learning: rule acquisition or statistical learning? *Trends in Cognitive Sciences*, *3*(8), 289-290.
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *The Behavioral and Brain Sciences*, *31*(5), 489-508; discussion 509-558.
<https://doi.org/10.1017/S0140525X08004998>
- Christiansen, M. H., & Chater, N. (2015). The language faculty that wasn't: a usage-based account of natural language recursion. *Frontiers in Psychology*, *6*.
<https://doi.org/10.3389/fpsyg.2015.01182>
- Ciochon, R. L., & Fleagle, J. G. (1987). *Primate Evolution and Human Origins*. Transaction Pub.
- Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The Syntax and Meaning of Wild Gibbon Songs. *PLOS ONE*, *1*(1), e73. <https://doi.org/10.1371/journal.pone.0000073>
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology. General*, *120*(3), 235-253.
- Comins, J. A., & Gentner, T. Q. (2013). Perceptual categories enable pattern generalization in songbirds. *Cognition*, *128*(2), 113-118.
<https://doi.org/10.1016/j.cognition.2013.03.014>

- Comins, J. A., & Gentner, T. Q. (2014). Auditory temporal pattern learning by songbirds using maximal stimulus diversity and minimal repetition. *Animal Cognition*, *17*(5), 1023-1030. <https://doi.org/10.1007/s10071-014-0732-5>
- Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences*, *5*(12), 539-546. [https://doi.org/10.1016/S1364-6613\(00\)01800-3](https://doi.org/10.1016/S1364-6613(00)01800-3)
- Corballis, M. C. (2007). Recursion, language, and starlings. *Cognitive Science*, *31*(4), 697-704. <https://doi.org/10.1080/15326900701399947>
- Corballis, M. C. (2009). Do rats learn rules? *Animal Behaviour*, *78*(4), e1-e2. <https://doi.org/10.1016/j.anbehav.2009.05.001>
- Corp, N., & Byrne, R. W. (2002). Leaf Processing by Wild Chimpanzees: Physically Defended Leaves Reveal Complex Manual Skills. *Ethology*, *108*(8), 673-696. <https://doi.org/10.1046/j.1439-0310.2002.00801.x>
- Corsi, P. M. (1972). Human Memory and the Medial Temporal Region of the Brain. *Dissertation Abstracts International*, *34*.
- Coye, C., Zuberbühler, K., & Lemasson, A. (2016). Morphologically structured vocalizations in female Diana monkeys. *Animal Behaviour*, *115*(Supplement C), 97-105. <https://doi.org/10.1016/j.anbehav.2016.03.010>
- Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1119-1130. <https://doi.org/10.1037/0278-7393.30.5.1119>
- Culy, C. (1985). The Complexity of the Vocabulary of Bambara. In *The Formal Complexity of Natural Language* (p. 349-357). Springer, Dordrecht. https://doi.org/10.1007/978-94-009-3401-6_14
- de la Mora, D. M., & Toro, J. M. (2013). Rule learning over consonants and vowels in a non-human animal. *Cognition*, *126*(2), 307-312. <https://doi.org/10.1016/j.cognition.2012.09.015>

- de Vries, M. H., Monaghan, P., Knecht, S., & Zwitserlood, P. (2008). Syntactic structure and artificial grammar learning: the learnability of embedded hierarchical structures. *Cognition*, *107*(2), 763-774. <https://doi.org/10.1016/j.cognition.2007.09.002>
- Dobson, C. W., & Lemon, R. E. (1979). Markov Sequences in Songs of American Thrushes. *Behaviour*, *68*(1/2), 86-105.
- Eimer, M., Goschke, T., Schlaghecken, F., & Stürmer, B. (1996). Explicit and implicit learning of event sequences: evidence from event-related brain potentials. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *22*(4), 970-987.
- Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, *105*(2), 247-299. <https://doi.org/10.1016/j.cognition.2006.09.010>
- Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes' edge: positional learning in chimpanzees and humans. *Animal Cognition*, *13*(3), 483-495. <https://doi.org/10.1007/s10071-009-0299-8>
- Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition*, *105*(3), 577-614. <https://doi.org/10.1016/j.cognition.2006.12.014>
- Endress, A. D., Nespors, M., & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences*, *13*(8), 348-353. <https://doi.org/10.1016/j.tics.2009.05.005>
- Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., & Bolhuis, J. J. (2015). Structures, Not Strings: Linguistics as Part of the Cognitive Sciences. *Trends in Cognitive Sciences*, *19*(12), 729-743. <https://doi.org/10.1016/j.tics.2015.09.008>
- Everett, D. L. (2005). Cultural Constraints on Grammar and Cognition in Pirahã Another Look at the Design Features of Human Language. *Current Anthropology*, *46*(4), 621-646. <https://doi.org/10.1086/431525>
- Fagot, J., & Barbet, I. (2006). Grouping and segmentation of visual objects by baboons (*Papio papio*) and humans (*Homo sapiens*). In E. A. Wasserman & T. R. Zentall (Éd.), *Comparative Cognition: Experimental Explorations of Animal Intelligence* (p. 11-24). New York: Oxford University Press.

- Fagot, Joël, & De Lillo, C. (2011). A comparative study of working memory: Immediate serial spatial recall in baboons (*Papio papio*) and humans. *Neuropsychologia*, *49*(14), 3870-3880. <https://doi.org/10.1016/j.neuropsychologia.2011.10.003>
- Fagot, Joël, Malassis, R., & Medam, T. (2018). The processing of positional information in a two-item sequence limits the emergence of symmetry in baboons (*Papio papio*), but not in humans (*Homo sapiens*). *Learning & Behavior*, *46*(1), 67-78. <https://doi.org/10.3758/s13420-017-0290-1>
- Ferdinand, N. K., Mecklinger, A., & Kray, J. (2008). Error and deviance processing in implicit and explicit sequence learning. *Journal of Cognitive Neuroscience*, *20*(4), 629-642. <https://doi.org/10.1162/jocn.2008.20046>
- Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports*, *6*, srep25434. <https://doi.org/10.1038/srep25434>
- Fiebach, C. J., & Schubotz, R. I. (2006). Dynamic Anticipatory Processing of Hierarchical Sequential Events: a Common Role for Broca's Area and Ventral Premotor Cortex Across Domains? *Cortex*, *42*(4), 499-502. [https://doi.org/10.1016/S0010-9452\(08\)70386-1](https://doi.org/10.1016/S0010-9452(08)70386-1)
- Fischmeister, F. P., Martins, M. J. D., Beisteiner, R., & Fitch, W. T. (2017). Self-similarity and recursion as default modes in human cognition. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *97*, 183-201. <https://doi.org/10.1016/j.cortex.2016.08.016>
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(3), 458-467.
- Fitch, W. T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics (p. 73-90). <https://doi.org/10.1017/CBO9780511817755.005>
- Fitch, W. T. (2014). Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition. *Physics of Life Reviews*, *11*(3), 329-364. <https://doi.org/10.1016/j.plrev.2014.04.005>
- Fitch, W. T., & Friederici, A. D. (2012). Artificial grammar learning meets formal language theory: an overview. *Philosophical Transactions of the Royal Society of London*.

- Series B, Biological Sciences*, 367(1598), 1933-1955.
<https://doi.org/10.1098/rstb.2012.0103>
- Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science (New York, N.Y.)*, 303(5656), 377-380.
<https://doi.org/10.1126/science.1089401>
- Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2), 179-210.
<https://doi.org/10.1016/j.cognition.2005.02.005>
- Forkstam, C., Elwér, A., Ingvar, M., & Petersson, K. M. (2008). Instruction effects in implicit artificial grammar learning: a preference for grammaticality. *Brain Research*, 1221, 80-92. <https://doi.org/10.1016/j.brainres.2008.05.005>
- Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., & Petersson, K. M. (2006). Neural correlates of artificial syntactic structure classification. *NeuroImage*, 32(2), 956-967.
<https://doi.org/10.1016/j.neuroimage.2006.03.057>
- Friederici, A. D. (2011). The Brain Basis of Language Processing: From Structure to Function. *Physiological Reviews*, 91(4), 1357-1392.
<https://doi.org/10.1152/physrev.00006.2011>
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2458-2463. <https://doi.org/10.1073/pnas.0509389103>
- Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., & von Cramon, D. Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 16(12), 1709-1717.
<https://doi.org/10.1093/cercor/bhj106>
- Froehlich, A. L., Herbranson, W. T., Loper, J. D., Wood, D. M., & Shimp, C. P. (2004). Anticipating by Pigeons Depends on Local Statistical Information in a Serial Response Time Task. *Journal of Experimental Psychology: General*, 133(1), 31-45.
<https://doi.org/10.1037/0096-3445.133.1.31>

- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117-125. <https://doi.org/10.1016/j.tics.2014.12.010>
- Frost, R. L. A., & Monaghan, P. (2016). Simultaneous segmentation and generalisation of non-adjacent dependencies from continuous speech. *Cognition, 147*, 70-74. <https://doi.org/10.1016/j.cognition.2015.11.010>
- Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science (New York, N.Y.), 165*(3894), 664-672.
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review, 16*(3), 486-490. <https://doi.org/10.3758/PBR.16.3.486>
- Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*(7088), 1204-1207. <https://doi.org/10.1038/nature04675>
- Gentner, T. Q., Fenn, K., Margoliash, D., & Nusbaum, H. (2010). Simple stimuli, simple strategies. *Proceedings of the National Academy of Sciences, 107*(16), E65-E65. <https://doi.org/10.1073/pnas.1000501107>
- Gentner, T. Q., & Hulse, S. H. (1998). Perceptual mechanisms for individual vocal recognition in European starlings, *Sturnus vulgaris*. *Animal Behaviour, 56*(3), 579-594. <https://doi.org/10.1006/anbe.1998.0810>
- Gervain, J., Macagno, F., Cogoi, S., Pena, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences, 105*(37), 14222-14227. <https://doi.org/10.1073/pnas.0806530105>
- Gervain, Judit, & Werker, J. F. (2013). Learning non-adjacent regularities at age 0 ; 7. *Journal of Child Language, 40*(4), 860-872. <https://doi.org/10.1017/S0305000912000256>
- Ghirlanda, S., Lind, J., & Enquist, M. (2017). Memory for stimulus sequences: a divide between humans and other animals? *Royal Society Open Science, 4*(6), 161011. <https://doi.org/10.1098/rsos.161011>

- Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of « sameness » and « difference » in an insect. *Nature*, *410*(6831), 930-933.
<https://doi.org/10.1038/35073582>
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431-436.
- Gómez, R. L., & Gerken, L. (2000). Infant Artificial Language Learning and Language Acquisition. *Trends in Cognitive Sciences*, *4*(5), 178-186.
- Gómez, R., & Maye, J. (2005). The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy*, *7*(2), 183-206. https://doi.org/10.1207/s15327078in0702_4
- Gonzalez-Gomez, N., & Nazzi, T. (2012). Acquisition of Nonadjacent Phonological Dependencies in the Native Language During the First Year of Life. *Infancy*, *17*(5), 498-524. <https://doi.org/10.1111/j.1532-7078.2011.00104.x>
- Green, S. R., Mercado, E., Pack, A. A., & Herman, L. M. (2011). Recurring patterns in the songs of humpback whales (*Megaptera novaeangliae*). *Behavioural Processes*, *86*(2), 284-294. <https://doi.org/10.1016/j.beproc.2010.12.014>
- Greenfield, P. M. (1991). Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, *14*(4), 531-551. <https://doi.org/10.1017/S0140525X00071235>
- Hakes, D. T., Evans, J. S., & Brannon, L. L. (1976). Understanding sentences with relative clauses. *Memory & Cognition*, *4*(3), 283-290. <https://doi.org/10.3758/BF03213177>
- Hauser, Marc D, Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science (New York, N.Y.)*, *298*(5598), 1569-1579.
<https://doi.org/10.1126/science.298.5598.1569>
- Hauser, Marc D, Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition*, *78*(3), B53-B64. [https://doi.org/10.1016/S0010-0277\(00\)00132-3](https://doi.org/10.1016/S0010-0277(00)00132-3)
- Hauser, Marc D., & Watumull, J. (2017). The Universal Generative Faculty: The source of our expressive power in language, mathematics, morality, and music. *Journal of Neurolinguistics*, *43*(Part B), 78-94. <https://doi.org/10.1016/j.jneuroling.2016.10.005>

- Hauser, Marc David, & Glynn, D. (2009). Can free-ranging rhesus monkeys (*Macaca mulatta*) extract artificially created rules comprised of natural vocalizations? *Journal of Comparative Psychology*, *123*(2), 161-167. <https://doi.org/10.1037/a0015584>
- Herbranson, W. T., & Shimp, C. P. (2003). "Artificial grammar learning" in pigeons: A preliminary analysis. *Learning & Behavior*, *31*(1), 98-106. <https://doi.org/10.3758/BF03195973>
- Herbranson, W. T., & Shimp, C. P. (2008). Artificial grammar learning in pigeons. *Learning & Behavior*, *36*(2), 116-137.
- Hoch, L. (2010, juillet 9). *Perception et apprentissage des structures musicales et langagières : études des ressources cognitives partagées et des effets attentionnels*. Lyon 2.
- Hochmann, J.-R., Azadpour, M., & Mehler, J. (2008). Do Humans Really Learn An Bn Artificial Grammars From Exemplars? *Cognitive Science: A Multidisciplinary Journal*, *32*(6), 1021-1036. <https://doi.org/10.1080/03640210801897849>
- Hohle, B., Schmitz, M., Santelmann, L. M., & Weissenborn, J. (2006). The Recognition of Discontinuous Verbal Dependencies by German 19-Month-Olds: Evidence for Lexical and Structural Influences on Children's Early Processing Capacities. *Language Learning and Development*, *2*(4), 277-300. https://doi.org/10.1207/s15473341lld0204_3
- Holy, T. E., & Guo, Z. (2005). Ultrasonic Songs of Male Mice. *PLOS Biology*, *3*(12), e386. <https://doi.org/10.1371/journal.pbio.0030386>
- Hopkins, W. D., Meguerditchian, A., Coulon, O., Misiura, M., Pope, S., Mareno, M. C., & Schapiro, S. J. (2017). Motor skill for tool-use is associated with asymmetries in Broca's area and the motor hand area of the precentral gyrus in chimpanzees (*Pan troglodytes*). *Behavioural Brain Research*, *318*, 71-81. <https://doi.org/10.1016/j.bbr.2016.10.048>
- Hopkins, W. D., Russell, J. L., & Cantalupo, C. (2007). Neuroanatomical Correlates of Handedness for Tool Use in Chimpanzees (*Pan troglodytes*). *Psychological science*, *18*(11), 971-977. <https://doi.org/10.1111/j.1467-9280.2007.02011.x>
- Hosino, T., & Okanoya, K. (2000). Lesion of a higher-order song nucleus disrupts phrase level complexity in Bengalese finches. *Neuroreport*, *11*(10), 2091-2095.

- Humboldt, W. F. von. (1836). *Über die Verschiedenheit des menschlichen Sprachbaues: und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Druckerei der Königlichen Akademie der Wissenschaften.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*(4), 665-679. <https://doi.org/10.1162/jocn.2006.18.4.665>
- Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, *130*(4), 658-680. <https://doi.org/10.1037/0096-3445.130.4.658>
- Inoue, M., & Mikami, A. (2006). Prefrontal activity during serial probe reproduction task: encoding, mnemonic, and retrieval processes. *Journal of Neurophysiology*, *95*(2), 1008-1041. <https://doi.org/10.1152/jn.00552.2005>
- Jäger, G., & Rogers, J. (2012). Formal language theory: refining the Chomsky hierarchy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1956-1970. <https://doi.org/10.1098/rstb.2012.0077>
- Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., & Slemmer, J. A. (2009). Abstract Rule Learning for Visual Sequences in 8- and 11-Month-Olds. *Infancy: the official journal of the International Society on Infant Studies*, *14*(1), 2-18. <https://doi.org/10.1080/15250000802569611>
- Kakishita, Y., Sasahara, K., Nishino, T., Takahasi, M., & Okanoya, K. (2009). Ethological data mining: an automata-based approach to extract behavioral units and rules. *Data Mining and Knowledge Discovery*, *18*(3), 446-471. <https://doi.org/10.1007/s10618-008-0122-1>
- Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. *Journal of Linguistics*, *43*(2), 365-392. <https://doi.org/10.1017/S0022226707004616>
- Karlsson, F. (2010). Syntactic recursion and iteration. In H. van der Hulst (Éd.), *Recursion and Human Language* (p. 43-67). Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110219258>
- Kershenbaum, A., Bowles, A. E., Freeberg, T. M., Jin, D. Z., Lameira, A. R., & Bohn, K. (2014). Animal vocal sequences: not the Markov chains we thought they were.

- Proceedings of the Royal Society of London B: Biological Sciences*, 281(1792), 20141370. <https://doi.org/10.1098/rspb.2014.1370>
- Kinsella, A. R. (2010). 10. Was recursion the key step in the evolution of the human language faculty? In *Recursion and Human Language*. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110219258.177>
- Koechlin, E., & Jubault, T. (2006). Broca's Area and the Hierarchical Organization of Human Behavior. *Neuron*, 50(6), 963-974. <https://doi.org/10.1016/j.neuron.2006.05.017>
- Lai, J., & Poletiek, F. H. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center-embedded hierarchical structures. *Cognition*, 118(2), 265-273. <https://doi.org/10.1016/j.cognition.2010.11.011>
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical prerequisites. Volume I*. Stanford University Press.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar: descriptive application. Volume II*. Stanford University Press.
- Lany, J., & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning. *Psychological Science*, 19(12), 1247-1252. <https://doi.org/10.1111/j.1467-9280.2008.02233.x>
- Levelt, W. J. M. (2008). *Formal Grammars in Linguistics and Psycholinguistics: Volume 1: An Introduction to the Theory of Formal Languages and Automata, Volume 2: Applications in Linguistic Theory, Volume 3: Psycholinguistic Applications*. John Benjamins Publishing.
- Lobina, D. J. (2011). "A Running Back" and Forth: A Review of Recursion and Human Language. *BIOLINGUISTICS*, 5(1-2), 151-169.
- Lyn, H. (2017). The Question of Capacity: Why Enculturated and Trained Animals have much to Tell Us about the Evolution of Language. *Psychonomic Bulletin & Review*, 24(1), 85-90. <https://doi.org/10.3758/s13423-016-1129-z>
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4(5), 540-545. <https://doi.org/10.1038/87502>
- Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the*

National Academy of Sciences, 106(20), 8362-8367.

<https://doi.org/10.1073/pnas.0810928106>

Malassis, R., Rey, A., & Fagot, J. (in press). Non-adjacent Dependencies Processing in Human and Non-Human Primates. *Cognitive Science*.

Manza, L., & Bornstein, R. F. (1995). Affective discrimination and the implicit learning process. *Consciousness and Cognition*, 4(4), 399-409.

<https://doi.org/10.1006/ccog.1995.1047>

Marcus, G. F. (2001). *The Algebraic Mind*. MIT Press.

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: the exception that proves the rule. *Cognitive Psychology*, 29(3), 189-256.

<https://doi.org/10.1006/cogp.1995.1015>

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, 283(5398), 77-80.

<https://doi.org/10.1126/science.283.5398.77>

Marcus, Gary F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, 18(5), 387-391. <https://doi.org/10.1111/j.1467-9280.2007.01910.x>

Marks, L. E. (1968). Scaling of grammaticalness of self-embedded English sentences. *Journal of Verbal Learning and Verbal Behavior*, 7(5), 965-967.

[https://doi.org/10.1016/S0022-5371\(68\)80106-9](https://doi.org/10.1016/S0022-5371(68)80106-9)

Martins, M. D. (2012). Distinctive signatures of recursion. *Phil. Trans. R. Soc. B*, 367(1598), 2055-2064. <https://doi.org/10.1098/rstb.2012.0097>

Meguerditchian, A., Gardner, M. J., Schapiro, S. J., & Hopkins, W. D. (2012). The sound of one-hand clapping: handedness and perisylvian neural correlates of a communicative gesture in chimpanzees. *Proceedings of the Royal Society of London B: Biological Sciences*, 279(1735), 1959-1966. <https://doi.org/10.1098/rspb.2011.2485>

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1007-1028.

- Miller, G.A. (1967). Project grammarama. In G.A. Miller (Éd.), *Psychology of communication*. New York: Basic Books.
- Miller, George A., & Chomsky, N. (1963). Finitary Models of Language Users. In D. Luce (Éd.), *Handbook of Mathematical Psychology* (p. 2–419). John Wiley & Sons.
- Miller, George A., & Isard, S. (1964). Free recall of self-embedded english sentences. *Information and Control*, 7(3), 292-303. [https://doi.org/10.1016/S0019-9958\(64\)90310-9](https://doi.org/10.1016/S0019-9958(64)90310-9)
- Milne, A. E., Mueller, J. L., Männel, C., Attaheri, A., Friederici, A. D., & Petkov, C. I. (2016). Evolutionary origins of non-adjacent sequence processing in primate brain potentials. *Scientific Reports*, 6, 36259. <https://doi.org/10.1038/srep36259>
- Minier, L., Fagot, J., & Rey, A. (2016). The Temporal Dynamics of Regularity Extraction in Non-Human Primates. *Cognitive Science*, 40(4), 1019-1030. <https://doi.org/10.1111/cogs.12279>
- Misyak, J. B., & Christiansen, M. H. (2012). Statistical Learning and Language: An Individual Differences Study. *Language Learning*, 62(1), 302-331. <https://doi.org/10.1111/j.1467-9922.2010.00626.x>
- Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, 109(39), 15953-15958. <https://doi.org/10.1073/pnas.1204319109>
- Murphy, R. A., Mondragón, E., & Murphy, V. A. (2008). Rule Learning by Rats. *Science*, 319(5871), 1849-1851. <https://doi.org/10.1126/science.1151564>
- Nazzi, T., Barrière, I., Goyet, L., Kresh, S., & Legendre, G. (2011). Tracking irregular morphophonological dependencies in natural language: evidence from the acquisition of subject-verb agreement in French. *Cognition*, 120(1), 119-135. <https://doi.org/10.1016/j.cognition.2011.03.004>
- Negishi, M. (1999). Do infants learn grammar with algebra or statistics? *Science (New York, N.Y.)*, 284(5413), 435; author reply 436-437.
- Nelson, M. J., Karoui, I. E., Giber, K., Yang, X., Cohen, L., Koopman, H., ... Dehaene, S. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences*, 114(18), E3669-E3678. <https://doi.org/10.1073/pnas.1701590114>

- Nespor, M., Peña, M., & Mehler, J. (2003). On the Different Roles of Vowels and Consonants in Speech Processing and Language Acquisition. *Lingue e linguaggio*, (2/2003).
<https://doi.org/10.1418/10879>
- Neubert, F.-X., Mars, R. B., Thomas, A. G., Sallet, J., & Rushworth, M. F. S. (2014). Comparison of Human Ventral Frontal Cortex Areas for Cognitive Control and Language with Areas in Monkey Frontal Cortex. *Neuron*, 81(3), 700-713.
<https://doi.org/10.1016/j.neuron.2013.11.012>
- Nevins, A. I., Pesetsky, D., & Rodrigues, C. (2009). Pirahã Exceptionality: A Reassessment. Consulté à l'adresse <https://dash.harvard.edu/handle/1/3597237>
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127-162.
- Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, 49(2), 85-117. <https://doi.org/10.1016/j.cogpsych.2003.12.002>
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1-32.
[https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)
- O'Donnell, T. J., Hauser, M. D., & Fitch, W. T. (2005). Using mathematical models of language experimentally. *Trends in Cognitive Sciences*, 9(6), 284-289.
<https://doi.org/10.1016/j.tics.2005.04.011>
- Ojima, S., & Okanoya, K. (2014). The Non-Hierarchical Nature of the Chomsky Hierarchy-Driven Artificial-Grammar Learning. *BIOLINGUISTICS*, 8(0), 163-180.
- Onnis, L., Monaghan, P., Christiansen, M. H., & Chater, N. (2005). Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. In *PROCEEDINGS OF THE TWENTY-SIXTH ANNUAL CONFERENCE OF THE COGNITIVE SCIENCE SOCIETY*.
- Onnis, Luca, Monaghan, P., Richmond, K., & Chater, N. (2005). Phonology impacts segmentation in online speech processing. *Journal of Memory and Language*, 53(2), 225-237. <https://doi.org/10.1016/j.jml.2005.02.011>
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National*

- Academy of Sciences*, 106(51), 22026-22031.
<https://doi.org/10.1073/pnas.0908118106>
- Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(1), 80-96. <https://doi.org/10.1037/0278-7393.34.1.80>
- Pacton, S., Sobaco, A., & Perruchet, P. (2015). Is an attention-based associative account of adjacent and nonadjacent dependency learning valid? *Acta Psychologica*, 157, 195-199. <https://doi.org/10.1016/j.actpsy.2015.03.002>
- Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, 108(6), 2522-2527. <https://doi.org/10.1073/pnas.1018711108>
- Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science (New York, N.Y.)*, 298(5593), 604-607.
<https://doi.org/10.1126/science.1072901>
- Perruchet, P., & Pacteau, C. (1990). Synthetic Grammar Learning: Implicit Rule Abstraction or Explicit Fragmentary Knowledge? *Journal of Experimental Psychology: General*, 119, 264-275. <https://doi.org/10.1037/0096-3445.119.3.264>
- Perruchet, P., & Rey, A. (2005). Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review*, 12(2), 307-313.
- Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: no need for algebraic-like computations. *Journal of Experimental Psychology: General*, 133(4), 573-583. <https://doi.org/10.1037/0096-3445.133.4.573>
- Pettersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, 28(3), 383-407.
<https://doi.org/10.1016/j.cogsci.2003.12.003>
- Pettersson, K.-M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, 120(2), 83-95.
<https://doi.org/10.1016/j.bandl.2010.08.003>

- Petkov, C. I., & Wilson, B. (2012). On the pursuit of the brain network for proto-syntactic learning in non-human primates: conceptual issues and neurobiological hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1598), 2077-2088. <https://doi.org/10.1098/rstb.2012.0073>
- Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *The European Journal of Neuroscience*, 16(2), 291-310.
- Petrides, Michael, Cadoret, G., & Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature*, 435(7046), 1235-1238. <https://doi.org/10.1038/nature03628>
- Petrides, Michael, & Pandya, D. N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biology*, 7(8), e1000170. <https://doi.org/10.1371/journal.pbio.1000170>
- Pinker, S. (1991). Rules of language. *Science (New York, N.Y.)*, 253(5019), 530-535.
- Pinker, Steven, & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, 95(2), 201-236. <https://doi.org/10.1016/j.cognition.2004.08.004>
- Poletiek, F. H., Fitz, H., & Bocanegra, B. R. (2016). What baboons can (not) tell us about natural language grammars. *Cognition*, 151, 108-112. <https://doi.org/10.1016/j.cognition.2015.04.016>
- Pothos, E. M. (2005). The rules versus similarity distinction. *The Behavioral and Brain Sciences*, 28(1), 1-14; discussion 14-49.
- Procyk, E., Ford Dominey, P., Amiez, C., & Joseph, J. P. (2000). The effects of sequence structure and reward schedule on serial reaction time learning in the monkey. *Brain Research. Cognitive Brain Research*, 9(3), 239-248.
- Ravignani, A., & Sonnweber, R. (2017). Chimpanzees process structural isomorphisms across sensory modalities. *Cognition*, 161, 74-79. <https://doi.org/10.1016/j.cognition.2017.01.005>
- Ravignani, A., Sonnweber, R.-S., Stobbe, N., & Fitch, W. T. (2013). Action at a distance: dependency sensitivity in a New World primate. *Biology Letters*, 9(6), 20130852. <https://doi.org/10.1098/rsbl.2013.0852>

- Ravignani, A., Westphal-Fitch, G., Aust, U., Schlumpp, M. M., & Fitch, W. T. (2015). More than one way to see it: Individual heuristics in avian visual computation. *Cognition*, *143*, 13-24. <https://doi.org/10.1016/j.cognition.2015.05.021>
- Reber, A. (1969). Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology*, *81*(1), 115-119. <https://doi.org/10.1037/h0027454>
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855-863. [https://doi.org/10.1016/S0022-5371\(67\)80149-X](https://doi.org/10.1016/S0022-5371(67)80149-X)
- Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (in press). Regularity extraction across species: associative learning mechanisms shared by human and non-human primates. *Topics in Cognitive Science*.
- Rey, A., Perruchet, P., & Fagot, J. (2012). Centre-embedded structures are a by-product of associative learning and working memory constraints: evidence from baboons (*Papio Papio*). *Cognition*, *123*(1), 180-184. <https://doi.org/10.1016/j.cognition.2011.12.005>
- Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, *11*(4), 426-428. <https://doi.org/10.1038/nn2072>
- Rogers, J., & Pullum, G. K. (2011). Aural Pattern Recognition Experiments and the Subregular Hierarchy. *Journal of Logic, Language and Information*, *20*(3), 329-342. <https://doi.org/10.1007/s10849-011-9140-2>
- Romanski, L. M., & Averbeck, B. B. (2009). The Primate Cortical Auditory System and Neural Representation of Conspecific Vocalizations. *Annual review of neuroscience*, *32*, 315-346. <https://doi.org/10.1146/annurev.neuro.051508.135431>
- Romberg, A. R., & Saffran, J. R. (2013). All Together Now: Concurrent Learning of Multiple Structures in an Artificial Language. *Cognitive Science*, *37*(7), 1290-1320. <https://doi.org/10.1111/cogs.12050>
- Roy, A. C., & Arbib, M. A. (2005). The syntactic motor system. *Gesture*, *5*(1), 7-37. <https://doi.org/10.1075/gest.5.1.03roy>
- Saffran, J., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, *107*(2), 479-500. <https://doi.org/10.1016/j.cognition.2007.10.010>

- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science (New York, N.Y.)*, *274*(5294), 1926-1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27-52.
- Saffran, Jenny R. (2001). The Use of Predictive Dependencies in Language Learning. *Journal of Memory and Language*, *44*(4), 493-515. <https://doi.org/10.1006/jmla.2000.2759>
- Saffran, Jenny R., Newport, E. L., & Aslin, R. N. (1996). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, *35*(4), 606-621. <https://doi.org/10.1006/jmla.1996.0032>
- Saffran, Jenny R., & Thiessen, E. D. (2007). Domain-General Learning Capacities. In E. Hoff & rilyn Shatz (Éd.), *Blackwell Handbook of Language Development* (p. 68-86). Blackwell Publishing Ltd. <https://doi.org/10.1002/9780470757833.ch4>
- Sandoval, M., & Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(5), 511-522. <https://doi.org/10.1002/wcs.1244>
- Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition*, *69*(2), 105-134. [https://doi.org/10.1016/S0010-0277\(98\)00060-2](https://doi.org/10.1016/S0010-0277(98)00060-2)
- Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in Broca's area. *NeuroImage*, *37*(1), 8-17. <https://doi.org/10.1016/j.neuroimage.2007.04.047>
- Santolin, C., Rosa-Salva, O., Regolin, L., & Vallortigara, G. (2016). Generalization of visual regularities in newly hatched chicks (*Gallus gallus*). *Animal Cognition*, *19*(5), 1007-1017. <https://doi.org/10.1007/s10071-016-1005-2>
- Santolin, C., & Saffran, J. R. (2018). Constraints on Statistical Learning Across Species. *Trends in Cognitive Sciences*, *22*(1), 52-63. <https://doi.org/10.1016/j.tics.2017.10.003>
- Savage-Rumbaugh, E. S., & Lewin, R. (1996). *Kanzi: the ape at the brink of the human mind*. Wiley.
- Schel, A. M., Tranquilli, S., & Zuberbühler, K. (2009). The alarm call system of two species of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*).

- Journal of Comparative Psychology* (Washington, D.C.: 1983), 123(2), 136-150.
<https://doi.org/10.1037/a0014280>
- Schenker, N. M., Hopkins, W. D., Spocter, M. A., Garrison, A. R., Stimpson, C. D., Erwin, J. M., ... Sherwood, C. C. (2010). Broca's Area Homologue in Chimpanzees (Pan troglodytes): Probabilistic Mapping, Asymmetry, and Comparison to Humans. *Cerebral Cortex*, 20(3), 730-742. <https://doi.org/10.1093/cercor/bhp138>
- Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., ... Zuberbühler, K. (2014). Monkey semantics: two 'dialects' of Campbell's monkey alarm calls. *Linguistics and Philosophy*, 37(6), 439-501.
<https://doi.org/10.1007/s10988-014-9155-7>
- Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J. P., Kuhn, J., ... Zuberbühler, K. (2016a). Formal monkey linguistics : the debate. <http://dx.doi.org/10.1515/tl-2016-0010>
- Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., ... Zuberbühler, K. (2016b). Formal monkey linguistics. *Theoretical Linguistics*, 42(1-2), 1-90.
<https://doi.org/10.1515/tl-2016-0001>
- Schlenker, P., Chemla, E., & Zuberbühler, K. (2016). What Do Monkey Calls Mean? *Trends in Cognitive Sciences*, 20(12), 894-904. <https://doi.org/10.1016/j.tics.2016.10.004>
- Seidenberg, M. S., & Elman, J. L. (1999). Do Infants Learn Grammar with Algebra or Statistics? *Science*, 284(5413), 433-433.
<https://doi.org/10.1126/science.284.5413.433f>
- Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2002). Neuroscience. Does grammar start where statistics stop? *Science (New York, N.Y.)*, 298(5593), 553-554.
<https://doi.org/10.1126/science.1078094>
- Seki, Y., Suzuki, K., Osawa, A. M., & Okanoya, K. (2013). Songbirds and humans apply different strategies in a sound sequence discrimination task. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00447>
- Servan-Schreiber, E., & R. Anderson, J. (1990). Learning Artificial Grammars With Competitive Chunking. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 16(4), 592-608. <https://doi.org/10.1037/0278-7393.16.4.592>

- Seyfarth, R. M., & Cheney, D. L. (2014). The evolution of language from social cognition. *Current Opinion in Neurobiology*, 28, 5-9. <https://doi.org/10.1016/j.conb.2014.04.003>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27, 623-656.
- Shima, K., Isoda, M., Mushiake, H., & Tanji, J. (2007). Categorization of behavioural sequences in the prefrontal cortex. *Nature*, 445(7125), 315-318.
<https://doi.org/10.1038/nature05470>
- Shultz, T. R., & Bale, A. C. (2001). Neural Network Simulation of Infant Familiarization to Artificial Sentences: Rule-Like Behavior Without Explicit Rules and Variables. *Infancy*, 2(4), 501-536. https://doi.org/10.1207/S15327078IN0204_07
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 372(1711).
<https://doi.org/10.1098/rstb.2016.0059>
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105-120.
<https://doi.org/10.1016/j.jml.2015.02.001>
- Soderstrom, M., White, K. S., Conwell, E., & Morgan, J. L. (2007). Receptive Grammatical Knowledge of Familiar Content Words and Inflection in 16-Month-Olds. *Infancy*, 12(1), 1-29. <https://doi.org/10.1111/j.1532-7078.2007.tb00231.x>
- Sonnweber, R., Ravignani, A., & Fitch, W. T. (2015). Non-adjacent visual dependency learning in chimpanzees. *Animal Cognition*, 18(3), 733-745.
<https://doi.org/10.1007/s10071-015-0840-x>
- Spierings, M. J., & ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27), E3977-3984.
<https://doi.org/10.1073/pnas.1600483113>
- Stadler, M. A., & Frensch, P. (Éd.). (1998). *Handbook of Implicit Learning*. Sage Publications.

- Steedman, M. J. (1984). A Generative Grammar for Jazz Chord Sequences. *Music Perception: An Interdisciplinary Journal*, 2(1), 52-77.
<https://doi.org/10.2307/40285282>
- Steiper, M. E., Young, N. M., & Sukarna, T. Y. (2004). Genomic data support the hominoid slowdown and an Early Oligocene estimate for the hominoid–cercopithecoid divergence. *Proceedings of the National Academy of Sciences of the United States of America*, 101(49), 17021-17026. <https://doi.org/10.1073/pnas.0407270101>
- Stobbe, N., Westphal-Fitch, G., Aust, U., & Fitch, W. T. (2012). Visual artificial grammar learning: comparative research on humans, kea (*Nestor notabilis*) and pigeons (*Columba livia*). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 367(1598), 1995-2006. <https://doi.org/10.1098/rstb.2012.0096>
- Suge, R., & Okanoya, K. (2010). Perceptual chunking in the self-produced songs of Bengalese finches (*Lonchura striata* var. *domestica*). *Animal Cognition*, 13(3), 515-523. <https://doi.org/10.1007/s10071-009-0302-4>
- Suzuki, R., Buck, J. R., & Tyack, P. L. (2006). Information entropy of humpback whale songs. *The Journal of the Acoustical Society of America*, 119(3), 1849-1866.
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, 7, 10986. <https://doi.org/10.1038/ncomms10986>
- Takahasi, M., Yamada, H., & Okanoya, K. (2010). Statistical and Prosodic Cues for Song Segmentation Learning by Bengalese Finches (*Lonchura striata* var. *domestica*). *Ethology*, 116(6), 481-489. <https://doi.org/10.1111/j.1439-0310.2010.01772.x>
- ten Cate, C. (2017). Assessing the uniqueness of language: Animal grammatical abilities take center stage. *Psychonomic Bulletin & Review*, 24(1), 91-96. <https://doi.org/10.3758/s13423-016-1091-9>
- ten Cate, C., & Ballintijn, M. R. (1996). Dove Coos and Flashed Lights: Interruptibility of « Song » in a Nonsongbird. *Journal of Comparative Psychology*, 110(3), 267-275. <https://doi.org/10.1037/0735-7036.110.3.267>
- ten Cate, C., Lachlan, R., & Zuidema, W. (2013). *Analyzing the structure of bird vocalizations and language: finding common ground*.

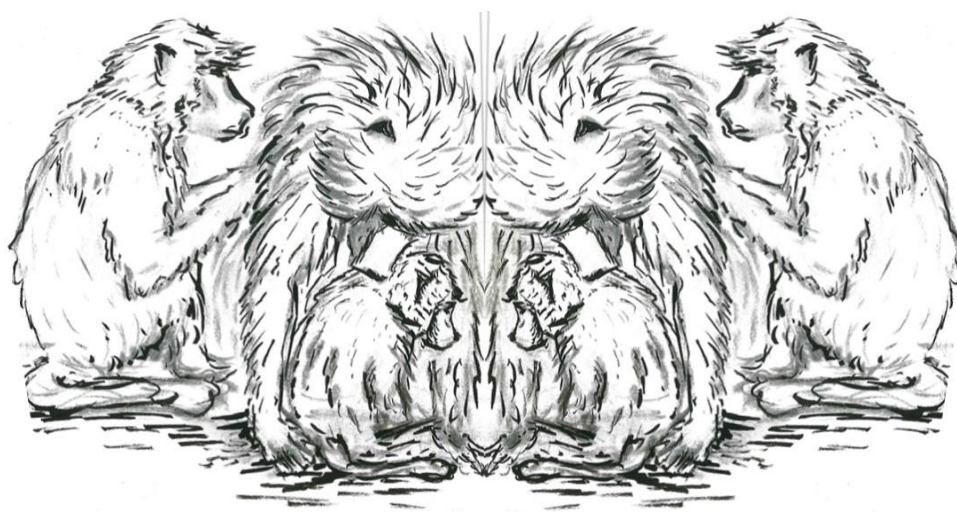
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1598), 1984-1994.
<https://doi.org/10.1098/rstb.2012.0055>
- Terrace, H. S., Petitto, L. A., Sanders, R. J., & Bever, T. G. (1979). Can an ape create a sentence? *Science (New York, N.Y.)*, 206(4421), 891-902.
- Todt, D., & Hultsch, H. (1998). How songbirds deal with large amounts of serial information: retrieval rules suggest a hierarchical song memory. *Biological Cybernetics*, 79(6), 487-500. <https://doi.org/10.1007/s004220050498>
- Tomalin, M. (2007). Reconsidering recursion in syntactic theory. *Lingua*, 117(10), 1784-1800. <https://doi.org/10.1016/j.lingua.2006.11.001>
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.
- Toro, J. M., Nespore, M., Mehler, J., & Bonatti, L. L. (2008). Finding words and rules in a speech stream: functional differences between vowels and consonants. *Psychological Science*, 19(2), 137-144. <https://doi.org/10.1111/j.1467-9280.2008.02059.x>
- Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, 67(5), 867-875.
<https://doi.org/10.3758/BF03193539>
- Truswell, R. (2017). Dendrophobia in Bonobo Comprehension of Spoken English. *Mind & Language*, 32(4), 395-415. <https://doi.org/10.1111/mila.12150>
- Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1), 230-265. <https://doi.org/10.1112/plms/s2-42.1.230>
- Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology. General*, 134(4), 552-564.
<https://doi.org/10.1037/0096-3445.134.4.552>
- Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., ... Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: an rTMS study. *Brain Research*, 1224, 69-78. <https://doi.org/10.1016/j.brainres.2008.05.070>

- Uddén, J., Ingvar, M., Hagoort, P., & Petersson, K. M. (2012). Implicit acquisition of grammars with crossed and nested non-adjacent dependencies: investigating the push-down stack model. *Cognitive Science*, *36*(6), 1078-1101.
<https://doi.org/10.1111/j.1551-6709.2012.01235.x>
- Uhrig, L., Dehaene, S., & Jarraya, B. (2014). A hierarchy of responses to auditory regularities in the macaque brain. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(4), 1127-1132. <https://doi.org/10.1523/JNEUROSCI.3165-13.2014>
- van Heijningen, C. A. A., Chen, J., van Laatum, I., van der Hulst, B., & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: which rule? *Animal Cognition*, *16*(2), 165-175. <https://doi.org/10.1007/s10071-012-0559-x>
- van Heijningen, C. A. A. van, Visser, J. de, Zuidema, W., & ten Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences*, pnas.0908113106.
<https://doi.org/10.1073/pnas.0908113106>
- van Heugten, M., & Johnson, E. K. (2010). Linking infants' distributional learning abilities to natural language acquisition. *Journal of Memory and Language*, *63*(2), 197-209.
<https://doi.org/10.1016/j.jml.2010.04.001>
- van Kampen, A., Parmaksiz, G., Van de Vijver, R., & Hoehle, B. (2008). Metrical and statistical cues for word segmentation: the use of vowel harmony and word stress as a cue to word boundaries by 6- and 9-month-old Turkish learners.
- Versace, E., Spierings, M. J., Caffini, M., ten Cate, C., & Vallortigara, G. (2017). Spontaneous generalization of abstract multimodal patterns in young domestic chicks. *Animal Cognition*, *20*(3), 521-529. <https://doi.org/10.1007/s10071-017-1079-5>
- Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammar. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*(2), 328-344. <https://doi.org/10.1037/0278-7393.18.2.328>
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental Science*, *7*(3), 270-276.

- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159-164.
<https://doi.org/10.1111/j.1467-7687.2007.00549.x>
- Vries, M. de, Christiansen, M., & Petersson, K. M. (2011). Learning Recursion: Multiple Nested and Crossed Dependencies. *BIOLINGUISTICS*, *5*(1-2), 010-035.
- Vuong, L. C., Meyer, A. S., & Christiansen, M. H. (2016). Concurrent Statistical Learning of Adjacent and Nonadjacent Dependencies. *Language Learning*, *66*(1), 8-30.
<https://doi.org/10.1111/lang.12137>
- Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of Numerical and Sequential Patterns in Macaque and Human Brains. *Current Biology*, *25*(15), 1966-1974. <https://doi.org/10.1016/j.cub.2015.06.035>
- Wang, M. D. (1970). The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior*, *9*(4), 398-404.
[https://doi.org/10.1016/S0022-5371\(70\)80079-2](https://doi.org/10.1016/S0022-5371(70)80079-2)
- Watumull, J., Hauser, M. D., Roberts, I. G., & Hornstein, N. (2014). On recursion. *Language Sciences*, *4*, 1017. <https://doi.org/10.3389/fpsyg.2013.01017>
- Wilson, B., Kikuchi, Y., Sun, L., Hunter, D., Dick, F., Smith, K., ... Petkov, C. I. (2015). Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. *Nature Communications*, *6*, 8901.
<https://doi.org/10.1038/ncomms9901>
- Wilson, B., Marslen-Wilson, W. D., & Petkov, C. I. (2017). Conserved Sequence Processing in Primate Frontal Cortex. *Trends in Neurosciences*, *0*(0).
<https://doi.org/10.1016/j.tins.2016.11.004>
- Wilson, B., Slater, H., Kikuchi, Y., Milne, A. E., Marslen-Wilson, W. D., Smith, K., & Petkov, C. I. (2013). Auditory Artificial Grammar Learning in Macaque and Marmoset Monkeys. *The Journal of Neuroscience*, *33*(48), 18825-18835.
<https://doi.org/10.1523/JNEUROSCI.2414-13.2013>
- Wilson, B., Smith, K., & Petkov, C. I. (2015). Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies. *The European Journal of Neuroscience*, *41*(5), 568-578. <https://doi.org/10.1111/ejn.12834>

- Wright, A. A., Rivera, J. J., Katz, J. S., & Bachevalier, J. (2003). Abstract-concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology. Animal Behavior Processes*, 29(3), 184-198.
- Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences of the United States of America*, 110(16), 6324-6327.
<https://doi.org/10.1073/pnas.1216803110>
- Zaccarella, E., & Friederici, A. D. (2016). The neurobiological nature of syntactic hierarchies. *Neuroscience and Biobehavioral Reviews*.
<https://doi.org/10.1016/j.neubiorev.2016.07.038>
- Zimmerer, V. C., Cowell, P. E., & Varley, R. A. (2011). Individual behavior in learning of an artificial grammar. *Memory & Cognition*, 39(3), 491-501.
<https://doi.org/10.3758/s13421-010-0039-y>
- Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour*, 63(2), 293-299. <https://doi.org/10.1006/anbe.2001.1914>

ANNEXES



I. Etude IV: Supplementary Information

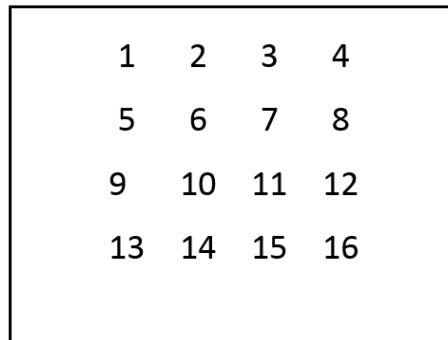


Figure S1 Locations on the touchscreen. Note that letters are given for illustrative purposes but were not displayed during the experiment.

Table S2 Lists of all the sequences used in this study. Numbers denote different locations on the touchscreen as shown in Figure S1.

Set	Mirror	Copy	Set	Mirror	Copy
	4.1.1.4	1.4.1.4		15.2.1.1.2.15	1.2.15.1.2.15
	11.1.1.11	1.11.1.11	6-target baseline sequences (Test 1)	13.10.2.2.10.13	2.10.13.2.10.13
	12.1.1.12	1.12.1.12		1.15.3.3.15.1	3.15.1.3.15.1
	15.1.1.15	1.15.1.15		3.1.10.10.1.3	10.1.3.10.1.3
	1.2.2.1	2.1.2.1		10.3.13.13.3.10	13.3.10.13.3.10
	7.2.2.7	2.7.2.7		2.13.15.15.13.2	15.13.2.15.13.2
	8.2.2.8	2.8.2.8			
	9.2.2.9	2.9.2.9		11.3.1.1.3.11	1.3.11.1.3.11
	1.3.3.1	3.1.3.1		6.4.1.1.4.6	1.4.6.1.4.6
	4.3.3.4	3.4.3.4		4.6.1.1.6.4	1.6.4.1.6.4
	5.3.3.5	3.5.3.5		12.10.1.1.10.12	1.10.12.1.10.12
	11.3.3.11	3.11.3.11		3.11.1.1.11.3	1.11.3.1.11.3
	6.4.4.6	4.6.4.6		10.12.1.1.12.10	1.12.10.1.12.10
4-target baseline sequences	11.4.4.11	4.11.4.11		12.8.2.2.8.12	2.8.12.2.8.12
	14.4.4.14	4.14.4.14		8.12.2.2.12.8	2.12.8.2.12.8
	16.4.4.16	4.16.4.16		11.1.3.3.1.11	3.1.11.3.1.11
	4.5.5.4	5.4.5.4	6-target baseline sequences (Tests 2-3)	1.11.3.3.11.1	3.11.1.3.11.1
	10.5.5.10	5.10.5.10		6.1.4.4.1.6	4.1.6.4.1.6
	13.5.5.13	5.13.5.13		1.6.4.4.6.1	4.6.1.4.6.1
	16.5.5.16	5.16.5.16		14.9.4.4.9.14	4.9.14.4.9.14
	1.6.6.1	6.1.6.1		16.11.4.4.11.16	4.11.16.4.11.16
	3.6.6.3	6.3.6.3		9.14.4.4.14.9	4.14.9.4.14.9
	11.6.6.11	6.11.6.11		11.16.4.4.16.11	4.16.11.4.16.11
	16.6.6.16	6.16.6.16		16.7.5.5.7.16	5.7.16.5.7.16
	5.7.7.5	7.5.7.5		7.16.5.5.16.7	5.16.7.5.16.7
	6.7.7.6	7.6.7.6		4.1.6.6.1.4	6.1.4.6.1.4
	10.7.7.10	7.10.7.10		1.4.6.6.4.1	6.4.1.6.4.1
	14.7.7.14	7.14.7.14		16.5.7.7.5.16	7.5.16.7.5.16
	5.8.8.5	8.5.8.5		5.16.7.7.16.5	7.16.5.7.16.5
9.8.8.9	8.9.8.9		12.2.8.8.2.12	8.2.12.8.2.12	

12.8.8.12	8.12.8.12	2.12.8.8.12.2	8.12.2.8.12.2
15.8.8.15	8.15.8.15	15.13.8.8.13.15	8.13.15.8.13.15
4.9.9.4	9.4.9.4	13.15.8.8.15.13	8.15.13.8.15.13
7.9.9.7	9.7.9.7	14.4.9.9.4.14	9.4.14.9.4.14
12.9.9.12	9.12.9.12	4.14.9.9.14.4	9.14.4.9.14.4
13.9.9.13	9.13.9.13	12.1.10.10.1.12	10.1.12.10.1.12
1.10.10.1	10.1.10.1	1.12.10.10.12.1	10.12.1.10.12.1
8.10.10.8	10.8.10.8	3.1.11.11.1.3	11.1.3.11.1.3
12.10.10.12	10.12.10.12	1.3.11.11.3.1	11.3.1.11.3.1
14.10.10.14	10.14.10.14	16.4.11.11.4.16	11.4.16.11.4.16
5.11.11.5	11.5.11.5	4.16.11.11.16.4	11.16.4.11.16.4
10.11.11.10	11.10.11.10	10.1.12.12.1.10	12.1.10.12.1.10
14.11.11.14	11.14.11.14	8.2.12.12.2.8	12.2.8.12.2.8
16.11.11.16	11.16.11.16	2.8.12.12.8.2	12.8.2.12.8.2
2.12.12.2	12.2.12.2	1.10.12.12.10.1	12.10.1.12.10.1
3.12.12.3	12.3.12.3	15.8.13.13.8.15	13.8.15.13.8.15
6.12.12.6	12.6.12.6	8.15.13.13.15.8	13.15.8.13.15.8
13.12.12.13	12.13.12.13	9.4.14.14.4.9	14.4.9.14.4.9
2.13.13.2	13.2.13.2	4.9.14.14.9.4	14.9.4.14.9.4
3.13.13.3	13.3.13.3	13.8.15.15.8.13	15.8.13.15.8.13
8.13.13.8	13.8.13.8	8.13.15.15.13.8	15.13.8.15.13.8
10.13.13.10	13.10.13.10	11.4.16.16.4.11	16.4.11.16.4.11
2.14.14.2	14.2.14.2	7.5.16.16.5.7	16.5.7.16.5.7
8.14.14.8	14.8.14.8	5.7.16.16.7.5	16.7.5.16.7.5
9.14.14.9	14.9.14.9	4.11.16.16.11.4	16.11.4.16.11.4
15.14.14.15	14.15.14.15		
2.15.15.2	15.2.15.2		
6.15.15.6	15.6.15.6		
7.15.15.7	15.7.15.7		
13.15.15.13	15.13.15.13		
3.16.16.3	16.3.16.3		
7.16.16.7	16.7.16.7		
9.16.16.9	16.9.16.9		
15.16.16.15	16.15.16.15		

Set	Mirror		Copy	
	Consistent	Inconsistent	Consistent	Inconsistent
	1.14.14.1	1.14.1.14	1.14.1.14	1.14.14.1
	1.9.9.1	1.9.1.9	1.9.1.9	1.9.9.1
	10.3.3.10	10.3.10.3	10.3.10.3	10.3.3.10
	10.6.6.10	10.6.10.6	10.6.10.6	10.6.6.10
	11.12.12.11	11.12.11.12	11.12.11.12	11.12.12.11
	11.13.13.11	11.13.11.13	11.13.11.13	11.13.13.11
	12.11.11.12	12.11.12.11	12.11.12.11	12.11.11.12
	12.15.15.12	12.15.12.15	12.15.12.15	12.15.15.12
	13.11.11.13	13.11.13.11	13.11.13.11	13.11.11.13
	13.7.7.13	13.7.13.7	13.7.13.7	13.7.7.13
	14.1.1.14	14.1.14.1	14.1.14.1	14.1.1.14
	14.16.16.14	14.16.14.16	14.16.14.16	14.16.16.14
	15.12.12.15	15.12.15.12	15.12.15.12	15.12.12.15
	15.4.4.15	15.4.15.4	15.4.15.4	15.4.4.15
	16.14.14.16	16.14.16.14	16.14.16.14	16.14.14.16
Pre-test	16.2.2.16	16.2.16.2	16.2.16.2	16.2.2.16
1	2.16.16.2	2.16.2.16	2.16.2.16	2.16.16.2
	2.5.5.2	2.5.2.5	2.5.2.5	2.5.5.2
	3.10.10.3	3.10.3.10	3.10.3.10	3.10.10.3
	3.7.7.3	3.7.3.7	3.7.3.7	3.7.7.3
	4.15.15.4	4.15.4.15	4.15.4.15	4.15.15.4
	4.8.8.4	4.8.4.8	4.8.4.8	4.8.8.4
	5.2.2.5	5.2.5.2	5.2.5.2	5.2.2.5
	5.9.9.5	5.9.5.9	5.9.5.9	5.9.9.5
	6.10.10.6	6.10.6.10	6.10.6.10	6.10.10.6
	6.8.8.6	6.8.6.8	6.8.6.8	6.8.8.6
	7.13.13.7	7.13.7.13	7.13.7.13	7.13.13.7
	7.3.3.7	7.3.7.3	7.3.7.3	7.3.3.7
	8.4.4.8	8.4.8.4	8.4.8.4	8.4.4.8
	8.6.6.8	8.6.8.6	8.6.8.6	8.6.6.8
	9.1.1.9	9.1.9.1	9.1.9.1	9.1.1.9
	9.5.5.9	9.5.9.5	9.5.9.5	9.5.5.9

Set	Mirror		Copy	
	Consistent	Inconsistent	Consistent	Inconsistent
Pre-test 2	1.14.14.1	1.14.14.16	1.14.1.14	1.14.1.9
	1.9.9.1	1.9.9.5	1.9.1.9	1.9.1.14
	10.3.3.10	10.3.3.7	10.3.10.3	10.3.10.6
	10.6.6.10	10.6.6.8	10.6.10.6	10.6.10.3
	11.12.12.11	11.12.12.15	11.12.11.12	11.12.11.13
	11.13.13.11	11.13.13.7	11.13.11.13	11.13.11.12
	12.11.11.12	12.11.11.13	12.11.12.11	12.11.12.15
	12.15.15.12	12.15.15.4	12.15.12.15	12.15.12.11
	13.11.11.13	13.11.11.12	13.11.13.11	13.11.13.7
	13.7.7.13	13.7.7.3	13.7.13.7	13.7.13.11
	14.1.1.14	14.1.1.9	14.1.14.1	14.1.14.16
	14.16.16.14	14.16.16.2	14.16.14.16	14.16.14.1
	15.12.12.15	15.12.12.11	15.12.15.12	15.12.15.4
	15.4.4.15	15.4.4.8	15.4.15.4	15.4.15.12
	16.14.14.16	16.14.14.1	16.14.16.14	16.14.16.2
	16.2.2.16	16.2.2.5	16.2.16.2	16.2.16.14
	2.16.16.2	2.16.16.14	2.16.2.16	2.16.2.5
	2.5.5.2	2.5.5.9	2.5.2.5	2.5.2.16
	3.10.10.3	3.10.10.6	3.10.3.10	3.10.3.7
	3.7.7.3	3.7.7.13	3.7.3.7	3.7.3.10
	4.15.15.4	4.15.15.12	4.15.4.15	4.15.4.8
	4.8.8.4	4.8.8.6	4.8.4.8	4.8.4.15
	5.2.2.5	5.2.2.16	5.2.5.2	5.2.5.9
	5.9.9.5	5.9.9.1	5.9.5.9	5.9.5.2
	6.10.10.6	6.10.10.3	6.10.6.10	6.10.6.8
	6.8.8.6	6.8.8.4	6.8.6.8	6.8.6.10
	7.13.13.7	7.13.13.11	7.13.7.13	7.13.7.3
	7.3.3.7	7.3.3.10	7.3.7.3	7.3.7.13
	8.4.4.8	8.4.4.15	8.4.8.4	8.4.8.6
	8.6.6.8	8.6.6.10	8.6.8.6	8.6.8.4
9.1.1.9	9.1.1.14	9.1.9.1	9.1.9.5	
9.5.5.9	9.5.5.2	9.5.9.5	9.5.9.1	

Set	Mirror		Copy	
	Consistent	Inconsistent	Consistent	Inconsistent
Test 1	11.14.5.5.14.11	11.14.5.5.11.14	11.12.8.11.12.8	11.12.8.11.8.12
	11.7.6.6.7.11	11.7.6.6.11.7	11.8.12.11.8.12	11.8.12.11.12.8
	12.16.14.14.16.12	12.16.14.14.12.16	12.5.6.12.5.6	12.5.6.12.6.5
	12.8.11.11.8.12	12.8.11.11.12.8	12.6.5.12.6.5	12.6.5.12.5.6
	14.11.5.5.11.14	14.11.5.5.14.11	14.12.16.14.12.16	14.12.16.14.16.12
	14.7.9.9.7.14	14.7.9.9.14.7	14.16.12.14.16.12	14.16.12.14.12.16
	16.12.14.14.12.16	16.12.14.14.16.12	16.4.9.16.4.9	16.4.9.16.9.4
	16.5.8.8.5.16	16.5.8.8.16.5	16.9.4.16.9.4	16.9.4.16.4.9
	4.9.16.16.9.4	4.9.16.16.4.9	4.6.8.4.6.8	4.6.8.4.8.6
	4.9.7.7.9.4	4.9.7.7.4.9	4.8.6.4.8.6	4.8.6.4.6.8
	5.16.8.8.16.5	5.16.8.8.5.16	5.11.14.5.11.14	5.11.14.5.14.11
	5.6.12.12.6.5	5.6.12.12.5.6	5.14.11.5.14.11	5.14.11.5.11.14
	6.5.12.12.5.6	6.5.12.12.6.5	6.11.7.6.11.7	6.11.7.6.7.11
	6.8.4.4.8.6	6.8.4.4.6.8	6.7.11.6.7.11	6.7.11.6.11.7
	7.11.6.6.11.7	7.11.6.6.7.11	7.4.9.7.4.9	7.4.9.7.9.4
	7.14.9.9.14.7	7.14.9.9.7.14	7.9.4.7.9.4	7.9.4.7.4.9
	8.12.11.11.12.8	8.12.11.11.8.12	8.16.5.8.16.5	8.16.5.8.5.16
	8.6.4.4.6.8	8.6.4.4.8.6	8.5.16.8.5.16	8.5.16.8.16.5
	9.4.16.16.4.9	9.4.16.16.9.4	9.14.7.9.14.7	9.14.7.9.7.14
	9.4.7.7.4.9	9.4.7.7.9.4	9.7.14.9.7.14	9.7.14.9.14.7

Set	Mirror		Copy	
	Consistent	Inconsistent	Consistent	Inconsistent
Test 2	10.13.5.5.13.10	10.13.5.5.10.13	10.14.8.10.14.8	10.14.8.10.8.14
	13.10.5.5.10.13	13.10.5.5.13.10	10.8.14.10.8.14	10.8.14.10.14.8
	13.3.12.12.3.13	13.3.12.12.13.3	12.13.3.12.13.3	12.13.3.12.3.13
	14.8.10.10.8.14	14.8.10.10.14.8	12.3.13.12.3.13	12.3.13.12.13.3
	15.2.14.14.2.15	15.2.14.14.15.2	14.15.2.14.15.2	14.15.2.14.2.15
	16.3.6.6.3.16	16.3.6.6.16.3	14.2.15.14.2.15	14.2.15.14.15.2
	2.15.14.14.15.2	2.15.14.14.2.15	15.6.7.15.6.7	15.6.7.15.7.6
	3.13.12.12.13.3	3.13.12.12.3.13	15.7.6.15.7.6	15.7.6.15.6.7
	3.16.6.6.16.3	3.16.6.6.3.16	2.7.9.2.7.9	2.7.9.2.9.7
	4.5.3.3.5.4	4.5.3.3.4.5	2.9.7.2.9.7	2.9.7.2.7.9
	5.4.3.3.4.5	5.4.3.3.5.4	3.4.5.3.4.5	3.4.5.3.5.4
	6.7.15.15.7.6	6.7.15.15.6.7	3.5.4.3.5.4	3.5.4.3.4.5
	7.6.15.15.6.7	7.6.15.15.7.6	5.10.13.5.10.13	5.10.13.5.13.10
	7.9.2.2.9.7	7.9.2.2.7.9	5.13.10.5.13.10	5.13.10.5.10.13
	8.14.10.10.14.8	8.14.10.10.8.14	6.16.3.6.16.3	6.16.3.6.3.16
	9.7.2.2.7.9	9.7.2.2.9.7	6.3.16.6.3.16	6.3.16.6.16.3
Test 3 & Control	1.13.7.7.13.1	1.13.7.7.1.13	1.5.9.1.5.9	1.5.9.1.9.5
	1.16.8.8.16.1	1.16.8.8.1.16	1.9.5.1.9.5	1.9.5.1.5.9
	10.4.15.15.4.10	10.4.15.15.10.4	10.16.2.10.16.2	10.16.2.10.2.16
	12.15.5.5.15.12	12.15.5.5.12.15	10.2.16.10.2.16	10.2.16.10.16.2
	13.1.7.7.1.13	13.1.7.7.13.1	10.3.9.10.3.9	10.3.9.10.9.3
	15.12.5.5.12.15	15.12.5.5.15.12	10.9.3.10.9.3	10.9.3.10.3.9
	15.9.11.11.9.15	15.9.11.11.15.9	11.15.9.11.15.9	11.15.9.11.9.15
	16.1.8.8.1.16	16.1.8.8.16.1	11.9.15.11.9.15	11.9.15.11.15.9
	16.2.10.10.2.16	16.2.10.10.16.2	15.10.4.15.10.4	15.10.4.15.4.10
	2.16.10.10.16.2	2.16.10.10.2.16	15.4.10.15.4.10	15.4.10.15.10.4
	3.9.10.10.9.3	3.9.10.10.3.9	5.12.15.5.12.15	5.12.15.5.15.12
	4.10.15.15.10.4	4.10.15.15.4.10	5.15.12.5.15.12	5.15.12.5.12.15
	5.9.1.1.9.5	5.9.1.1.5.9	7.1.13.7.1.13	7.1.13.7.13.1
	9.15.11.11.15.9	9.15.11.11.9.15	7.13.1.7.13.1	7.13.1.7.1.13
9.3.10.10.3.9	9.3.10.10.9.3	8.1.16.8.1.16	8.1.16.8.16.1	
9.5.1.1.5.9	9.5.1.1.9.5	8.16.1.8.16.1	8.16.1.8.1.16	

II. Etude sur la Prévalence des Informations Ordinales et Relationnelles : Données Comparatives chez le Babouin et l'Humain

Fagot, J., Malassis, R., & Medam, T. (2018). The processing of positional information in a two-item sequence limits the emergence of symmetry in baboons (*Papio papio*), but not in humans (*Homo sapiens*). *Learning & Behavior*, 46(1), 67-78. <https://doi.org/10.3758/s13420-017-0290-1>

Cette étude est issue d'une collaboration au sein de l'équipe. La question abordée ne traitait pas directement des régularités non-adjacentes et l'extraction de structure, mais portait sur la prévalence de certains modes d'encodage des séquences sur d'autres, lorsque plusieurs stratégies sont disponibles. Nous avons utilisé l'adaptation de la tâche SRT visuelle que j'ai développé pour la deuxième étude de ma thèse.

Nous avons cherché à quantifier l'importance relative pour les babouins des propriétés ordinales (e.g. A apparaît en 1^{er}, B en 2^{ème}) et relationnelles (e.g. A est associé à B) caractérisant un set de séquences. Après une période d'exposition à trois séquences constituées chacune de deux formes visuelles associées (notées AB, CD, EF), nous avons présenté aux babouins des réarrangements de ces séquences (BA, ED, FC). Dans ces séquences test, les informations ordinales et relationnelles étaient manipulées de manière orthogonale :

Exposition	Test	Relation	Position
A-B	B-A	+	-
C-D E-F	E-D	-	+
C-D E-F	F-C	-	-

L'examen des temps de réponse des babouins sur le deuxième élément de ces séquences test a indiqué que les babouins réagissaient davantage aux modifications de la *position* des stimuli dans la séquence qu'aux modifications des *relations* entre ces stimuli. La même expérience menée chez l'humain dans des conditions similaires d'apprentissage révèle que ceux-ci réagissent de manière similaire aux deux types de modifications. Ces résultats suggèrent que lorsque des séquences de stimuli sont caractérisées à la fois par des régularités ordinales et relationnelles, les babouins donnent priorité à ces premières.

The processing of positional information in a two-item sequence limits the emergence of symmetry in baboons (*Papio papio*), but not in humans (*Homo sapiens*)

Joël Fagot^{1,2} · Raphaele Malassis^{1,2} · Tiphaine Medam^{1,2}

© Psychonomic Society, Inc. 2017

Abstract When trained to associate Stimulus A to Stimulus B, humans can derive the untrained symmetrical B to A relation while nonhuman animals have much more difficulties. Urcuioli (2008, *Journal of the Experimental Analysis of Behavior*, 90, 257–282; 2015, *Conducta*, 3, 4–25) proposed that the apparent difficulty of animals in symmetry testing reflects their double encoding of the information on the stimuli (identity and relation) and their positional (i.e., spatial and temporal/ordinal) characteristics. This comparative study tested the emergence of symmetry in humans and baboons in a task in which the position of the stimuli was manipulated independently of their relation. Humans and baboons initially learned to associate pairs of visual shapes on a touch screen in a specific order. Three pairs of (A-B, C-D, and E-F) stimuli were used in training. After training, the two species were tested with the B-A, F-C, and E-D pairs. The B-A pairs preserved the association initially learned with A-B but reversed the positional information relative to training. The F-C pair neither preserved the association nor the positional information of the training pairs, and positional information were the only cues preserved in the E-D pair. Humans showed a response time advantage for B-A, suggesting symmetry, but also for E-D, suggesting that they also process positional information. In baboons, the advantage was found only for E-D, suggesting that they only process positional information. These

results confirm that the processing of stimulus pairs differ between nonhuman animals to humans.

Keywords Associative learning · Equivalence classes · Language · Categorization

Derived relational responding is a cognitive phenomenon in which trained relations between arbitrary stimuli result in the spontaneous emergence of untrained relations. This cognitive phenomenon is best demonstrated in humans by experiments on stimulus equivalence (Sidman & Tailby, 1982). In such experiments, the participants must typically learn by trial and error that the visual sample Stimulus A must be associated in a conditional matching-to-sample task to the comparison Stimulus B. They must also learn that the same B stimulus now used as sample in the matching task must be associated to the comparison Stimulus C. After such (A-B, B-C) training, the subjects demonstrate stimulus equivalence when they spontaneously exhibit the relation of reflexivity, symmetry, and transitivity. Reflexivity implies the ability to spontaneously associate each stimulus to itself, for instance, to indicate that A goes with A. The relation of symmetry implies that any given association learned during the training phase can be reversed. In other words, if the subject has learned that A goes to B, he or she should be able to indicate without training that B also goes to A. The relation of transitivity involves that the stimuli that have common associates go together. In other words, after being trained with A-B and B-C, the subject showing transitivity should be able to match A to C and C to A. Stimulus equivalence, as defined by the ability to process all these three relations, was initially theorized and demonstrated by Sidman and collaborators using this kind of matching-to-sample protocols (Sidman, 1971; Sidman & Tailby, 1982), and the fact that our species has this ability

✉ Joël Fagot
joel.fagot@univ-amu.fr

¹ Laboratory of Cognitive Psychology, Aix Marseille University, CNRS, Marseille, France

² Brain and Language Research Institute, Aix Marseille University, CNRS, Aix-en-Provence, France

has been largely confirmed since their initial studies (e.g., Adams, Fields, & Verhave, 1993; Carp & Petursdottir, 2015; Devany, Hayes, & Nelson, 1986). One reason for which stimulus equivalence is interesting is that it appears to be closely linked with human language. First, our flexible use of words implies a bidirectional (i.e., symmetrical) relation between the word and its referent: the word refers to the object (e.g., in requesting contexts) as the object evokes the word (e.g., in naming contexts). Second, stimulus equivalence emerges at approximately 2 years of age (e.g., Boelens, van den Broek, & van Klarenbosch, 2000; Lipkens, Hayes, & Hayes, 1993) and is thus concurrent with language learning. Third, children with language disabilities or other clinical groups with linguistic deficiencies (e.g., children with autism) often express difficulties in tasks of stimulus equivalence (Devany et al., 1986; Eikeseth & Smith, 1992; O'Donnell & Saunders, 2003; but see Carr, Wilkinson, Blackman, & McIlvane, 2000). The apparent relations between language and stimulus equivalence call for comparative studies on stimulus equivalence in non-linguistic species.

From a comparative point of view, stimulus equivalence has been an elusive phenomenon when tested in nonhuman animals, such as pigeons (e.g., R. Lipkens, Kop, & Matthijs, 1988) and nonhuman primates (e.g., rhesus monkeys and baboons; Sidman, Rauzin, et al., 1982). Positive results have been repeatedly reported regarding reflexivity (Barros, Galvão, & McIlvane, 2002; Sweeney & Urcuioli, 2010) and transitivity (D'Amato, Salmon, Loukas, & Tomie, 1985; Kuno, Kitadate, & Iwamoto, 1994; Zentall, Clement, & Weaver, 2003), but the large majority of the studies have reported negative results regarding the symmetry relation. Lionello-DeNolf (2009) reviewed 24 investigations of symmetry in nonhuman animals and concluded that only three of these studies found clear evidence of symmetry in their subjects (pigeons: Frank & Wasserman, 2005; sea lions: Kastak, Schusterman, & Kastak, 2001; Schusterman & Kastak, 1993). Since this review has been published, a few additional papers (e.g., Urcuioli, 2015) have confirmed that symmetry is in the scope of pigeons. Nevertheless, the general picture of the existing literature remains that symmetry seems an exceptionally difficult feat in nonhuman species. The main goal of our research was to identify which factor may best explain the general difficulties expressed by nonhuman animals during symmetry testing.

Urcuioli (2008, 2015) proposed that the difficulty experienced by pigeons in symmetry testing would result from their mode of encoding of two types of information that are (1) the information on the identity of the nominal stimuli and their relation and (2) the information on their position in the spatial and temporal dimensions. According to Urcuioli, a Stimulus A appearing as sample in a fixed spatial position #1 (e.g., typically, the center response key) would be encoded as Stimulus A in temporal/ordinal position #1 at spatial position

#1, while a Stimulus B appearing as a comparison in a fixed spatial position #2 (e.g., left or right response key) would be encoded as Stimulus B in temporal/ordinal position #2 at spatial location #2. This mode of encoding would prevent generalization in symmetry trials because Stimulus A in temporal/ordinal position #1 and spatial position #1 would be treated as functionally different from the same stimulus presented in the temporal/ordinal position #2 and spatial position #2.

The idea that nonhuman animals focus too strongly their attention on positional information during the learning of associations between pairs of items is mostly supported by Frank and Wasserman's (2005) and Urcuioli's (2008, Experiment 3) studies in pigeons. These studies had two interesting distinguishing features that may explain their positive results in symmetry trials. First, they both used a successive matching-to-sample task rather than the two-alternative matching tasks (e.g., Lipkens et al., 1988). In their successive matching task, the sample and comparison stimuli are shown on the same location on the screen, and the pigeons have to peck (go response) or not (no-go response) the comparison stimulus depending on the identity of the sample. This aspect of their task contrasts strongly with the two-alternative matching-to-sample task in which the sample and comparisons are shown on different locations on the displays (see Lionello-DeNolf, 2009, for a review of negative findings obtained in this task; see also Urcuioli, 2008, Experiments 1 and 2). Second, both Frank and Wasserman (2005) and Urcuioli (2008, Experiment 3) mixed identity (A-A, B-B) and symbolic matching trials (A-B) during the training phase. One consequence of this intermixing of the identity and symbolic matching trials is that each stimulus has also been presented in the first or second temporal/ordinal position during training. This counterbalancing might have reduced the importance given to temporal/ordinal information in the pigeons' mode of encoding, allowing a stronger focus on relational information at the origin of a successful performance in symmetry trials. Said differently, the two most convincing lines of evidence of symmetry in nonhuman animals were obtained with a test design that minimized the possibility to encode positional (spatial and temporal/ordinal) information in pigeons. This hypothesis may also explain why symmetry training, involving exposure to the reversal of the initially learned associations, seems to promote generalized symmetry (e.g., Pepperberg & Gordon, 2005; Schusterman & Kastak, 1993).

Our study aimed at testing further the general idea that nonhuman animals encode in priority the positional information to the detriment of the relational information in A-B types of sequences, and that this mode of encoding is the main source of failures in symmetry tests involving B-A types of trials. We developed for that purpose a novel prediction task, inspired from the serial response time task (Misyak, Christiansen, & Bruce Tomblin, 2010; Nissen & Bullemer, 1987), which was given to both humans (*Homo sapiens*) and

Guinea baboons (*Papio papio*). The main advantage of this procedure is to offer fine-grained measures of participants' expectations on critical trials rather than merely reflecting the result of a decision process in a binary choice task, as in the aforementioned studies. Our task required participants to learn in a training phase the association between items in three pairs of shapes. A test phase then presented some novel pairings of the stimuli used during the training phase. One test condition preserved positional information but disrupted relational information. One other test condition preserved relational information but disrupted positional information. A last test condition disrupted both kinds of information. We predicted that the baboons will demonstrate an advantage for the stimulus pairs preserving positional information, while humans would demonstrate an advantage for the stimulus pair preserving the relations. To our knowledge, this study is the first to extend Urcuioli's line of research with nonhuman primate species.

Method

Participants

Data were collected on a total of 14 Guinea baboons (*Papio papio*; age range: 4–12 years; seven females) living in a larger social group of 24 individuals located at the CNRS Station de Primatologie, Rousset-sur-Arc, France. This group of baboons is maintained within a 700 m² outdoor enclosure connected to an indoor area used at night, and two test rooms in which the test systems were presented ad libitum (see below and Fagot & Bonté, 2010, for a more detailed description of their living conditions). The baboons were all marked with a radio-frequency identification (RFID) microchip implanted in each arm. Water was provided ad libitum during the experiment. Animals were not food deprived. They received their regular portion of food (monkey chow and fruits) every day at 5 p.m. All baboons had previously participated to numerous computerized experiments using the ALDM test systems and had already been involved in an unpublished experiment that used the same prediction task as in this study. However, this previous experiment never trained or tested our baboons with symmetrical pairs. Note also that some of our subjects have already been involved in Medam, Marzouki, Montant, and Fagot's (2016) study on the effect of category training on symmetry. The use of a two-alternative matching-to-sample procedure in this study failed to reveal symmetry in these individuals.

The human data were collected on a total of 21 adults (age range: 19–28 years, 16 females) who were recruited at our university and participated for payment. These participants were all unaware of the purpose of the research and have never been involved in any of our research programs.

Apparatus

The baboons had an unrestricted access during the experiment to 10 automated learning devices for monkeys (ALDMs; Fagot & Bonté, 2010; Fagot & Paleressompouille, 2009), which were installed in two trailers directly accessible from their outside enclosure. Each ALDMs consisted of a 70 cm × 70 cm × 80 cm testing chamber with a view port and two hand ports. The view port allowed to see a 19-in. touch screen (1939L Open-Frame Touchmonitor, Elo Touch Solutions) installed in front of the chamber, and the hand port was used to manipulate it. The RFID microchip implanted in the arm of each subject served to automatically identify each baboon once it introduced one of its arms through the hand port. The baboon could use the ALDM test system of its choice, whenever it wanted to use it, because the automatic reading of the subject identity was used to prompt the program to resume the trial list at the appropriate place for that subject. The ALDMs automatically dispensed several grains of dry wheat inside the test chamber once a correct response had been registered. Incorrect responses were not food rewarded but gave rise to a 3-s time-out indicated by a green screen. The experiment was programmed using E-Prime (Version 2.0, Psychology Software Tools, Pittsburgh, PA) software.

The apparatus for humans consisted of a 19-in. computer touch screen (ET1915L, Elo Touch) installed on a table in a quiet room. Correct responses in the task were indicated by a short 1-s black screen; incorrect responses were followed by a 3-s time-out with a green screen.

Stimuli

This experiment used a total of six different yellow visual shapes that served to create three different pairs of stimuli (hereafter referred to as pair A-B, C-D, and E-F). The individual stimuli were shown on a 1024 × 768 pixel background and measured approximately 180 × 180 pixels on the screen. The stimuli were drawn for the purpose of the experiment, with the sole constraint that the discrimination of the items of the set should be easy for the two species.

General design of the task

The two species were tested in an adapted version of the serial reaction time task, in which the subjects had to select and touch two stimuli consecutively. The procedure is illustrated in Fig. 1. Each trial began with a display of two items presented one above the other (locations balanced) on the left side of the screen. These two stimuli were first presented during 200 ms, after which one of them disappeared for a short period of 80 ms. This short disappearance of the stimulus gives the impression that the stimulus blinks once. The stimulus disappearing for 80 ms was the positive stimulus (S+) that

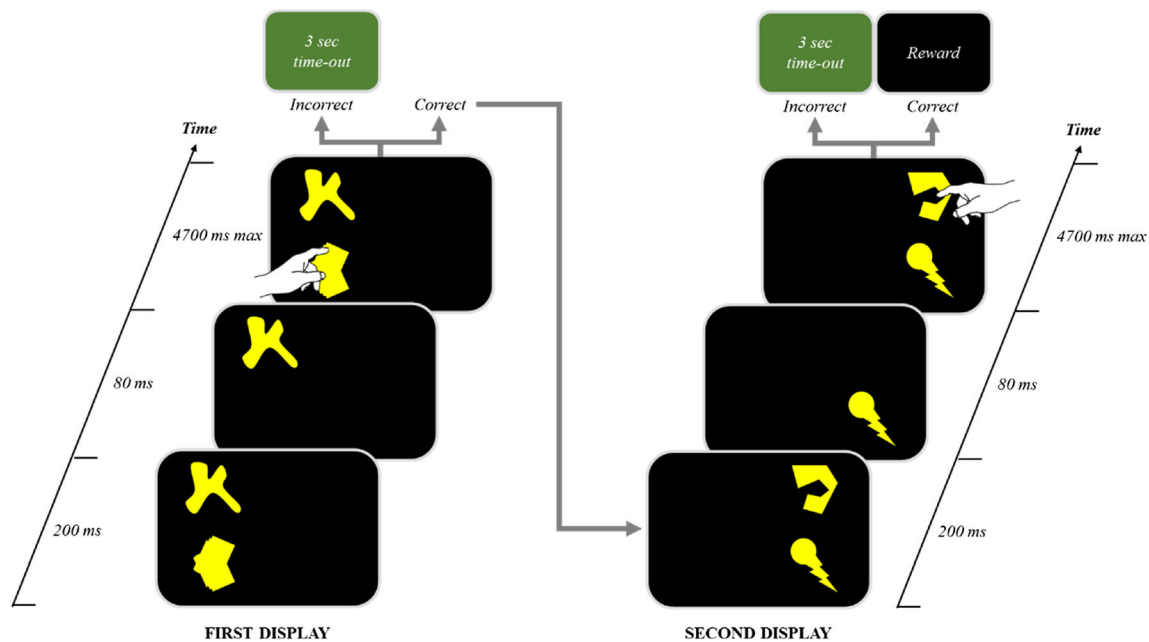


Fig. 1 Illustration of the experimental procedure used in baboons. This same procedure was also used in humans, except that humans did not receive a food reward after each successful response

the subject had to touch in this first display. A total of 4,700 ms was allowed for touching this stimulus. A correct touch of S+ at this stage caused the display to disappear immediately to be replaced by the second display. An incorrect selection of the negative stimulus (S−) triggered a 3-s green screen, and the trial was aborted and scored as incorrect. Once the S+ had been touched, the subject saw the second display containing two other stimuli that were presented one above the other (location balanced) on the right part of the screen. Again, one of these two stimuli disappeared for 80 ms, and the subject then had to touch this S+ stimulus, while avoiding S−. Again, 4,700 ms were allowed for responding at this stage. A correct selection of S+ on this second display triggered a positive feedback (a black screen and food reward for monkey, and black screen only for humans), and an incorrect response produced the 3-s green time-out screen, indicating that an error has been made. Trials in which the participants failed to touch the screen during the time window allotted for responding to the first or second display were systematically aborted. These aborted trials were represented immediately.

The procedure outlined above was used in a first training phase and in the following test phase. The pairs used in each phase are indicated in Table 1. The training phase aimed to teach the subjects that the stimuli were organized in pairs. In the training trials, each first stimulus of a given pair (e.g., Stimulus A in the A-B pair) served as S+ in the first display of the training trials using this pair, and the second stimulus of the pair (i.e., B) served as S+ in the second display. The first and second stimuli of another pair (e.g., C-D or E-F pair) served as S− in the first and the second displays, respectively.

The test sessions presented the same stimuli as in training sessions but used the three stimulus pairings indicated in Table 1. The first one (i.e., B-A) tested for the emergence of symmetry, because this pairing is a reversal of the A-B pairing of the training phase. The remaining two pairings (i.e., F-C and E-D) assessed the effect of stimulus ordering, because the position of these stimuli was either preserved (E-D) or reversed (F-C) in comparison to training. Within a trial, the first and second stimulus of a given pair (e.g., Stimuli B and A, in the B-A pair) served as S+ in the first and second display, respectively, while the first and second stimulus of another pair (e.g., E-D or F-C pair) served as S− in the first and the second displays, respectively. The dependent variables used for statistical analyses were the number of correct responses to the first and the second display of each kind of trial, as well as the response times to S+ in the first (hereafter referred to as

Table 1 Experimental design in the training and test phase

Training Pairs	Test Pairs	Associative cue	Positional cue
A-B	B-A	+	−
C-D	E-D	−	+
E-F	F-C	−	−

Note. Each letter corresponds to one stimulus in the set. For each pair, the first letter indicates the stimuli used for the first display, and the second letter indicate the stimulus used for second display. Bold characters indicate the stimulus of the training pairs under scrutiny in the corresponding test pair. A “+” sign indicates that the associative (third column) and/or positional cues (fourth column) present in the training pair were preserved in the test pairs. A “−” sign indicates that it was not.

RT1) and second display of the trials (RT2). However, because very few errors were recorded in the task, our attention was mostly focused on response times variations that provided fine-grained information on learning processes and on the effect of the different tested factors. With this test design, comparison of the response to the second display of the B-A and F-C test pairs should reveal if the participants process the trained associations as if they are symmetrical. If they do so, then they should express a better accuracy on the second object and/or faster RT2 for the A shape than for the C shape. In addition, a comparison of the response to the second display of the F-C and E-D pairs should also indicate if the subject process the stimulus position, independently of the trained associations. In that case, they should express a better accuracy and/or faster RT2 in D than in C.

Training procedure

During training, the human and baboon participants received blocks of 96 counterbalanced training trials, each containing 32 trials with the A-B, C-D, and E-F stimulus pairs. Each pair was presented in an identical number of trials as the S+ or the S- pair within a block. The ordering of the trials was randomized at the beginning of each block and therefore differed from one block to the next. The baboons received a total of 20 such blocks (1,920 trials) during the training phase. To avoid boredom or inattention to the task, the human participants only received five training blocks (480 trials).

Before the experiment, the baboons received two preliminary training phases, which both used different stimuli from those of the next training and testing phase. These two preliminary training phases were not presented to humans, who rapidly discovered by themselves that they had to touch the stimuli that blinked in the displays. Thus, in a first preliminary phase, the baboons were taught to touch the stimulus among two which disappeared for 80 ms from the screen. Once they reached 80% correct with two individual shapes, we proceeded to the second preliminary training phase. This phase introduced two displays in succession, which were presented following the same procedure as for the training trials (see above), but which used different stimuli from those of the final training and test phase. To obtain a reward, the baboon had to touch the stimulus that blinked once in each display, once again until they reached 80% correct with this design. The stimuli were organized in three pairs in this second preliminary training phases, to teach the subjects that the task involved some associations between stimuli. The training phase involving the three A-B, C-D, and E-F pairs was introduced immediately after these two preliminary training phases. The instruction given to humans prior to the experiments only explained that they had to touch some stimuli on the screen, and that response errors are indicated by a green screen.

Test procedure

During the test phase, the baboons received 10 test blocks of 96 trials each (960 trials). Each block involved 32 trials with each of the three (B-A, F-C, and E-D) test pairs. Human participants received five test blocks (480 trials). Importantly, the stimulus taking the role of the A, B, C, D, E, and F stimulus was counterbalanced across subjects. This counterbalancing of the visual shapes guaranteed that the variations in accuracy or response time are not tied to individual stimuli.

At the end of the test phase, the experimenter asked the human subjects to report verbally their understanding of the task. Thus, all the subjects were first asked, “What were you supposed to do?” Then, if the free report of the subjects did not include the following information, the subject was asked to answer the following questions in the following order: (1) “Which forms had to be touched?” (2) “Did you notice some relationships between and among the stimuli?” (3) “Did you notice any change during the experiment?” and (4) “Did you use verbal names to refer mentally to the stimuli?”

Data filtering

A filtering procedure was applied for the analysis of correct response times to neglect the outlier trials with excessively long response times that may reflect transient inattention to the task. Thus, we first discarded from the data set all the trials in which RT1 and/or RT2 was greater than 3 s. This first layer of filtering removed .11% of the trials for humans and .30% for the baboons. The second filter additionally removed all the trials for which RT2 was greater than two standard deviations from the RT2 mean (computed for each individual and block). This second filtering procedure removed 3.62% of the remaining trials for humans and 4.32% for the baboons. Statistical analyses considered separately the data from the two species because of an uneven number of training and test blocks in humans and baboons.

Results

Humans

The human participants performed consistently high in the task, with an average percentage of correct responses of 96.97% ($SD = 17%$) in the training phase, and 96.4% ($SD = 18.64%$) in the test phase. This performance precludes detailed statistical analyses on accuracy due to potential ceiling effects.

A preliminary inspection of the data indicated that some participants showed an acceleration in RT2 throughout the training phase, while others did not. These different response profiles suggest that only a subset of humans—those for

which the RT2 declined with training—could predict their response to the second item of the pair depending on the identity of the first item. We therefore analyzed the postexperimental verbal reports of the subjects, and split our sample of participants in two groups. The first group ($n = 10$, hereafter named AS+), includes all the participants who have reported verbally during the debriefing that the stimuli were organized in pairs. The second group ($n = 11$, hereafter named the group AS−) includes all the subjects who failed to report any association between or among the stimuli.

Training phase Figure 2 represents the RT1 and RT2 training data for the two groups of participants.

A two-way analysis of variance (ANOVA) was computed for each group, using the stimulus pair (A-B, C-D, and E-F) as well as the training block (1–5) as independent variables, and RT1 as a dependent variable. None of the main effects or the interactions emerged as significant, in either group (all $ps > .05$). We found in particular no reliable effect of stimulus pair for RT1, whatever the group considered: AS+, $F(2, 18) = .17$, $p > .10$, $\eta_p^2 = .02$; AS−, $F(2, 20) = 1.71$, $p > .10$, $\eta_p^2 = .15$. This finding is important because it demonstrates that the response to the first item of each pair did not differ in terms of perceptual or motor constraints.

The same ANOVAs conducted on the RT2 revealed no reliable main effects or interaction when the AS− group was considered (all $ps > .10$). In the case of the AS+ group, neither the effect of stimulus pair nor the interaction were significant, $F(2, 18) = 2.16$, $p > .1$, $\eta_p^2 = .19$; and $F(8, 72) = 1.07$, $p > .1$, $\eta_p^2 = .11$, respectively. However, the main effect of block was statistically significant for this AS+ group, $F(4, 36) = 12.59$, $p < .001$, $\eta_p^2 = .58$. As shown in the top right panel of Fig. 2, this main effect of block corresponds to a reduction of response

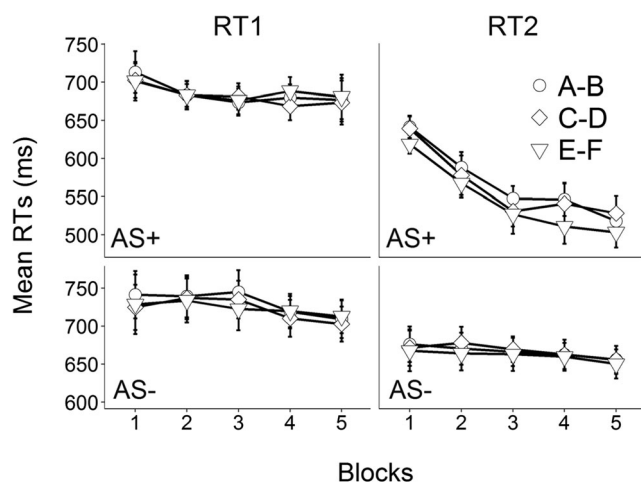


Fig. 2 Mean correct response times on the first (RT1) and the second (RT2) display obtained for each training pair, in the five training blocks presented to the human participants of the AS+ (top) and AS− (bottom) group. Note that the y -axes have different scales in the two groups and display only a subset of the range. Error bars indicate the standard errors centered on each participant mean RT

times with repeated training. One potential explanation of this decrease in RT2 is that it reflects a general improvement of performance for pointing the stimuli on the screen, but this hypothesis is ruled out because our analyses of RT1 showed that they remained stable across the training blocks. These findings, along with the verbal reports of the subjects, therefore suggest that AS+ is the sole group that learned that the stimuli were organized by pairs. For these individuals, the learning of the stimulus pairing that occurred across the training blocks gave rise to reduced RT2 because the identity of the second stimulus became progressively predictable with a continuous exposure to the training stimulus pairs.

Test phase Figure 3 represents the RT2 data obtained during the test phase in the AS+ and AS− groups.

Analyses for the AS+ group With our test procedure, we expected that the effects of symmetry (if any) will emerge in the AS+ group immediately after the different pairs were recombined, therefore during the first test trials. Two-tailed paired t tests compared the mean RT2 for the first 10 trials of each pair. These analyses revealed that the response times were shorter on average for the B-A pair (mean = 584 ms, $SD = 57$ ms) than for the F-C pair (620 ms, $SD = 60$ ms), $t(9) = 3.22$, $p < .05$, Cohen's $d = .81$. Importantly, the stimuli composing the B-A and F-C test pairs all changed their position in comparison to training, and in these two pairs, only B-A preserved the association learned during the training phase. The shorter RT2 obtained for B-A, in comparison to F-C, can therefore be explained by the fact that the subject expected to see an A in the second display of the symmetrical B-A pairs, while no similar prediction was possible in the F-C pairs. This

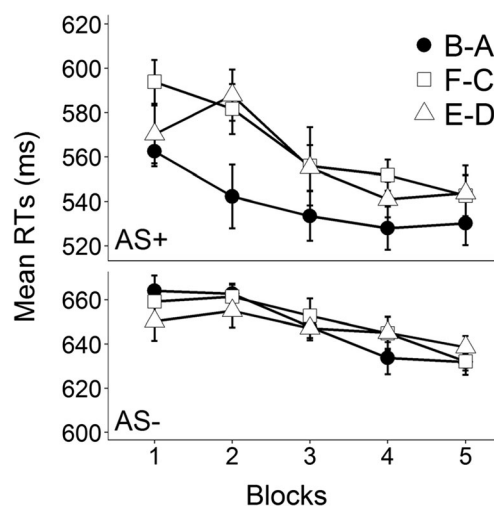


Fig. 3 Mean correct response times on the second display (RT2) obtained in the AS+ and AS− group for each test block and stimulus pair. Note that the y -axes have different scales in the two groups and display only a subset of the range. Error bars indicate the standard errors centered on each participant mean RT

result suggests that the human participants from the AS+ group processed the relation of symmetry.

Data analyses also revealed shorter RT2 on average for the E-D pair (mean = 543 ms, $SD = 32$ ms), compared to both the B-A, $t(9) = -2.55$, $p < .05$, Cohen's $d = .47$, and the F-C pairs, $t(9) = -4.66$, $p < .01$, Cohen's $d = .49$. This effect indicates that the subjects from the AS+ group also processed the position of the stimuli, in addition to their processing of symmetry reported above.

Another question is if these effects could persist in AS+ with repeated exposure to the five test blocks. This question was addressed with a stimulus pair (B-A, E-D, F-C) by test block (1–5) two-way ANOVA on RT2. This analysis revealed that the main effect of test block was significant, $F(4, 36) = 4.59$, $p < .01$, $\eta_p^2 = .33$. It showed that the RT2 decreased on average with repeated testing. The effect of stimulus pair approached significance, $F(2, 18) = 3.53$, $p < .06$, $\eta_p^2 = .28$, and the stimulus pair by test block interaction was not significant, $F(8, 72) = 1.12$, $p > .10$, $\eta_p^2 = .11$. The effect of stimulus pair, albeit not significant, is shown in Fig. 3 by systematically shorter RT2 in the five test blocks for B-A than for both F-C and E-D. The fact that these RT2 remained systematically shorter during the five test blocks suggests a relative persistence of the effect of symmetry in our group of participants. Conversely, the initial advantage in RT2 found for the E-D pair during the analysis of the 10 first test trials disappeared rapidly. Figure 3 indeed indicates that RT2 for the E-D pair is shorter than the RT2 for the F-C pair in the first test block only. The effect of position seems therefore more transitory in humans of the AS+ group than the effect of stimuli symmetry.

Analyses for the AS– group The same set of analyses conducted on the AS– group revealed shorter response times in the 10 first trials on average for the E-D pair (mean = 642 ms, $SD = 96$ ms) than for the F-C pair (mean = 679 ms, $SD = 76$ ms), $t(10) = 2.36$, $p = .04$. This difference suggests an encoding of the position within the first 10 trials which is confirmed by a nonsignificant trend for shorter RT2 on average for E-D than for B-A (mean = 673 ms, $SD = 76$ ms), $t(10) = 2.05$, $p = .068$. No reliable difference emerged between B-A and F-C, $t(10) = 0.39$, $p > .10$.

The ANOVA performed on the five blocks of the test phase revealed a main effect of block, $F(4, 40) = 5.07$, $p < .01$, $\eta_p^2 = .37$, but no reliable effect of condition, $F(2, 20) = 0.14$, $p > .100$, $\eta_p^2 = .01$, and no significant interaction between these two factors, $F(8, 80) = 1.10$, $p > .100$, $\eta_p^2 = .10$. Taken together, results of the AS– group confirm that participants from the AS– group did not learn that the item that blinks once in the first display predicts a specific item in the second display. It seems that these participants encoded only the positional properties of the individual stimuli. However, as for the AS+ group, this encoding seems transitory as it only appeared in

the first 10 test trials, but disappeared when the entire test phase was considered.

Verbal reports All the subjects from the AS+ group, but none of the AS– group, reported verbally after the experiment that they had noticed a change during the experiment (i.e., the changes between the training and testing phases). Within AS+, three subjects indicated that one association (i.e., for A-B) was preserved after the change. These data suggest variations among the AS+ subjects, some of them seem to process the symmetry relation implicitly, while the three subjects who noticed that the A-B association was preserved, processed it more explicitly. Despite these variations, all the AS+ subjects reported that they have given labels to the object stimuli.

Baboons

The baboons performed 76.30% correct on average during the training phase ($SD = 42.48\%$) and therefore made more errors than humans. On that basis, we report below statistical analyses on both response times and accuracy.

Response times of the training phase We found earlier that the human subjects who noticed that the stimuli were organized in pairs during the training phase (AS+ participants) showed a progressive decline in RT2 during that phase. Because this decline in RT2 is in humans an indicator of learning, our first analyses of the baboons' data distinguished two groups of baboons. The first group (referred to below as the AS+ group) was composed of the individuals who showed reduced RT2 with training, and the second group (referred to below as the AS– group) was composed of the individuals who did not show this effect. This was achieved by way of one-tailed t tests ($p < .05$) comparing the mean RT2 obtained, for each individual, during the first and last blocks of the training phase. From these analyses, eight baboons were assigned to the AS+ group, the remaining six subjects were assigned to the AS– group.

Figure 4 represents the mean RT1 and RT2 obtained in the AS+ and AS– groups during training. Because the number of errors was greater in baboons than in humans, reducing the number of RT measures per block, statistical analyses of correct RT1 and RT2 considered blocks of 192 trials made from the grouping of two initial contiguous 96-trial blocks. RT1 and RT2 data were analyzed independently for the AS– and AS+ groups, using two-way ANOVAS, considering the stimulus pair (A-B, C-D, E-F) and the training block (ten 196-trial blocks) as independent variables.

For the AS+ group, the two-way ANOVA on RT1 revealed a main effect of block, $F(9, 63) = 3.40$, $p < .002$, $\eta_p^2 = .33$. The main effect of stimulus pair and the interaction between the pair and block factors were not significant (all $ps > .10$).

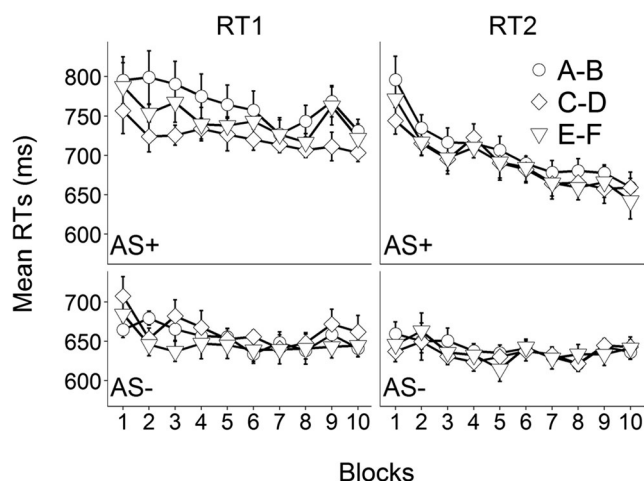


Fig. 4 Mean RT1 and RT2 obtained during the training phase for the AS+ and AS- groups of baboons. The y-axes have different scales in the two groups and display only a subset of the range. Error bars represent the standard errors centered on each participant mean RT

Similar results were obtained in AS+ for RT2, with a reliable effect of block being the only significant effect, $F(9, 63) = 12.37, p < .001, \eta_p^2 = .64$.

For the AS- group, the two-way ANOVA conducted on RT1 showed a main effect of block, $F(9, 45) = 2.69, p < .02, \eta_p^2 = .35$, but no reliable effect of stimulus pair and no stimulus pair by block interaction (all $ps > .1$). When computed on RT2, the same ANOVA revealed no reliable main effect or interaction in AS- (all $ps > .10$).

In sum, the individuals from the AS+ groups showed an effect of block for both RT1 and RT2. Regarding RT1, we can safely assume that the effect of block indicates a general effect of practice, reducing the response times with continuous training. Regarding RT2, the question arises of whether block effect only corresponds to a general effect of practice, as per RT1, or can also reflect the learning of the associations between the stimuli belonging to each pair. To address this question, we ran a complementary two-way ANOVA on the response times in which the training block (1–10) and the target of the response (Object 1 vs. Object 2) were used as factors. This ANOVA revealed a general effect of block, $F(9, 63) = 10.36, p < .001, \eta_p^2 = .60$, as well as a main effect of target, $F(1, 7) = 13.44, p < .01, \eta_p^2 = .66$. Of the most importance for our purpose, these two factors interacted significantly, $F(9, 63) = 2.54, p < .05, \eta_p^2 = .27$. As shown in Fig. 4, training has a stronger effect on the time to respond to the second than to the first object of the stimulus pairs. This result is a strong indicator that the subjects of the AS+ group have learned that the stimuli are organized in pairs. By contrast, no decrease in RT2 was observed in AS-. This effect confirms that the AS+ and AS- groups behaved differently in the task, regarding their response to the second display. We will see below that these differences in behavior correspond to cognitive differences during the test phase.

Accuracy of the training phase Figure 5 reports the mean accuracy, for each group and training conditions. Following the same logic as above, the accuracy to the first object (Object 1) and second (Object 2) objects of the stimulus pairs were analyzed separately with block (1–10) by stimulus pairs (A-B, C-D, E-F) two-way ANOVAs computed independently for the two groups of baboons.

For the AS+ groups, the two-way ANOVA on Object 1 only found a reliable effect of block, $F(9, 63) = 4.41, p < .01, \eta_p^2 = .39$. A similar result was obtained when the response to Object 2 was considered, with the effect of block being the only significant effect, $F(9, 63) = 5.63, p < .001, \eta_p^2 = .43$. Because the effect of block was significant in AS+ for both Objects 1 and 2, the question arises of whether the increased accuracy with blocks for Object 2 can be accounted for only by a general effect of practice, as observed for Object 1, or if this effect can also be explained by the learning of the associations between the first and second stimulus of each pair. To assess this idea, we ran an additional block (1–10) by object (Object 1 vs. Object 2) two-way ANOVA, using accuracy as the dependent variable. This analysis indicated no main effect of object ($p > .10$) but a main effect of block, $F(9, 63) = 6.65, p < .001, \eta_p^2 = .49$, showing that the performance increased with training. More importantly for our purpose, the Object \times Block interaction was also significant, $F(9, 63) = 2.68, p < .02, \eta_p^2 = .28$, indicating that the effect of block was stronger for Object 2 than for Object 1. These findings therefore converge perfectly with those reported above on both RT1 and RT2. They confirm that the baboons from AS+ progressively learned the associations between the first and second objects of each stimulus pair.

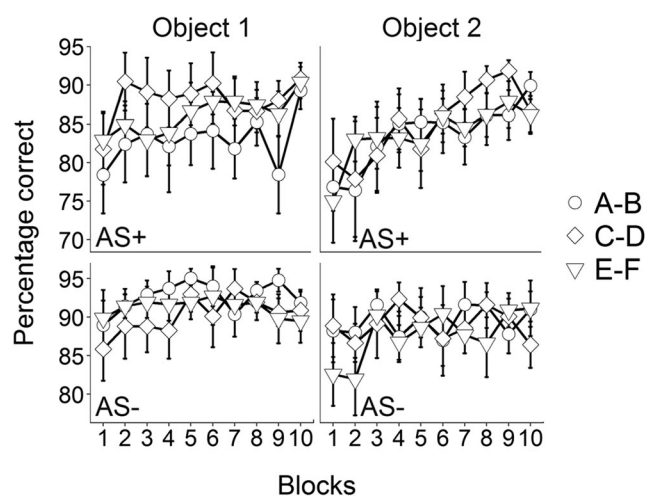


Fig. 5 Mean accuracy obtained during the training phase in baboons, depending on the group of baboons (AS+ and AS-), 196-trial test block, and stimulus pair. The y-axes have different scales in the two groups and display only a subset of the range. Error bars represent the standard errors

For the AS⁻ group, the two-way ANOVA using the accuracy of the response to Object 1 revealed no significant main effect or interaction (all $ps > .05$), and these results were replicated in the analyses of the response to Object 2 (all $ps > .05$). Taken together, these data reinforce the idea that the baboons from AS⁻ did not learn that the stimuli were organized in pairs and were therefore unable to predict the identity of Object 2 depending on the identity of Object 1.

Response times of the test trials Figure 6 represents the mean response times obtained during the test in AS⁺ and AS⁻ groups, for each block and stimulus pair. With our procedure, we first expected that the effects of symmetry (if any) will be of maximal amplitude during the first test trials. These effects might eventually decrease in amplitude or even disappear because of the learning of the novel stimulus pairing that may occur with the repetition of the test blocks.

Analyses for the AS⁺ group To analyze the response times obtained for AS⁺ at the onset of the test, we computed two-tailed paired t tests comparing the mean RT2 obtained in the first 10 trials for the different stimulus pairs. These analyses indicated shorter RT2 for the stimulus pair E-D (mean = 674 ms, $SD = 54$ ms) than for F-C (mean = 768 ms, $SD = 77$ ms), $t(7) = -4.17$, $p < .01$, Cohen's $d = 0.58$. The differences between B-A (mean = 727 ms, $SD = 136$ ms) and E-D, or between B-A and F-C were not statistically significant ($ps > .10$). These results suggest that the AS⁺ baboons processed stimulus positional information. However, they provide no indication that these baboons were able to process the relation of symmetry.

To analyze how RT2 evolved with the repetition of the test blocks, we ran a block (1–5) by stimulus pair (B-A, F-C, E-D) two-way ANOVA on RT2 using blocks of 192-trials made from the grouping of two initial 96-trial blocks. This analysis showed a reliable effect of block, indicating that the RT2 decreased through continuous training, $F(4, 28) = 3.73$, $p < .02$, $\eta_p^2 = .35$, but no effect of stimulus pair, $F(2, 14) = 2.60$, $p > .10$, and no reliable interaction $F(8, 56) = 1.24$, $p > .10$. Note the lack of effect of stimulus pair departs from the data obtained from the analysis of the 10 first trials. This discrepancy suggests that the effect of the preservation of the position of the stimuli in the training pairs declined with repeated testing.

Analyses for the AS⁻ group The same analyses conducted on the AS⁻ group revealed no difference in mean RT2 between the stimulus pairs when the first 10 test trials were considered (B-A: mean = 656 ms, $SD = 81$ ms; F-C: mean = 664 ms, $SD = 38$ ms; E-D = 668 ms, $SD = 76$ ms; all $ps > .10$). In addition, the two-way ANOVA showed no reliable main effect of stimulus pair or block, and interaction between these two factors ($ps > .10$).

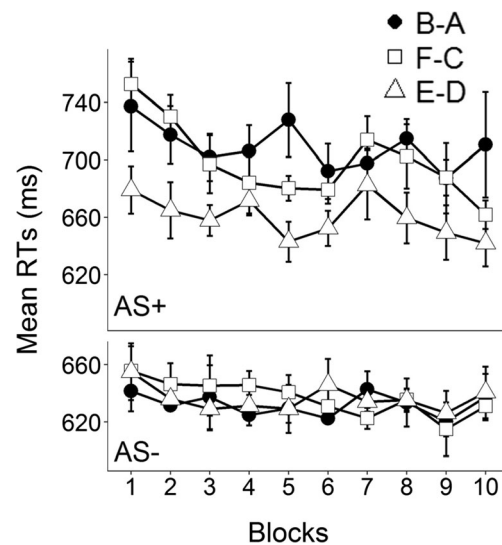


Fig. 6 Mean responses times on the second object (RT2) obtained in the AS⁺ and AS⁻ groups, for each stimulus pair and test block. The y-axes differ between the two groups. Error bars represent the standard errors centered on each participant mean RT

Accuracy of the test trials Figure 7 represents how the accuracy varied across the test blocks and stimulus pairs. Accuracy analyses followed the same logic as for the analyses of RT2. First, paired two-tailed t tests compared the mean performance on Object 2 obtained during the first 10 test trials of each stimulus pair. The next analyses focused our attention on all the test blocks.

Analyses for the AS⁺ group This analysis revealed a trend for a greater performance at the onset of the test phase with E-D (mean = 92% correct, $SD = 11\%$) than with both F-C (mean = 71%, $SD = 19\%$), $t(7) = 2.35$, $p = .05$, Cohen's $d = -0.37$, and B-A (mean = 66%, $SD = 30\%$), $t(7) = 2.20$, $p = .064$, Cohen's $d = -1.16$. No difference was found between F-C and

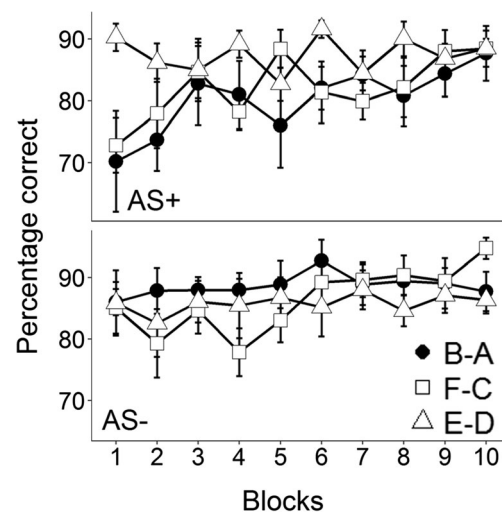


Fig. 7 Mean accuracy obtained in the two groups of baboons for each stimulus pair and test block. The y-axes differ between the two groups. Error bars represent the standard errors

B-A ($p > .10$). The results obtained from the analyses of the accuracy and mean RT2 were therefore highly consistent as the advantage found above for E-D in terms of response times converge with an accuracy advantage. Because the E-D pair is the sole test pair for which the order of the stimulus was preserved in comparison to training, this advantage for E-D confirms a processing of the positional information by the baboons from the AS+ group.

The second set of analyses on the five test blocks revealed a main effect of block, $F(4, 28) = 6.71$, $p < .001$, $\eta_p^2 = .49$, as well as a significant block by pair interaction, $F(8, 56) = 2.34$, $p < .05$, $\eta_p^2 = .25$. No main effect of stimulus pair was found ($p > .10$). Post hoc analyses revealed that the baboons' performance with the E-D pair was higher compared to the B-A pair in the first test block, but that no difference was observed for the next blocks. These results therefore converge with those reported above on RT2. Altogether, these results indicate that the initial advantage for E-D in terms of RT and accuracy progressively disappeared during the test.

Analyses for the AS− group The analyses conducted on the data from the AS− group confirmed that no difference in accuracy was observed for this group during the first 10 test trials between the B-A (mean = 93% correct, $SD = 6\%$), F-C (mean = 91% correct, $SD = 9\%$), and E-D (mean = 87% correct, $SD = 15\%$) pairs ($ps > .10$). The two-way ANOVA considering the five blocks of tests showed no reliable main effect nor interaction ($ps > .10$).

Discussion

This comparative human/baboon experiment examined the encoding of pairs of stimuli in sequences to identify potential species differences in their processing of the symmetry relation. A serial response time task involving pairs of objects was developed and presented to the two species for that purpose. In our task, we presumed that the individuals of the two species who learned the association between the objects to be touched on the displays should demonstrate an RT2 acceleration with training. We distinguished two groups of subjects on that basis and categorized all the individuals showing an RT2 acceleration during training in the AS+ group; all the remaining subjects were categorized in the AS− group.

The results of the test phase confirmed our expectation that the baboon and human subjects of the AS− groups did not process the associations between the target stimuli. In the AS− baboons, we obtained no reliable effect of stimulus pair and no Block \times Stimulus Pair interaction in the test phase. We propose that the baboons of the AS− group only processed the blink in each display but did not process the fact that the stimuli which blinked once within a trial were organized in pairs. In the AS− humans, the only effect emerging from the

test trials is a response time advantage for the E-D pair occurring at the onset of the test. This advantage suggests that the AS− human subjects encoded at least minimally the identity of the stimuli and their position. The processing of the positional information might therefore occur in the AS− human group along with the processing of the information on the blink mentioned above for the baboons.

Test trials identified the cognitive strategy employed by the members of the AS+ groups of each species. In humans, we found shorter RT2 in the B-A than in the F-C pair. The shorter RT2 obtained for B-A in Test Block 1, in comparison to F-C, can be explained by the preservation of the associative link between the stimuli in the B-A pair only. We take this result as demonstrating that humans from the AS+ group processed the associations between the stimuli, and furthermore processed A-B and B-A as symmetrical relations in our study. Importantly, humans from this AS+ group also demonstrated a RT2 advantage for E-D over B-A. This advantage for E-D suggests that this group also processed positional information, along with the information of the associations.

These results in humans contrast sharply with those obtained in baboons from the AS+ group. In these baboons, the stimulus pair giving rise to the shortest RT2 was E-D, not B-A, as found in humans. We also found a reliable difference between E-D and F-C. The advantage for the stimulus pair E-D, which emerged in the first 10 trials, is confirmed by a better accuracy for E-D, in comparison to both F-C and B-A. Critically, neither E-D nor F-C preserved the relations learned during the training phase, and the positional information is preserved in E-D only. Therefore, while baboons from the AS+ group processed the associations between the stimuli in the training phase, they did not process A-B and B-A as symmetrical relations in our study. Therefore, our results suggest an encoding of positional information in baboons of the AS+ group. This encoding of the positional information is further confirmed by a lack of statistical difference between B-A and F-C, which are the two stimulus pairs reversing their position in comparison to training. In turn, results suggest that the baboons from the AS− group have neither encoded the positional nor the relational information. Apparently, these subjects only paid attention to the blink cue, as discussed above.

The fact that the baboons from the AS+ group give priority to the positional information converges with previous findings and conclusions from Urcuioli (2015). Our results also add to the literature this idea that prioritization of positional over relational information is a pervasive effect that can be replicated in pigeons and monkeys, at least within the nonhuman animal species. The demonstration that nonhuman animals give priority to perceptual cues over relational ones is not new in the comparative literature. Such effects were already found with relational stimuli of different types (hierarchical stimuli: Fagot & Deruelle, 1997; artificial relational stimuli:

Fagot, Kruschke, Dépy, & Vauclair, 1998; Wasserman, Fagot, & Young, 2001; ABA/ABB sequential patterns: Chen, van Rossum, & ten Kate, 2015). Nevertheless, our study makes it even more striking, following Urcuioli (2015), that such prioritization has major impacts on the nonhuman animals' ability to both processes symmetry and to form equivalence classes based on symmetrical relations.

The current research has practical consequences for further studies on symmetry and equivalence class formation in animals. As already suggested by Urcuioli (2015), it confirms that the search for symmetry in animals requires a strict control of the positional cues that might be available in the task. One limitation of our study is that spatial and temporal/ordinal information covaried during the training and test phases, because the first stimulus of each pair was always presented on the left side of the screen, the second one being presented on the right side. Our study therefore provides no information to distinguish the exact influence of these two factors (but see Lionello-DeNolf, 2009, for a discussion of this issue). Nevertheless, the current results as well as those of previous research all point to the same conclusion: They suggest that this kind of perceptual cues are very important to consider in future experimental designs. Note, however, that such effects do not rule out the possibility that other factors also intervene. For instance, Medam et al. (2016) proposed that experimental designs mixing symbolic and identity matching trials (as in Frank & Wasserman, 2005) promote cognitive flexibility, because the subjects can learn that the stimuli can have different roles in the task. We suggest that a more flexible encoding of the stimulus pair should also be favored in future experimental designs.

The cognitive ability to solve symmetry problems seems highly adaptive, and experts of primate cognition can easily mention situations in which animals apparently treat relations between items as symmetrical. One classic example of this kind is the case of vervet monkeys, who can both produce alarm calls in response to leopards (leopard → alarm call) and can meanwhile show specific behavioral responses when hearing these calls, suggesting that they also treat the call → leopard symmetrical association (Seyfarth, Cheney, & Marler, 1980). Although this kind of situations probably occurs in nature, they are neither as variable nor as frequent as found during interpersonal verbal communication in humans involving a continuous shift between speech production and speech perception. When communicating verbally with others, our comprehension of others' speech requires that the words we perceived are mapped to their referent, while speech production requires the reversed referent to word mapping. In other words, verbal communication is an exercise for which it is central to treat the associations/relationships between and among items as being bidirectional. This importance of symmetry for our communicative system might explain previous demonstrations of a strong relationship between the ability to

develop equivalence classes and language skills (e.g., Devany et al., 1986). An important result in our study is that the AS+ humans who used labels to code the stimuli demonstrated symmetry. By contrast, the AS+ baboons and AS− humans who did not use labels to code the stimuli failed to process the stimulus pairing and relied exclusively on positional cues to solve the task. These results suggest an effect of stimulus labeling on the processing of the task. More experimental work will be needed to better understand this effect.

Most of the associations occurring in the environment are unidirectional. This is true in both humans and nonhuman animals (e.g., a horn predicts a car in humans, the animal keeper predicts food delivery in baboons, and the reverse associations are not true), but only humans have extra training with symmetrical relations via language. Other animal species that are missing a human-like communication system, and thus the extensive bidirectional training that humans have with their exposure to language, may require specific training and test conditions to learn that the stimuli go together in a bidirectional way; to neglect information such as the stimuli position, which is irrelevant to the task; and to be able to derive untrained (symmetrical) relations from the trained ones.

Acknowledgements The authors thank the staff of the CNRS primate center (Rousset-sur-Arc, France) as well as Julie Gullstrand for technical supports. This research was supported by the project Premilang2 (#ANR-13-BSH2-0002), labex BLRI (ANR-11-LABX-0036), and the Convergence ILCB institute.

Compliance with ethical standards

Ethical standard This research conformed to the Standard of the American Psychological Association's Ethical Principles of Psychologist and Code of Conduct, and received ethical approval from the French Ministry of Education (approval APAFIS#2717-2015111708173794 v3).

References

- Adams, B. J., Fields, L., & Verhave, T. (1993). Formation of generalized equivalence classes. *Psychological Record*, *43*, 553–566.
- Barros, R. S., Galvão, O. F., & McIlvane, W. J. (2002). Generalized identity matching-to-sample in *Cebus apella*. *The Psychological Record*, *52*, 441–460.
- Boelens, H., van den Broek, M., & van Klarenbosch, T. (2000). Symmetric matching to sample in 2-year-old children. *Psychological Record*, *50*, 293–304.
- Carp, C. L., & Petursdottir, A. L. (2015). Intraverbal naming and equivalence class formation in children. *Journal of the Experimental Analysis of Behavior*, *104*, 223–240.
- Carr, D., Wilkinson, K. M., Blackman, D., & McIlvane, W. J. (2000). Equivalence classes in individuals with minimal verbal repertoires. *Journal of Experimental Analysis of Behavior*, *74*, 101–114.
- Chen, J., van Rossum, D., & ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition*, *18*(1), 151–164. doi:10.1007/s10071-014-0786-4

- D'Amato, M. R., Salmon, D. P., Loukas, E., & Tomie, A. (1985). Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*). *Journal of the Experimental Analysis of Behavior*, 44, 35–47.
- Devany, J. M., Hayes, S. C., & Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *Journal of the Experimental Analysis of Behavior*, 46, 243–257.
- Eikeseth, S., & Smith, T. (1992). The development of functional and equivalence classes in high-functioning autistic children: The role of naming. *Journal of Experimental Analysis of Behavior*, 58, 123–133.
- Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free ranging baboons. *Behavior Research Methods*, 42, 507–516.
- Fagot, J., & Deruelle, C. (1997). Processing of global and local visual information and hemispheric specialization in humans (*Homo sapiens*) and baboons (*Papio papio*). *Journal of Experimental Psychology: Human Perception and Performance*, 23, 429–442.
- Fagot, J., Kruschke, J. K., Dépy, D., & Vaclair, J. (1998). Associative learning in humans (*Homo sapiens*) and baboons (*Papio papio*): Species differences in learned attention to features. *Animal Cognition*, 1, 123–133.
- Fagot, J., & Paleressompouille, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, 41, 396–404.
- Frank, A. J., & Wasserman, E. A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *Journal of the Experimental Analysis of Behavior*, 84, 147–165.
- Kastak, C. R., Schusterman, R. J., & Kastak, D. (2001). Equivalence classification by California sea lions using class-specific reinforcers. *Journal of the Experimental Analysis of Behavior*, 76, 131–158.
- Kuno, H., Kitadate, T., & Iwamoto, T. (1994). Formation of transitivity in conditional matching to sample by pigeons. *Journal of the Experimental Analysis of Behavior*, 62, 399–408.
- Lionello-DeNolf, K. M. (2009). The search for symmetry: 25 years in review. *Learning & Behavior*, 37, 188–203.
- Lipkens, G., Hayes, S. C., & Hayes, L. J. (1993). Longitudinal study of derived stimulus relations in an infant. *Journal of Experimental Child Psychology*, 56, 201–239.
- Lipkens, R., Kop, P. F., & Matthijs, W. (1988). A test of symmetry and transitivity in the conditional discrimination performances of pigeons. *Journal of the Experimental Analysis of Behavior*, 49, 395–409.
- Medam, T., Marzouki, Y., Montant, M., & Fagot, J. (2016). Categorization does not promote symmetry in Guinea baboons (*Papio papio*). *Animal Cognition*, 19, 987–998.
- Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science*, 2(1), 138–153.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32.
- O'Donnell, J., & Saunders, K. J. (2003). Equivalence relations in individuals with language limitations and mental retardation. *Journal of Experimental Analysis of Behaviour*, 80, 131–157.
- Pepperberg, I. M., & Gordon, J. D. (2005). Number comprehension by a grey parrot (*Psittacus erithacus*), including a zero-like concept. *Journal of Comparative Psychology*, 119, 197–209.
- Schusterman, R. J., & Kastak, D. (1993). A California sea lion (*Zalophus californianus*) is capable of forming equivalence relations. *Psychological Record*, 43, 823–839.
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210, 801–803.
- Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, 14, 5–13.
- Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982). A search for symmetry in the conditional discriminations of rhesus monkeys, baboons, and children. *Journal of the Experimental Analysis of Behavior*, 37, 23–44.
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37, 5–22.
- Sweeney, M. M., & Urcuioli, P. J. (2010). Reflexivity in pigeons. *Journal of the Experimental Analysis of Behavior*, 94, 267–282.
- Urcuioli, P. J. (2008). Associative symmetry, antisymmetry, and a theory of pigeons' equivalence-class formation. *Journal of the Experimental Analysis of Behavior*, 90, 257–282.
- Urcuioli, P. J. (2015). A successful search for symmetry. *Conductal*, 3, 4–25.
- Wasserman, E. A., Fagot, J., & Young, M. E. (2001). Same-different conceptualization by baboons (*Papio papio*): The role of entropy. *Journal of Comparative Psychology*, 115, 42–52.
- Zentall, T. R., Clement, T. S., & Weaver, J. E. (2003). Symmetry training in pigeons can produce functional equivalences. *Psychonomic Bulletin & Review*, 10, 387–391.

III. Extraction de Régularités Séquentielles : Données Comparatives chez le Babouin et l'Humain

Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (*sous presse*). Regularity extraction across species: associative learning mechanisms shared by human and non-human primates. *Topics in Cognitive Science*.

Regularity extraction across species:

associative learning mechanisms shared by human and non-human primates

Arnaud Rey, Laure Minier, Raphaëlle Malassis,

Louisa Bogaerts, and Joël Fagot

Laboratoire de Psychologie Cognitive

CNRS & Aix-Marseille Université, Marseille, France

Abstract

Extracting the regularities of our environment is a core cognitive ability in human and non-human primates. Comparative studies may provide information of strong heuristic value to constrain the elaboration of computational models of regularity learning. The current study illustrates this point by testing human and non-human primates (Guinea baboons, *Papio papio*) with the same experimental paradigm, using a novel online learning measure. For *local* co-occurrence regularities, we found similar patterns of regularity extraction in baboons and humans. However, only humans extracted the more *global* sequence structure. It is proposed that only the first result that is common to both species should be used to constrain models of regularity learning. The second result indicates that the extraction of global regularities cannot be accounted for by mere associative learning mechanisms and suggests that humans probably benefit from their language recoding abilities for extracting these regularities. We propose to use a comparative approach to address a series of remaining theoretical questions, which will contribute to the development of a general theory of regularity learning.

Keywords: regularity extraction; statistical learning; implicit learning; associative learning; verbal recoding

Statistical learning – the ability to extract and encode environmental regularities – appears as a key feature of human cognitive systems and has been a domain of intensive research over the last 20 years (see (R. Frost, Armstrong, Siegelman, & Christiansen, 2015b), and Thiessen, Kronstein & Hufnagle, 2013, for reviews, but also e.g., Estes, 1950; Restle, 1970, for earlier pioneer work). It is thought to play a key role in the segmentation of continuous input, prediction, discrimination and categorization, shaping our basic representations of the environment (Frost et al., 2015).

Historically, before the ‘statistical learning’ wave, similar issues were addressed in the field of ‘implicit learning’. The latter term captures the observation that some forms of learning occur after the mere repeated exposure to regularities (or rule systems, as in artificial grammar learning experiments, e.g., Reber, 1967), with no intention of learning and no clear awareness of what has been learned (Cleeremans, Destrebecqz, & Boyer, 1998; Shanks, 2005). Some authors, such as Conway and Christiansen (2006; see also Christiansen, this issue), have proposed to merge the two terms (i.e., ‘implicit statistical learning’) to cover the general learning ability to extract and encode different forms of environmental regularities.

Others, like Perruchet (2005; see also Perruchet & Poulin-Charronnat, 2012), have argued that ‘statistical learning’ is nothing more than a new terminology corresponding, in fact, to the old domain of associative learning (e.g., Hebb, 1961; Mitchell & Le Pelley, 2010; Shanks, 1995). In this view, the extraction of regularities is understood as the formation of new larger units also called ‘chunks’. This process is supported by associative learning mechanisms, which itself is considered a by-product of the concurrent attentional processing of multiple elements (or small units). Learning can also be qualified as ‘implicit’ as it doesn’t require awareness of what is learnt nor the intention to learn (see also Perruchet & Pacton, 2006, Perruchet & Vinter, 2002)³⁷. In that perspective, implicit statistical learning and associative learning are considered as two sides of the same coin, the former having a broader scope (including issues related to various aspects of language processing) compared to the latter that was more concerned by the formation of simple associations or by operant conditioning behaviors.

Basic associative (or implicit statistical) learning mechanisms are certainly shared by human and non-human primates (e.g., Wilson, Marslen-Wilson, & Petkov, 2017), and possibly

³⁷ In Perruchet and Vinter’s (2002) theory, learning involves attention or, more precisely, attentional processing of the learned information. In that sense, attention or attentional processing are equivalent to consciousness, but not awareness. According to this view, one can learn a regularity by paying attention to that regularity while being unaware of that regularity (i.e., not being able *to report* the exact content of that regularity).

by all species equipped with a nervous system (e.g., chicks: Santolin, Rosa-Salva, Regolin, & Vallortigara, 2016; songbirds: Chen, van Rossum, & ten Cate, 2015; rodents: Toro, Nespore & Gervain, 2016). This raises the question which aspects of the human ability for regularity extraction can be related to those in non-human primates. The present study aims to address this question by a direct behavioral comparison between humans and Guinea baboons (*Papio papio*) in the same experimental paradigm. This paradigm provides an online measure of regularity extraction at different levels of integration. We argue that the use of an online learning measure on the one hand and the study of non-human primates on the other hand form a valuable approach to overcome two difficulties that characterize the existing literature.

The first difficulty is that the temporal dynamics of regularity extraction is frequently not considered. The seminal study of Saffran, Aslin, and Newport (1996) is often considered as the critical starting point in the study of statistical learning, and several theoretical accounts and computational models of this phenomenon have been proposed (e.g., Frank, Goldwater, Griffiths, & Tenenbaum, 2010; French, Addyman, & Mareschal, 2011; Perruchet & Vinter, 1998; Pothos, 2007). The empirical evidence used to elaborate these models is mainly derived from studies using the same artificial language learning paradigm. In this task, participants are exposed to an artificial language composed of few polysyllabic nonsense words (e.g., 4) that are auditorily presented without any pause between words. After this exposure phase, participants are asked to indicate which one of two test sequences looks more familiar. The main dependent variable used to test models is therefore an offline measure that informs us about which kind of regularities can be learned but that does not provide any information regarding the temporal dynamics of learning (see e.g., Siegelman, Bogaerts, & Frost, 2017, for a detailed discussion of the problems of offline learning measures). In contrast to offline measures, online measures are obtained by recording the performance of participants during the exposure to the regularities and can therefore provide information on the temporal dynamics of regularity extraction. The serial response time task (Nissen & Bullemer, 1987) is one good example of such online measure. In this task, participants respond to sequences of stimuli that appear one-by-one at various locations on a computer screen. If the sequences include regularities (allowing for the prediction of upcoming locations), then participants speed their response times (RTs) if they manage to extract these regularities (see also Batterink, 2017; Christiansen, this issue; Misyak, Christiansen, & Tomblin, 2010; Siegelman, Bogaerts, Kronenfeld, & Frost, 2017, for other examples of online learning measures).

A second difficulty is paradoxically the strong focus on human (adult) participants. Indeed, during the exposure phase, participants may develop strategies or explicit learning procedures (like repeating internally some sequences of syllables) and their resulting performance is certainly influenced and biased by various forms of explicit recoding processes (Batterink, Reber, Neville, & Paller, 2015; Siegelman et al., 2017). Testing human infants without any advanced language and reasoning abilities (like the original Saffran et al. study did, followed by many others, e.g., Bulf, Johnson, & Valenza, 2011; Singh, Reznick, & Xuehua, 2012) could be a solution to overcome that difficulty. However, the study of infants also brings other procedural difficulties because the measures that can be recorded for this population (for example, the head-turn procedure) are indirect and much noisier (i.e., less reliable). Testing non-human primates allows us to study regularity extraction in the absence of any bias due to language experience and the related ability to verbally recode regularities.

Previous studies that adopted a comparative approach to study statistical learning were mainly concerned with discovering differences between human and non-human primates (e.g., Fitch & Hauser, 2004; Wang, Uhrig, Jarraya, & Dehaene, 2015). That logic is derived from a research tradition that searched for a critical distinctive feature that could distinguish animal communication systems from human language. For instance, Hauser, Chomsky, and Fitch (2002) put forward the ability to process and produce recursive structures as a fundamental distinctive feature that sets humans apart from other species. However, as argued by Pinker and Jackendoff (2005), there are certainly many other critical differences between human and non-human primates that can account for their distinct cognitive trajectories (see also, Fagot, Malassis, Medam, & Montant, 2017).

Here we propose an alternative perspective stating that we can increase our understanding of the precise dynamics of regularity learning processes by looking at the *common* patterns produced by human and non-human primates in online experimental paradigms. As mentioned above, it is obvious that human adults can make use of both implicit associative learning mechanisms and explicit recoding strategies based on their language abilities to perform regularity extraction tasks. Non-human primates miss the language tools that humans can use to explicitly recode regularities. One can therefore safely assume that non-human primates mainly rely on associative learning mechanisms in these tasks, and that these mechanisms are shared with humans. Therefore, any common pattern of performance between human and non-human primates should reflect common fundamental properties of these learning mechanisms. Alternatively, any difference could be related to the use of language

recoding abilities and should therefore be considered with caution for constraining the development of computational models.

We now illustrate this proposition by comparing the performance of human and non-human primates in a novel task that has been designed to combine the serial response time task and the artificial language paradigm (Franco & Destrebecqz, 2012). Human adults and baboons had to touch a red circle appearing successively at nine different locations of a computer screen. The dot to follow moved on the screen along paths containing statistical regularities. The use of this task allowed us (1) to test the two populations with the exact same experimental procedure, (2) to provide detailed information on the temporal dynamics of regularity extraction, and (3) to compare humans and baboons at different levels of integration, at the local level within the sequence (i.e., co-occurrences between individual elements), or at the level of the global sequence structure (see the Method section below for an operational definition of the distinction between local and more global regularities). Note that part of the primate data was reported in Minier, Fagot, and Rey (2015).

Methods

Participants were ten Guinea baboons (*Papio papio*, age range 3–15.5 years) from the CNRS primate facility in Rousset (France) and 5 human adults (age range 20–24 years) from Aix-Marseille University.

Baboons live within a larger group of 26 individuals, within a 700 m² outdoor enclosure and had a permanent access to ten Automated Learning Devices for Monkeys (ALDM, for a detailed description, see (Joël Fagot & Bonté, 2010b; Joël Fagot & Paleressompouille, 2009) equipped with a 19-inch touch screen and a food dispenser. The main feature of ALDM equipment is that a radio frequency identification reader (RFID) identifies each baboon via a microchip implanted in each arm. The baboons can therefore participate to the research at will, without being captured, as the test programs recognize them automatically. All baboons had previously participated to numerous computerized experiments using the ALDM test systems. The experiment was controlled by the E-Prime software (Version 2.0, Psychology Software Tools, Pittsburgh).

A trial began with the presentation of a fixation cross at the bottom of a touch screen. The cross stayed on the screen till it was pressed. After pressing on it, the fixation-cross disappeared and immediately 9 crosses representing locations were displayed on the screen,

one of them being replaced by the target, a red circle. When the target was touched, it disappeared and was replaced by the cross. The next cross/location in the sequence was then replaced by the red circle until the end of the sequence. Reward (a small pellet) was provided at the end of a sequence of nine touches.

If participants touched an inappropriate location (incorrect trial), or failed to touch the screen within 5,000ms after the red circle's appearance (aborted trial), a green screen was displayed for 3000ms (baboons) / 1000ms (humans) as a marker of failure. Aborted trials were not counted as trials and were therefore presented again, while incorrect trials were not. The time elapsed between the appearance of the red circle and the participant's touch on this circle was recorded as the response time (RT). To learn the task, baboons initially received trials that were rewarded after one touch, after which the number of touches in a trial was progressively increased up to nine. Humans were simply instructed to touch the red circle as quickly as possible. They were also informed that each trial began by the presentation of the yellow cross at the bottom of the screen followed by a sequence of nine touches on the moving red circle.

To equate for the motor difficulty of the sequences to be produced, each participant was first tested on 5 blocks (baboons) / 4 blocks (humans) of 200 random trials (called "Random I phase"), each composed of a random ordering of nine locations, without immediate repetition. A baseline measure for all possible transitions from one location to another was computed by calculating mean RTs for each transition at the group level (i.e., the entire group of baboons on the one hand and of human participants on the other hand).

After this Random I phase, participants were exposed to trials including statistical regularities (hereafter, "Pattern phase"). Three independent regularities, each composed of three fixed locations were constructed ($R_1 = 'A_1B_1C_1'$; $R_2 = 'A_2B_2C_2'$; $R_3 = 'A_3B_3C_3'$). To study the extraction dynamic of these regularities, we carefully selected these three patterns so that the mean RTs for the initial transition ($'A_iB_i'$) and final transition ($'B_iC_i'$) would not be statistically different considering the baseline measurements obtained for these transitions during the random phase. By matching these *local* adjacent dependencies at the onset of the Pattern phase, it was possible to study the fine-grained evolution of these RTs that provide a direct behavioral index of regularity extraction.

Moreover, to study the extraction dynamic at a more complex and *global* level, the three regular patterns appeared in each sequence of nine touches, increasing the predictability of each pattern in the sequence (1/3, 1/2, 1). For example, if participants are exposed to the regular patterns R3 followed by R1, then the last regular pattern will be R2. If they manage to extract

such a complex regularity, we should then observe a decrease in RTs for the first element of the last displayed regular pattern (in our example, for A₂).

During the Pattern phase, participants performed 10 (baboons) / 7 (humans) blocks of 200 trials each, one trial being composed of a random combination of the three regular patterns³⁸. Finally, participants performed a last series of 5 (baboons) / 1 (humans) random blocks (“Random II phase”) that had the same structure as the Random I phase (see Figure 1).

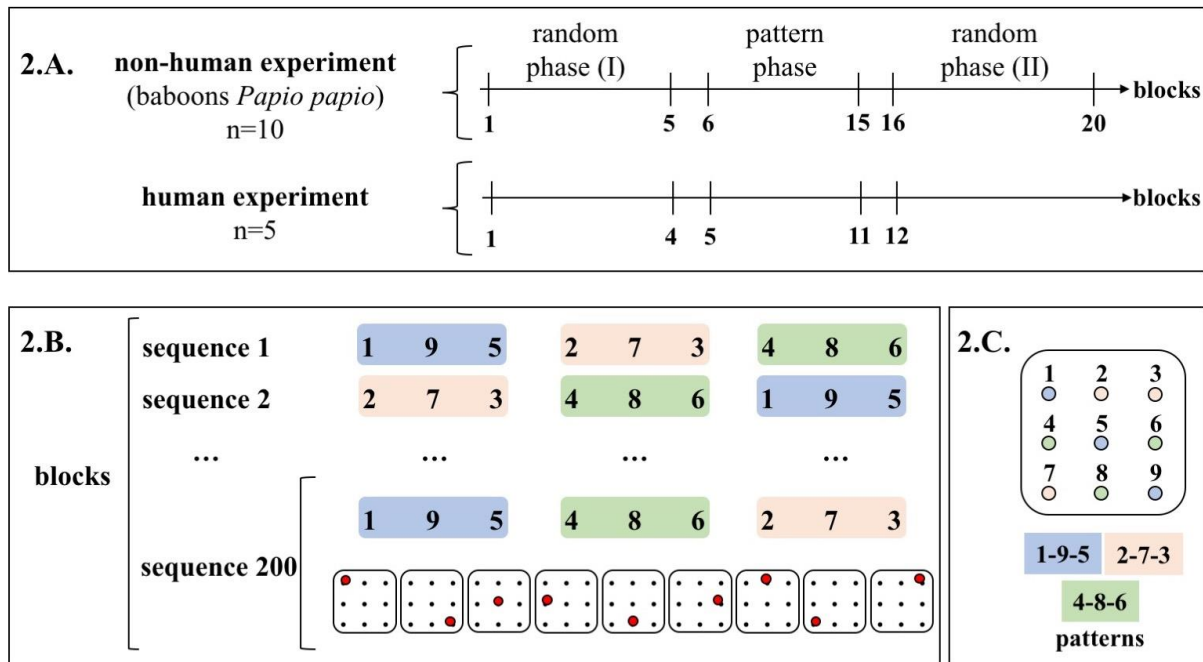


Figure 1 Experimental procedure for baboons and humans across the whole experiment (A). Experimental procedure at the block- and trial level (B). Patterns of locations on the touch screen used with humans (C).

Results

³⁸ The number of blocks in humans and baboons are not exactly the same, because if the baboons could do the task when they wanted, humans had to complete the blocks in one long session. That is the reason why the number of blocks for humans has been slightly reduced.

Incorrect trials were removed from the dataset (7.6% for baboons and 5.9% for humans), as well as RTs greater than two standard deviations from the mean (computed for each subject and each block, 4.2% for baboons and 3.6% for humans).

The first analysis concerned regularity extraction at the local level. Minier et al. (2015) reported an advantage of the second transition (i.e., BC) over the first one (i.e., AB) in baboons, with RTs on C decreasing faster than RTs on B. We therefore ran the same 2 (Transitions) * 7 (Block) repeated measures ANOVA on the mean RTs of humans. Mauchly's test indicated a violation of the sphericity assumption for the block factor ($\chi^2(44)=71.7$, $p<0.05$), so Greenhouse-Geisser estimates ($\epsilon=0.11$) were used to correct for degrees of freedom. We found a significant effect of Transition ($F(1,9)=14.9$, $p<0.01$, $\eta^2_p=0.62$), first transition RTs being slower than those of second transitions, and a significant effect of Block ($F(2.88, 25.9)= 36.5$, $p<0.001$, $\eta^2_p=0.80$) as the response latencies decreased through learning. Finally, the Transition * Block interaction was also significant ($F(9,81)=4.9$, $p<0.001$, $\eta^2_p=0.33$). Contrasts between first and second transitions for each block revealed a significant difference in blocks 1 to 8, but not in blocks 9 and 10. As shown in Figure 2, the human results match those of baboons with a faster decrease of RTs on C than on B.

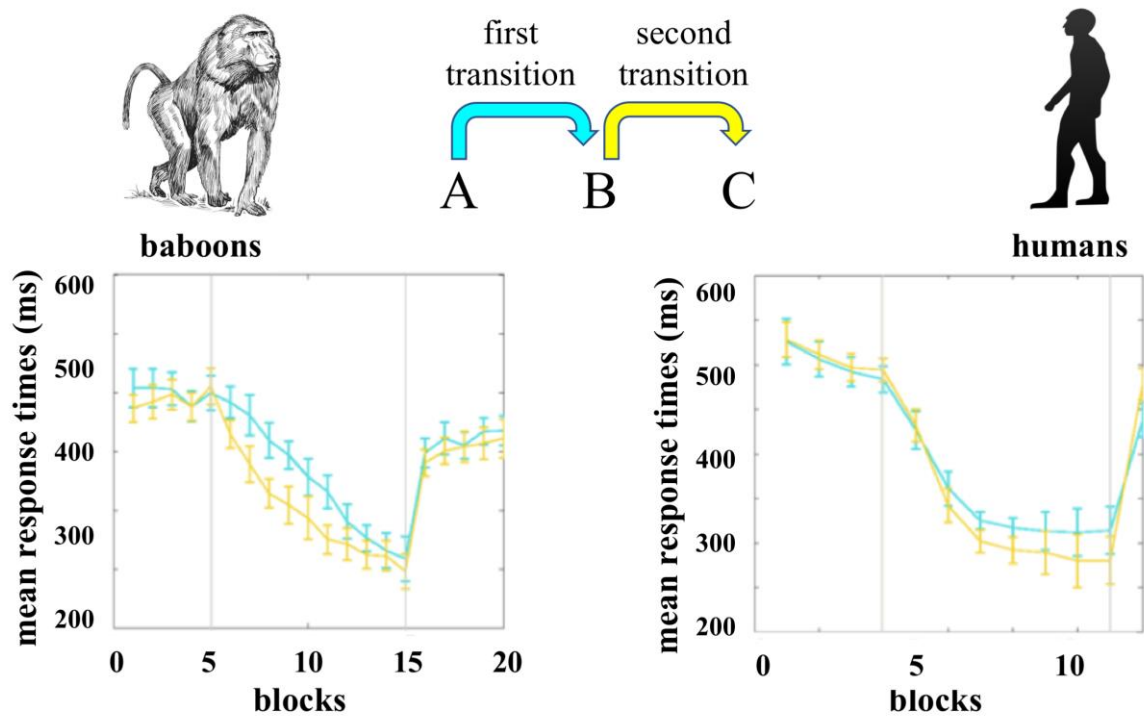


Figure 2 Evolution of mean RTs for the first (AB) and second (BC) transitions over blocks for baboons (on the left) and humans (on the right).

The second analysis concerned regularity extraction at a more global level, specifically, the predictability of the third pattern in the sequence given the first and second. As shown in Figure 3 (upper panel), the predictability within regular patterns is always equal to 1. Indeed, in each regular sequence, A_i was always followed by B_i that was always followed by C_i . However, the predictability of A varied across patterns depending on their position in the sequence of 9 touches. The first A of the first ABC regular pattern could occur with a probability of $1/3$, the second A with a probability of $1/2$, and the last A with a probability of 1. Of course, these probabilities depend on the ability of participants to extract the global statistical structure of the experiment.

We thus considered RTs for the three patterns on the three different positions in the sequence. Because humans and baboons did not receive the same number of blocks, we ran two different ANOVAs, both including Place in the pattern (first = A, second = B or third = C), Blocks (1 to 10 for baboons and 1 to 7 for humans) and Position of the patterns in the sequence (first, second or third) as within-subject factors.

For baboons, Mauchly's test indicated a violation of the assumption of sphericity for the Block factor ($\chi^2(44)=67.5, p<0.05$) as well as for the Position factor ($\chi^2(2)=6.3, p<0.05$), so Greenhouse-Geisser estimates ($\epsilon=0.31$ and $\epsilon=0.62$, respectively) were used to correct for degrees of freedom. We found a significant effect of Position ($F(1.23, 11.2)=8.0, p<0.013, \eta^2_p=0.47$), and a significant effect of Block ($F(2.8, 25.2)=3.7, p<0.001, \eta^2_p=0.29$). The interaction between Block and Position were found to be significant as well ($F(18,162)=1.8, p=0.026, \eta^2_p=0.17$). Helmert contrasts showed that the effect of Position was between the first position and the last two ($F(1,9)=8.9, p=0.02, \eta^2_p=0.5$) but not between the second position and the third ($F(1,9)=1.74, p=0.22, \eta^2_p=0.16$). Another ANOVA ran on the mean RTs of the first place of the patterns (i.e., A) with only the last learning block and Position as a within-participant factor with 3 levels (1st, 2nd and 3rd) showed a significant effect ($F(2,18)=4.9, p<0.05, \eta^2_p=0.35$). Helmert contrasts confirmed that this effect was significant between the first and two last positions ($F(1,9)=6.2, p=0.04, \eta^2_p=0.40$) but not between the second and the third positions ($F(1,9)=0.67, p=0.43, \eta^2_p=0.07$).

For humans, we found a significant effect of Position ($F(2,8)=36.83, p<0.001, \eta^2_p=0.90$), and a significant effect of Block ($F(6, 24)=7.0, p<0.001, \eta^2_p=0.64$). The interaction between Position and Block was also significant ($F(12,48)=13.57, p<0.001, \eta^2_p=0.77$). Helmert comparisons showed that the effect of Position was between the first position and the two last ($F(1,4)=51.5, p=0.02, \eta^2_p=0.98$), and between the second position and the third position ($F(1,4)=8.03, p=0.047, \eta^2_p=0.67$). A final ANOVA considered mean RTs of the first place in the patterns (i.e., A) in the last learning block only. Position was entered as within-participant factor with 3 levels (1st, 2nd and 3rd). This analysis showed a significant effect ($F(2,8)=36.7, p<0.0001, \eta^2_p=0.90$). Helmert contrasts indicated that this effect was significant between the first and two last positions ($F(1,4)=68.0, p<0.001, \eta^2_p=0.94$) and between the second and the third position as well ($F(1,4)=14.3, p=0.02, \eta^2_p=0.78$). The evolution of mean RTs for baboons and humans over each block, each position within the sequence, and each place within the regular pattern is provided in Figure 3 (lower panels). A clear difference between baboons and humans appears for the first element of the triplet when it occurred in Position 3 of the sequence (i.e., A₃): RTs decreased across learning blocks for humans while no decrease was observed for baboons.

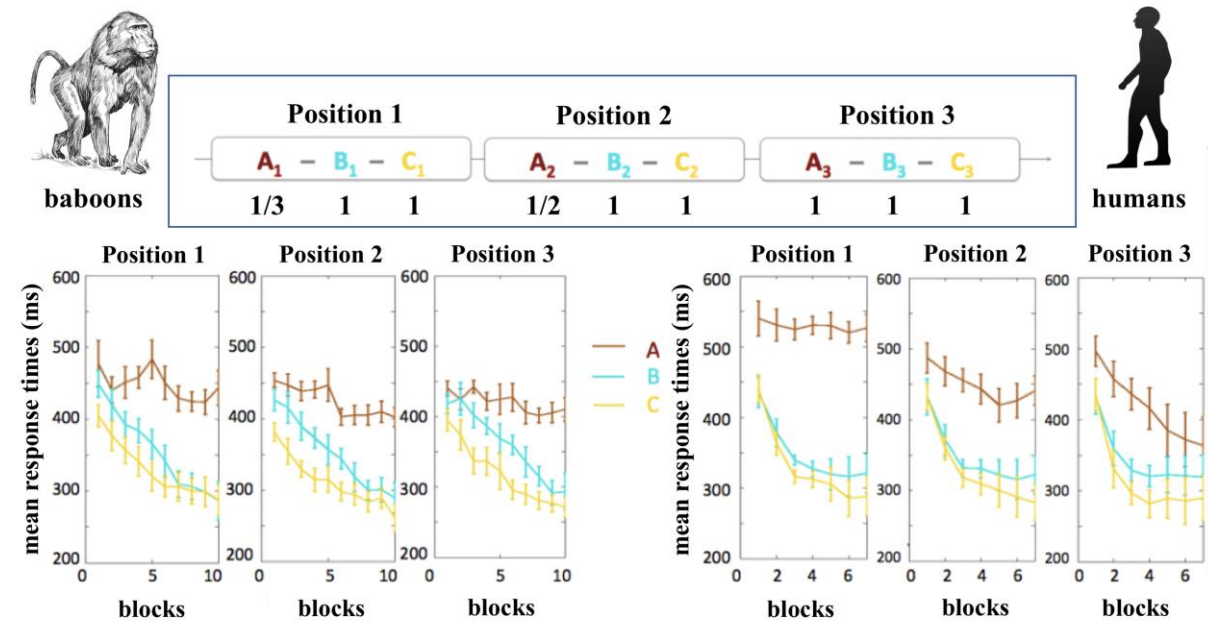


Figure 3 Upper central panel: Predictability of each location in the sequence of nine touches. Lower panels: Evolution of mean RTs over blocks (1-10 for baboons and 1-7 for humans), positions (1, 2, or 3) and places within the regular patterns (A, B, or C) for baboons (on the left) and humans (on the right).

Discussion

Two main findings were obtained in the present comparative study in which the performance of human adults and baboons were collected in the same regularity learning task. For regularities at the *local* level, human and non-human primates displayed a similar statistical learning dynamic: RTs decreased faster on the third location of a regular pattern (i.e., C) than on the second (i.e., B). At a more *global* level, we found different results for humans and baboons. In humans, RTs for the first element (i.e., A) of the third displayed regular pattern (i.e., the 7th element in the sequence, which should become predictable over the course of learning) decreased gradually across learning blocks while no such improvement was observed for baboons.

The first result, already reported for baboons by Minier et al. (2015), indicates that the two species had a similar dynamic of regularity extraction. In both species, RTs decreased faster on C than on B in ABC regular patterns, suggesting that it is a general property that should be captured by current computational models. As argued previously by Minier et al. (2015), this pattern of results is inconsistent with theoretical propositions resting on the notion of

transitional probabilities between adjacent elements. Indeed, in our task, the transitional probabilities from A to B and from B to C are all equal to one and are therefore insufficient to account for the present data. Similarly, the TRACX model by French et al. (2011) predicts that RTs should decrease faster on B compared to C given that the model is based on a left-to-right acquisition of statistical regularities. Only the SRN model (Elman, 1990) predicts the learning advantage of C over B. SRN assumes that learning is based on the development of associations between the successive elements of the input that are frequently and consistently encountered together. SRN encodes also the contextual information that precedes the successive elements, and therefore C benefits from richer contextual information (i.e., the preceding systematic sequence AB) while B benefits only from the systematic occurrence of A.

Our second result indicates that only humans capture the more complex regularity related to the predictability of the third regular pattern in the larger sequence. In that situation, following the logic that we proposed, the absence of a common behavioral pattern of results between human and non-human primates suggests that this kind of regularity cannot be extracted by mere associative learning mechanisms. Computational models are therefore not expected to account for this phenomenon that may require language recoding abilities.

Following a different line of reasoning, one may argue that the present results point to a possible limitation of the learning capacities in non-human primates. The extraction of these global regularities might indeed require a more complex learning mechanism or a larger memory span for preceding elements. One would need however to specify what is the precise nature of this learning mechanism and what would be the role of short-term memory span in facilitating the extraction of such global regularities.

Conclusions and avenues for future research

The comparative approach taken in the current paper suggests that the mechanisms underlying regularity extraction can be further understood by looking at patterns of performance that are common to human and non-human primates. Based on the present results we suggest that shared associative learning mechanisms can account for the extraction of *local* regularities (i.e., co-occurrences) and that a general theory of regularity learning could benefit from the large existing literature on associative learning (Mitchell & Le Pelley, 2010; Shanks, 1995). Moreover, the temporal dynamics of co-occurrence learning we observed in both species suggests a learning advantage for the final stimulus of a repeated (triplet) pattern, providing an

important empirical constraint for models of regularity extraction. Sensitivity to *global* regularities at the level of the entire sequence was observed for humans only and might call for a theory that accounts for the interaction between basic associative learning mechanisms and verbal recoding strategies.

Several important issues remain to be addressed to reach a general theory of associative learning mechanisms and their role for regularity extraction in different situations. We propose that the use of a comparative approach can be a fruitful avenue for future research to address a series of remaining theoretical questions, which we outline in more detail below.

A first series of questions concerns the role of the *timing* of co-occurrences in the sensory input. Suppose that A and B are two sources of information that co-occur systematically, A being always followed by B. Assuming no interfering information between A and B, how does the dynamic of regularity extraction change as a function of the delay between the presentation of A and the presentation of B? How many times A-B has to be repeated in order to observe an extraction of the co-occurrence? What is the role of the delay between two repetitions of A-B?

A second series of questions is related to the *sources of information surrounding* the A-B co-occurrence to be extracted. A general theory of regularity learning should be able to specify the role of factors such as the number and nature of interfering information between two repetitions of A-B, the role of the similarity between the interfering information and each element of the co-occurrence (A and B), the role of forward and backward inconsistencies (i.e., the role of the statistical co-occurrence of A with another source of information, e.g., C, and the interference of A-C on the extraction of A-B or, respectively, the co-occurrence of B with a preceding source of information, e.g., D, and the interference of D-B on the extraction of A-B), and finally, the effect of the presence of other co-occurrences (e.g., like E-F) appearing between two repetitions of A-B.

A third series of questions concerns the *number of co-occurring elements*. What is the regularity extraction dynamic when A-B is systematically followed by another information C (an issue that has started to be addressed in the present study)? Similarly, what will happen if we increase the chain of co-occurring elements? For ABCD sequences, for example, will we observe an even faster decrease of RTs on D compared to C (because of the richer context) or is the role of contextual information limited to a restricted time-window? And for longer sequences, what is the effect of repeating a previously encountered element (e.g., like in the sequence ABCAD)?

A fourth set of questions is related to the *learning of non-adjacent co-occurrences* (Gómez, 2002 ; Pacton & Perruchet, 2008 ; Pacton, Sobaco, & Perruchet, 2015; Wilson, Spierings, Ravignani, Mueller, Mintz, Wijnen, van der Kant, Smith, Rey, submitted). Under which conditions can we extract the co-occurrence of A and B when other sources of information are placed between A and B? Clearly, if co-occurrence learning is supported by associative learning mechanisms, the presence of interfering information between A and B will strongly decrease the probability to extract that regularity. However, baboons have shown some ability to process long-distance dependencies (Rey, Perruchet, & Fagot, 2012) indicating that the learning of non-adjacent regularities should certainly be accounted for by current computational models of regularity learning.

To conclude, we have argued that the domains of statistical learning and implicit learning, recently chunked under the term ‘implicit statistical learning’ (e.g., Christiansen, this issue), are fundamentally rooted in the domain of associative learning. We have also proposed an alternative way of using comparative studies to better comprehend the fundamental mechanisms supporting regularity extraction and the temporal dynamics of associative learning mechanisms. Rather than focusing on the differences between species, we suggest focusing on common patterns of performances to determine the empirical evidence that should constrain the elaboration of computational models. By employing online measures, we can generate fine-grained behavioral results that can inform models about the precise dynamics of regularity extraction mechanisms. We have suggested that this approach can be quite powerful in testing the predictions of computational models. Finally, we have outlined a list of questions that would need to be addressed to reach a deeper understanding of associative learning mechanisms in the context of regularity extraction.

Acknowledgments

This work, carried out within the Labex BLRI (ANR-11-LABX-0036) and the Institut Convergence ILCB (ANR-16-CONV-0002), has benefited from support from the French government, managed by the French National Agency for Research (ANR) and the Excellence Initiative of Aix-Marseille University (A*MIDEX). This research was also supported by the Premilang2 ANR-project (#ANR-13-BSH2-0002) and the Chunked ANR-project (#ANR-17-CE28-0013-02).

References

- Batterink, L. J. (2017). Rapid Statistical Learning Supporting Word Extraction From Continuous Speech. *Psychological Science*, 28, 921-928.
doi:10.1177/0956797617698226
- Batterink, L.J., Reber, P.J., Neville, H.J. & Paller, K.A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, 83, 62-78.
- Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, 121(1), 127-132.
- Chen, J., van Rossum, D., & ten Cate, C. (2015). Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Animal Cognition*, 18(1), 151-164.
doi:10.1126/science.284.5416.875a
- Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. *Trends in Cognitive Sciences*, 2(10), 406-416. doi:10.1016/S1364-6613(98)01232-7
- Christiansen, M. H. (this issue). Implicit Statistical Learning: A Tale of Two Literatures. *Topics in Cognitive Science*.
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science*, 17(10), 905-912.
- Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2), 179-211.
- Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, 57(2), 94-107. doi :10.1037/h0058559
- Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio papio*). *Behavior Research Methods*, 42(2), 507-516.
- Fagot, J., & Paleressompouille, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods*, 41(2), 396-404.
- Fagot, J., Malassis, R., Medam, T. & Montant, M. (2017). Comparing human and nonhuman animal performance on domain-general functions: Towards a multiple bottleneck scenario of language evolution. in *Origins of human language: continuities and splits*

- with nonhuman primates (pp 329-365)*. Editors: Louis-Jean Boë, Joël Fagot, Pascal Perrier, Jean-Luc Schwartz, Peter lang: Frankfurt.
- Fitch, W. T., and Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, *303*, 377-380.
- Franco, A., & Destrebecqz, A. (2012). Chunking or not chunking? How do we find words in artificial language learning? *Advances in Cognitive Psychology*, *8*(2), 144.
- Frank, M. C., Goldwater, S., Griffiths, T., & Tenenbaum, J. B. (2010). Modeling human performance in statistical word segmentation. *Cognition*, *117*, 107-125.
- French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A Recognition-Based Connectionist Framework for Sequence Segmentation and Chunk Extraction. *Psychological Review*, *118*(4), 614-636.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117-125. doi:10.1016/j.tics.2014.12.010
- Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*, 1569-1579.
- Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford, UK: Blackwell.
- Minier, L., Fagot, J., & Rey, A. (2015). The temporal dynamics of regularity extraction in nonhuman primates. *Cognitive Science*. doi:10.1111/cogs.12279
- Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*, 31. doi.org: 10.3389/fpsyg.2010.00031
- Mitchell, C. J., & Le Pelley, M. E. (Eds.). (2010). *Attention and associative learning: From brain to behaviour*. Oxford University Press, USA.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1-32.

- Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*, 80-96.
- Pacton, S., Sobaco, A., & Perruchet, P. (2015). Is an attention-based account of adjacent and nonadjacent dependency learning valid ? *Acta Psychologica*, *157*, 195-199.
- Perruchet, P. (2005). Statistical approaches to language acquisition and the self-organizing consciousness: A reversal of perspective. *Psychological Research*, *69*, 316-329.
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233-238.
doi:10.1016/j.tics.2006.03.006
- Perruchet, P., & Poulin-Charronnat, B. (2012). Word segmentation: Trading the (new, but poor) concept of statistical computation for the (old, but richer) associative approach. In Rebuschat, P., & Williams, J. N. (Eds.), *Statistical learning and language acquisition* (pp. 119-143). Berlin: De Gruyter Mouton.
- Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246-263.
- Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences*, *25*, 297–388. doi:10.1017/S0140525X02000067
- Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it?. *Cognition*, *95* (2), 201-236.
- Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, *133*(2), 227-244. doi: 10.1037/0033-2909.133.2.227
- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855-863. doi: 10.1016/S0022-5371(67)80149-X
- Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review*, *77*(6), 481-495. Doi: 10.1037/h0029964
- Rey, A., Perruchet, P., & Fagot, J. (2012). Centre-embedded structures are a by-product of associative learning and working memory constraints: Evidence from baboons (*Papio Papio*). *Cognition*, *123*(1), 180-184.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

- Santolin, C., Rosa-Salva, O., Regolin, L., & Vallortigara, G. (2016). Generalization of visual regularities in newly hatched chicks (*Gallus gallus*). *Animal Cognition*, *19*(5), 1007-1017. doi:10.1007/s10071-016-1005-2
- Shanks, D. R. (1995). *The Psychology of Associative Learning*. Cambridge: Cambridge University Press.
- Shanks, D. R. (2005). Implicit learning. In: Lamberts, K., and Goldstone, R., (eds.) *Handbook of Cognition*, pp. 202–220. London: Sage.
- Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior research methods*, *49*(2), 418-432. doi:10.3758/s13428-016-0719-z
- Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2017). Redefining “Learning” in Statistical Learning: What Does an Online Measure Reveal About the Assimilation of Visual Regularities? *Cognitive Science*. doi:10.1111/cogs.12556
- Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, *15*(4), 482-495. doi: 10.1111/j.1467-7687.2012.01141.x
- Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological bulletin*, *139*(4), 792-814. doi:10.1037/a0030801
- Toro, J. M., Nespors, M., & Gervain, J. (2016). Frequency-based organization of speech sequences in a nonhuman animal. *Cognition*, *146*, 1-7. doi:10.1016/j.cognition.2015.09.006
- Wang, L., Uhrig, L., Jarraya, B., & Dehaene, S. (2015). Representation of Numerical and Sequential Patterns in Macaque and Human Brains. *Current Biology*, *25*, 1966-1974.
- Wilson, B., Marslen-Wilson, W. D., & Petkov, C. I. (2017). Conserved sequence processing in primate frontal cortex. *Trends in Neurosciences*, *40*, 72-82. doi:10.1016/j.tins.2016.11.004

Résumé

Une hypothèse dominant actuellement les théories sur l'évolution des capacités syntaxiques est celle d'une spécificité humaine pour le traitement des grammaires supra-régulières. Cette hypothèse tend à être confirmée par les données comparatives actuellement disponibles, qui ne fournissent pas de démonstration non ambiguë de cette capacité chez une autre espèce que l'humain. Dans cette thèse, nous avons adopté une nouvelle approche consistant à étudier les prérequis cognitifs pour le traitement de ce type de grammaires. Notre hypothèse était que ces échecs précédents puissent être attribués à un biais chez ces espèces vers l'exploitation de régularités locales et des difficultés pour l'extraction de régularités plus distantes, plutôt qu'à une incapacité à traiter ces grammaires. Pour tester cette hypothèse, nous avons mené une série de quatre études chez le babouin de guinée (*Papio papio*) et l'humain. Dans une première étude, nous avons montré que le babouin nécessitait une quantité d'exposition beaucoup plus importante que l'humain pour apprendre des associations non-adjacentes. Dans une seconde étude, les babouins ont été capables de généraliser des patterns basés sur une répétition adjacente ou non-adjacente d'un élément, mais ils se sont montrés davantage sensibles aux répétitions adjacentes. Une troisième étude, corrélationnelle, révèle que les babouins parvenant se montrant sensibles aux régularités non-adjacentes dans différentes tâches ne sont pas ceux obtenant les meilleures performances pour l'apprentissage de dépendances adjacentes. Une dernière étude suggère que les babouins sont sensibles à une structure en miroir (impliquant des dépendances centrées-emboîtées), mais pas à une structure en copie (dépendances croisées). Cet ensemble de résultats met au jour une continuité des capacités syntaxiques au sein de la lignée des primates, mais révèle également des différences interspécifiques importantes dans les contraintes mnésiques pesant sur celles-ci.

Mots-clés : apprentissage de grammaires artificielles, régularités séquentielles, primates, cognition, psychologie comparée

Abstract

A current dominant hypothesis on the evolution of syntactic abilities propose that the processing of supra-regular grammars is a unique human capacity. In support of this hypothesis, artificial grammar learning studies conducted so far do not provide unambiguous demonstration of this capacity in a non-human species. In this thesis, we adopted a new approach by studying cognitive prerequisites for supra-regular grammar processing. Our hypothesis was that these previous failures could be attributed to a bias in these species towards the exploitation of local regularities and difficulties for processing more distant relationships, rather than an inability to master supra-regular grammars. We conducted a series of experiments in Guinea baboons (*Papio papio*) and humans to assess this hypothesis. In a first experiment, we show that baboons need much more exposure than humans to learn non-adjacent associations. In a second study, we show that baboons can generalize patterns involving an adjacent or a non-adjacent repetition of an element, but that they are more sensitive to the former. A third, correlational, study reveal that baboons succeeding to extract non-adjacent regularities are not those showing the best performance in learning local ones. A last study suggest that baboons are sensitive to a mirror structure (involving center-embedded dependencies), but not to a copy structure (crossed dependencies). Overall, our results reveal a stronger continuity in grammar processing capacities within the primate order than previously thought, but also highlight important species differences in memory constraints.

Keywords: artificial grammar learning, sequential regularities, primates, cognition, comparative psychology