



HAL
open science

Mécanismes d'apprentissage de nouveaux mots dans un contexte audiovisuel

Christine Nies

► **To cite this version:**

Christine Nies. Mécanismes d'apprentissage de nouveaux mots dans un contexte audiovisuel. Sciences de l'Homme et Société. 2020. dumas-03007337

HAL Id: dumas-03007337

<https://dumas.ccsd.cnrs.fr/dumas-03007337v1>

Submitted on 16 Nov 2020

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.

Mécanismes d'apprentissage de nouveaux mots dans un contexte audiovisuel

**Christine
NIES**

Sous la direction de Mathilde FORT et Anne VILAIN

Laboratoires : GIPSA-LAB et LPNC

UFR LLASIC
Département Sciences du langage et Français langue étrangère (FLE)

Mémoire de master 2 mention recherche - 30 crédits

Année universitaire 2019-2020

Mécanismes d'apprentissage de nouveaux mots dans un contexte audiovisuel

**Christine
NIES**

Sous la direction de Mathilde FORT et Anne VILAIN

Laboratoires : GIPSA-LAB et LPNC

UFR LLASIC
Département Sciences du langage et Français langue étrangère (FLE)

Mémoire de master 2 mention recherche - 30 crédits

Année universitaire 2019-2020

Remerciements

Je tiens à remercier tout d'abord mes directrices de mémoire, Mathilde Fort et Anne Vilain, de pour m'avoir permis de participer à ce projet passionnant, pour la confiance qu'elles m'ont accordée, ainsi que pour leurs disponibilités, leur bienveillance, les relectures et leurs recommandations. À Mathilde encore un merci tout particulier pour m'avoir accompagnée, rassurée et encouragée afin de mener à terme ce travail.

Merci à David Méary pour son précieux soutien lors de la mise en place de cette étude et pour sa grande patience lors de l'initiation à la technique de l'oculométrie.

Merci à Christophe Savariaux d'avoir apporté son expertise autant pour la préparation des stimuli et aussi dans la mise en place de ce projet.

Merci à tous les parents et enfants ayant participé à notre projet.

Merci également à Diana Gergova, pour son soutien et sa collaboration tout au long du projet.

Et enfin merci à tous, proches et amis, pour votre soutien moral et intellectuel ... et les relectures !

DÉCLARATION ANTI-PLAGIAT

1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

PRENOM : Christine

NOM : Nies

DATE : 28.07.2020

Sommaire

Remerciements	3
Introduction	6
PARTIE 1 - CADRE THEORIQUE	9
CHAPITRE 1. L'APPRENTISSAGE STATISTIQUE INCIDENT : UNE REPONSE INCOMPLETE POUR L'ACQUISITION DU LANGAGE	10
CHAPITRE 2. L'ATTENTION : UNE TACHE DIFFICILE D'UN POINT DE VUE DEVELOPPEMENTAL	14
CHAPITRE 3. L'ATTENTION AUX VISAGES PARLANTS : UNE STRATEGIE GAGNANTE POUR L'APPRENTISSAGE DE NOUVEAUX MOTS	17
1. Attention conjointe et suivi du regard : un outil pour résoudre le problème de la référence	17
2. Attention à la bouche du locuteur : un outil pour résoudre le problème de la représentation	18
CHAPITRE 4. HYPOTHESES POUR NOTRE ETUDE	20
PARTIE 2 - METHODOLOGIE	23
CHAPITRE 5. PARTICIPANTS ET ETHIQUE	24
CHAPITRE 6. QUESTIONNAIRE PARENTAL	25
CHAPITRE 7. STIMULI	27
1. Pseudo-mots	27
2. Enregistrement des stimuli	28
3. Objets	29
4. Contrebalancement	30
CHAPITRE 8. DISPOSITIF	31
CHAPITRE 9. PROCEDURE	32
CHAPITRE 10. PREDICTIONS EXPERIMENTALES	35
PARTIE 3 - ANALYSES ET RESULTATS	37
CHAPITRE 11. ANALYSES	38
1. Méthode analytique	38
2. Prétraitement des données et définition des aires d'intérêt (AOIs)	39
3. Nettoyage de données	40
CHAPITRE 12. RESULTATS	41
1. Phase de familiarisation	41
2. Phase d'apprentissage	42
3. Phases de test	44
4. Préférence pour la bouche et apprentissage	47
5. Meilleures performances et vocabulaire productif	49
6. Conclusion de la partie résultats	51
Discussion	52
Conclusion	60
Bibliographie	61
Table des illustrations	64
Table des annexes	65

Introduction

Les nourrissons apprennent leur langue avec une vitesse remarquable. Surtout pendant leur deuxième année de vie (« explosion du vocabulaire »). A l'âge de 24 mois ils connaissent environ 300-550 mots. Les mécanismes cognitifs à l'origine de ce processus restent peu décrits.

Les travaux de Saffran et ses collègues (1996) ont mis en évidence que notre système cognitif, dès la naissance, dispose d'un mécanisme d'apprentissage robuste. Ce mécanisme permet au nourrisson de trouver une structure et du sens dans les nombreuses informations qui lui parviennent lorsqu'il perçoit la parole. Par exemple, dans une situation expérimentale simple, pour apprendre le sens d'un mot, le nourrisson de 12 mois parvient à détecter la co-occurrence dans le temps et dans l'espace, d'un label sonore (e.g., « ballon ») et l'objet auquel il correspond (Smith & Yu, 2008).

Par contre, ce mécanisme d'apprentissage statistique « incident » ne peut expliquer l'acquisition du langage que de manière partielle. Les études sur l'apprentissage statistique ont été réalisées dans un contexte expérimental simplifié où la co-occurrence mot-objet est à elle seule un indice suffisant pour apprendre un mot prononcé par un seul locuteur. En revanche, dans leur vie quotidienne, les nourrissons doivent apprendre le langage dans des situations très riches sur le plan perceptif : les co-occurrences mot-objet sont complexes, et les locuteurs sont multiples, ce qui rend l'apprentissage de relations mot-objet beaucoup plus complexe que dans le contexte expérimental simplifié (Smith & Yu, 2013).

Puisque la déduction du sens d'un mot concret dans des situations de la vie quotidienne peut s'avérer être une tâche très complexe, des mécanismes cognitifs supplémentaires aux mécanismes d'apprentissage statistique seraient nécessaires pour apprendre un nouveau mot : les nourrissons doivent sélectionner (par des mécanismes attentionnels) et maintenir le temps de l'apprentissage (par des mécanismes de mémoire de travail) un sous-ensemble fini d'informations pertinentes à considérer. C'est sur la base de ce sous-ensemble qu'ils pourront ensuite réaliser leurs opérations statistiques, et en déduire l'association signifiant (mot)- signifié (objet).

Ces informations pertinentes à sélectionner pour l'apprentissage d'un nouveau mot sont présentes dans le signal de parole auditif et visuel ; en face-à-face avec un partenaire social. Les visages sont des stimuli très complexes et très riches sur le plan perceptif, qui ne fournissent pas que des informations linguistiques. Par exemple, les yeux donnent des

informations sur l'identité, le genre et l'état émotionnel du locuteur (Senju & Johnson, 2009). En revanche, la bouche fournit des informations auditives et visuo-articulatoires pertinentes. Regarder la bouche peut être une bonne stratégie pour améliorer la mémorisation des mots et des phonèmes qui les composent.

Des études ont montré que l'attention sélective aux visages parlants améliore considérablement la reconnaissance des mots chez les adultes (Fort et al., 2010, 2013). En revanche, les nourrissons semblent ne pas pouvoir profiter autant de cette multimodalité (Murray et al., 2016). Ce n'est qu'à partir de 8 à 12 mois que les structures corticales responsables de l'attention sélective commencent à être fonctionnelles, leur maturation se poursuivant jusque dans l'adolescence (Davidson et al., 2006; de Diego-Balaguer et al., 2016). Dès la deuxième année de vie, ce système va se perfectionner et se déclinier en différents mécanismes (e.g., la mémoire de travail). En conséquence, avant l'âge de 8 mois, les nourrissons passent beaucoup de temps à regarder la zone autour des yeux de leur interlocuteur (une région qui n'est pas importante pour l'encodage des mots et des phonèmes qui les composent). Il est intéressant de constater qu'ils dirigent leur attention vers la bouche du visage parlant au début de leur « explosion du vocabulaire », une stratégie qui pourrait booster leur performance d'apprentissage des mots (Fort et al., 2018; Lewkowicz & Hansen-Tift, 2012).

Plusieurs travaux ont mis en évidence le rôle important de l'attention aux visages parlants pour l'acquisition du langage. Suivre le regard permet à un nourrisson de diriger son attention visuelle vers les objets désignés et de mieux les mémoriser (Senju et al., 2006 ; Csibra, 2010). Se focaliser sur la bouche parlante de ses partenaires sociaux lui permet également de mieux encoder les phonèmes des mots de sa langue maternelle utilisée pour désigner ces objets (Lewkowicz et Hansen-Tift, 2012). D'un point de vue développemental, prêter attention aux yeux et à la bouche de son interlocuteur semble être une stratégie attentionnelle complémentaire pour acquérir des noms des objets dans sa langue maternelle. Ainsi, un nourrisson doit non seulement suivre le regard de son interlocuteur vers un objet et y maintenir le regard le temps de le mémoriser, mais il doit également apprendre à actualiser ses stratégies visuo-attentionnelles au cours du temps afin d'orienter et de maintenir le regard sur la bouche du locuteur lors de la prononciation du nom de l'objet.

Notre expérience vise à mieux comprendre les contraintes liées à l'immaturation du système visuo-attentionnel sur l'apprentissage de nouveaux mots. Plus précisément, nous nous intéressons au rôle que la multimodalité (le fait de prêter attention à différentes zones du visage d'un interlocuteur), et notamment l'attention à la bouche de l'interlocuteur, joue dans ce processus. Pour répondre à cette question, des nourrissons français au développement typique, âgés entre 20 et 28 mois, vont être testés sur une tâche d'apprentissage de nouveaux mots dans le laboratoire Babylab à Grenoble.

La première partie de ce travail sera consacré au cadre théorique de notre étude. Nous allons tout d'abord expliquer l'apprentissage statistique incident comme mécanisme d'apprentissage puissant mais incomplet, puis introduire des informations sur le développement de l'attention et parler de l'attention aux visages parlants comme stratégie gagnante pour l'apprentissage des mots (suivi regard, bouche). Ensuite, nous présenterons nos hypothèses pour le projet. La seconde partie présentera la méthodologie que nous avons suivie et nous exposerons la méthodologie d'analyse et les résultats obtenus lors de notre étude dans la troisième partie. Enfin, dans le dernier chapitre, nous discuterons les résultats en les mettant en parallèle avec ceux des différentes études menées précédemment. Nous terminerons avec la conclusion.

Partie 1

-

Cadre théorique

Chapitre 1. L'apprentissage statistique incident : une réponse incomplète pour l'acquisition du langage

S'il existe des études consensuelles (cf., Bergelson & Swingley, 2012; Vihman, 1996) sur les trajectoires du développement du langage, répondant aux questions « qu'est-ce qui est acquis et à quel moment », les connaissances sur les mécanismes cognitifs de l'apprentissage de la parole restent moins explorées.

Depuis une vingtaine d'années, de nombreux chercheurs se sont penchés sur un mécanisme d'apprentissage démontrant « comment » le langage est appris chez les nourrissons : l'apprentissage statistique (cf., Gomez & Gerken, 1999; Johnson & Tyler, 2010; Mintz, 1996; Saffran et al., 1996). Les travaux princeps sur cet apprentissage statistique du langage de Saffran et ses collègues (1996) ont démontré que, très tôt, le nourrisson est doté d'un mécanisme puissant pour détecter les différentes régularités dans son environnement linguistique. Ces auteurs ont montré que pour segmenter le flux de parole en mots, le nourrisson utilise les indices statistiques de co-occurrences entre syllabes. En effet, dans une langue donnée, certains enchaînements de syllabes sont plus probables que d'autres. Par exemple, dans la langue française, la séquence « gâteau » est traitée comme un mot distinct puisque la probabilité que *gâ* soit suivi de *teau* est élevée, tandis que la probabilité que *gâ* soit précédé ou *teau* suivi d'une syllabe donnée est relativement faible. Ces probabilités sont en lien avec le discours qui est présenté aux nourrissons dans leur vie quotidienne où de nombreux mots ou syllabes finales peuvent survenir avant le mot *gâteau* (e.g., *le gâteau*, *un gâteau*, *délicieux gâteau*, etc.) et de nombreux mots et syllabes initiales peuvent survenir suite au mot *gâteau* (e.g., *gâteau-là*, *gâteau au chocolat*, *gâteau pour toi*, etc.). L'enfant mémoriserait les co-occurrences entre les syllabes les plus fréquentes (par ex., *gâ* suivi de *teau*) comme mot distinct (Babineau & Peperkamp, 2017).

Dans leur expérience, Saffran et ses collègues (1996) ont mis en évidence l'importance de cet apprentissage statistique dans la segmentation de la parole en exposant des nourrissons de huit mois à un signal de parole artificielle. Ce signal a été composé de quatre mots trisyllabiques (« *pabiku* », « *tibudo* », « *tudaro* », « *pigola* »), présenté pendant deux minutes sans pause entre les mots et dans un ordre aléatoire. Les seuls indices de frontières de mot qui étaient présents dans ce flot continu étaient les probabilités

transitionnelles entre les paires de syllabes. Ces probabilités étaient de 1,0 pour les paires intra-mots (e.g., entre « pa » et « bi ») et 0,33 pour les paires inter-mots (e.g., entre « ku » et « ti »). Pour une illustration du déroulé de cette expérience (voir tableau 1).


Familiarisation	
...pabikutibudotudaropigola... 	
1.0 0.33	
Test	
Mot	Semi-mot
pabiku	kutibu

Tableau 1 : Illustration du déroulé de l'expérience de Saffran et al. (1996).
Source : Babineau et Peperkamp, 2017.

Lors de la phase de test, les nourrissons ont eu un temps d'écoute plus long pour les « mots partiels » comme « kugola » ce qui indique qu'ils ont mémorisé les « mots » auxquels ils ont été exposés au préalable et qu'ils étaient surpris d'entendre d'autres séquences en isolation. Ainsi, 2 minutes suffisent aux bébés de huit mois pour extraire des informations sur les statistiques séquentielles des syllabes afin de segmenter le signal de parole en mots. Ces résultats étayaient l'hypothèse que très tôt, les bébés peuvent regrouper des syllabes à partir de régularités statistiques (probabilité de transition entre syllabes) présentes dans le signal de parole.

La puissance et la complexité de ce mécanisme a continué à intéresser les chercheurs depuis le travail de Saffran et ses collègues. Les recherches effectuées sur ce mécanisme statistique ont pu démontrer que les nourrissons sont sensibles aux co-occurrences de différents éléments du langage. L'apprentissage statistique est un mécanisme incident qui mène les nourrissons bien plus loin que de segmenter simplement le discours en leur permettant de saisir les différents niveaux d'organisation d'une langue, tels que les sons, les mots et la syntaxe (Babineau & Peperkamp, 2017).

Bien que ce mécanisme d'apprentissage « incident » soit puissant, il ne permet d'expliquer l'apprentissage de nouveaux mots par le cerveau humain que de manière partielle. En effet, l'ensemble des preuves expérimentales en faveur de l'apprentissage statistique du langage a été obtenu en laboratoire, dans des situations simplifiées où la co-occurrence mot-objet est à elle seule un indice suffisant pour apprendre un mot prononcé par un seul locuteur. Or, dans la vie réelle, l'enfant doit apprendre le langage dans des situations très riches sur le plan perceptif : les co-occurrences mot-objet sont complexes et les locuteurs sont multiples, ce qui rend l'apprentissage de relation mot-objet beaucoup plus complexe que dans un contexte expérimental simplifié (voir figure 1) (Smith & Yu, 2013)

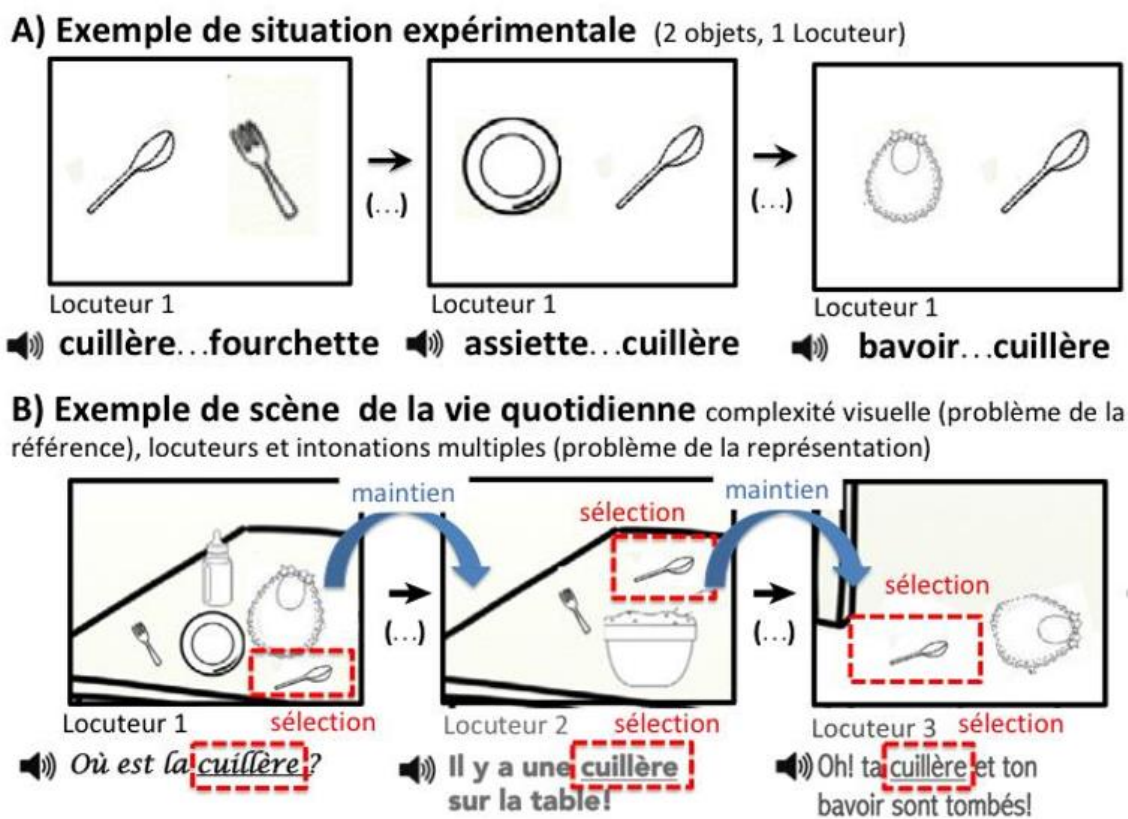


Figure 1 : A. Représentation schématique de trois essais non consécutifs dans le temps, dans une situation expérimentale simplifiée (mot prononcé par un seul locuteur et l'objet auquel il correspond). B. Représentation schématique d'une séquence de situations non consécutives du monde réel comme possiblement vu et entendu par un jeune enfant (co-occurrences mot-objets complexes et locuteurs multiples). Adapté de (Smith et al., 2014).

Pour donner un exemple, les travaux de Smith et Yu (2012, 2016) ont mis en évidence que la simple co-occurrence mot-objet n'est pas suffisante pour en apprendre l'association et que la complexité de la scène visuelle des enfants influence directement ce processus. Lors d'une tâche interactive d'apprentissage des nouveaux mots, l'enfant de 18 mois parvient à apprendre l'association mot-objet lorsque l'objet est nommé par l'adulte au moment où il domine la scène visuelle de l'enfant. En revanche, il ne réussit pas à apprendre cette association lorsque d'autres objets occupent la scène visuelle.

Pendant l'apprentissage d'un nouveau mot, le nourrisson se retrouve confronté à deux complexités : le problème de la référence et le problème de la représentation.

Cette première complexité a été formulée par Quine, comme **le problème de la référence** (Quine, 1961). D'après cette idée, l'enfant doit déduire qu'un mot fait référence à un objet (i.e., le signifié) tout en ignorant d'autres co-occurrences liées aux caractéristiques de l'objet (e.g., sa couleur, sa forme, etc.) ou à d'autres objets qui co-occurrent parallèlement. Les travaux de Smith et Yu (2012, 2016) mentionnés plus haut en donnent un exemple.

La deuxième complexité se trouve dans le signal acoustique dans lequel l'enfant doit trouver des mots utilisés pour désigner des objets (i.e., le signifiant). Cela implique de distinguer les sons qui le composent et de les maintenir en mémoire. Pour cela, l'enfant doit ignorer un grand nombre de dimensions non pertinentes présentées dans le signal acoustique comme des bruits parasites environnants ou des caractéristiques acoustiques propres à un locuteur (e.g. timbre, gravité de la voix). Cette complexité a été formalisée comme **problème de la représentation** (Kuhl, 2000).

Compte tenu du nombre de paramètres à considérer, parvenir à déduire le sens d'un mot concret dans des situations de la vie quotidienne peut s'avérer une tâche très complexe. Par conséquent, nous proposons dans notre étude que des mécanismes cognitifs supplémentaires aux mécanismes d'apprentissage statistique devraient être nécessaires pour apprendre un nouveau mot : les nourrissons doivent sélectionner, par des **mécanismes attentionnels**, et maintenir le temps de l'apprentissage, par des **mécanismes de mémoire de travail**, un sous-ensemble fini des dimensions pertinentes à considérer. C'est sur la base de ce sous-ensemble qu'ils pourraient réaliser ensuite leurs opérations statistiques, et en déduire l'association signifiant (mot) - signifié (objet).

Chapitre 2. L'attention : une tâche difficile d'un point de vue développemental

Le chapitre précédent montre que pour apprendre le langage l'enfant doit être attentif. L'attention, représentant un processus cognitif complexe, fait référence à sa capacité à se centrer sur une personne, un objet ou une activité pendant un certain temps. Les informations pertinentes à sélectionner pour apprendre un nouveau mot se présentent dans le signal de parole auditif et visuel. Exploiter la richesse d'informations multimodales peut s'avérer être une tâche difficile compte tenu de l'immaturation du système de l'attention chez les nourrissons.

Dans cette partie, nous présenterons l'état des connaissances sur les étapes clés du développement de l'attention en nous basant sur un des modèles du système attentionnel les plus influents développé par Posner et ses collaborateurs (Petersen & Posner, 2012). Dans leur modèle ils suggèrent l'existence de deux systèmes attentionnels différents : un système d'attention « ascendante », étant un système d'orientation dirigé par la saillance d'un stimulus externe, et un système d'attention « descendante » combinant un système d'attention soutenue et un système exécutif qui sont dirigés par les besoins d'une tâche.

Le système d'attention « ascendante » permet à un individu d'orienter son attention vers un endroit précis dans l'espace et dans le temps suite à un changement de saillance du stimulus extérieur (e.g., un bruit, un mouvement, etc.). Ce système se caractérise plutôt par un traitement des informations de nature globale (e.g., la saillance perceptive des stimuli), holistique, que par un traitement local centré sur des éléments. A cet égard, l'attention se concentre davantage sur les choses qui stimulent le sens. Son fonctionnement assuré par des structures sous-corticales atteint sa maturité à l'âge de 6 mois (Colombo, 2001; Gillet & Barthélémy, 2011).

En modalité auditive, cela correspond par exemple à la détection du signal de parole dans le bruit environnant tels que les signaux de parole concurrents ou les bruits de fonds. Il permet au nourrisson de s'orienter de manière automatique vers la source sonore la plus saillante (e.g., celle avec le plus d'intensité, ou celle avec le plus de changement d'intonation) dans son environnement perceptif. Cette source sonore correspond souvent à la parole qui lui est adressée (Newman, 2005). En modalité visuelle, cela correspond par exemple au fait que l'attention est tirée vers un stimulus à cause de sa saillance (i.e., identité-référence, forme,

couleur, taille relative, etc.) ou parce qu'il domine la scène visuelle de l'enfant (i.e., est présenté seul et visuellement proche) (Yu & Smith, 2012).

Le système d'attention « descendante » est directement lié à la capacité de sélectionner l'information pertinente pour une tâche donnée et de maintenir et manipuler cette information à l'esprit. Il est constitué d'un système d'attention soutenue et d'un système exécutif. Le système de l'attention soutenue est relié à la capacité à maintenir (sur une certaine durée) un certain état d'alerte permettant d'effectuer des efforts cognitifs importants (Colombo, 2001; Richards et al., 2010). Cette capacité permet au nourrisson d'accorder plus d'importance à certains éléments dans son environnement.

En modalité auditive, cela correspond par exemple à la capacité de se focaliser sur une source sonore tout en ignorant les autres. En modalité visuelle, il correspond par exemple à la capacité d'un nourrisson à maintenir l'attention sur une cible quand les distracteurs sont présents ou de tenir compte d'une des dimensions d'un stimulus (e.g., sa couleur, forme, etc.) tout en ignorant les autres.

Cette attention sélective va faciliter l'encodage et la mémorisation de cette cible. Ce système commence à être actif à partir de la deuxième partie de la première année de vie (Colombo, 2001). C'est à partir de la fin de la première année de vie (8 à 12 mois) que les structures corticales responsables du système exécutif commencent à être fonctionnelles. Le système exécutif correspond à trois grands mécanismes : la mémoire de travail, l'inhibition et la flexibilité attentionnelle.

La mémoire de travail sert par exemple à maintenir (temporairement) une information à l'esprit et de manipuler celle-ci par la suite. L'inhibition aide par exemple à empêcher les comportements inadéquats et à mettre à l'écart des événements (stimuli ou attributs) ou des réponses non pertinentes. Ce mécanisme permet au nourrisson de contrôler ses actions face aux situations qui surviennent dans son environnement. Il se développe grandement entre 3 et 7 ans (Reynolds & Richards, 2009).

La flexibilité attentionnelle sert par exemple à sélectionner l'information pertinente pour une tâche donnée et à changer d'une tâche à une autre. La maturation du système exécutif se poursuivant jusque dans l'adolescence (Davidson et al., 2006).

La figure 2 représente les différents systèmes attentionnels de la naissance jusqu'à l'âge adulte. Le système d'orientation est désigné par le système attentionnel « dirigé par la saillance du stimulus », mettant en exergue que les facteurs impliqués dans ce système sont présents dans le stimulus extérieur (système ascendant). Il renvoie à la notion de « saillance perceptive ». Le système d'attention soutenue et le système exécutif seront regroupés sous le terme général de système attentionnel « dirigé par les besoins de la tâche » (système descendant). Il renvoie à la notion de « pertinence informationnelle ». Il est directement relié à la capacité à reconnaître une ou plusieurs informations pertinentes pour la tâche langagière à effectuer, mais également à inhiber les informations non pertinentes pour cette même tâche.

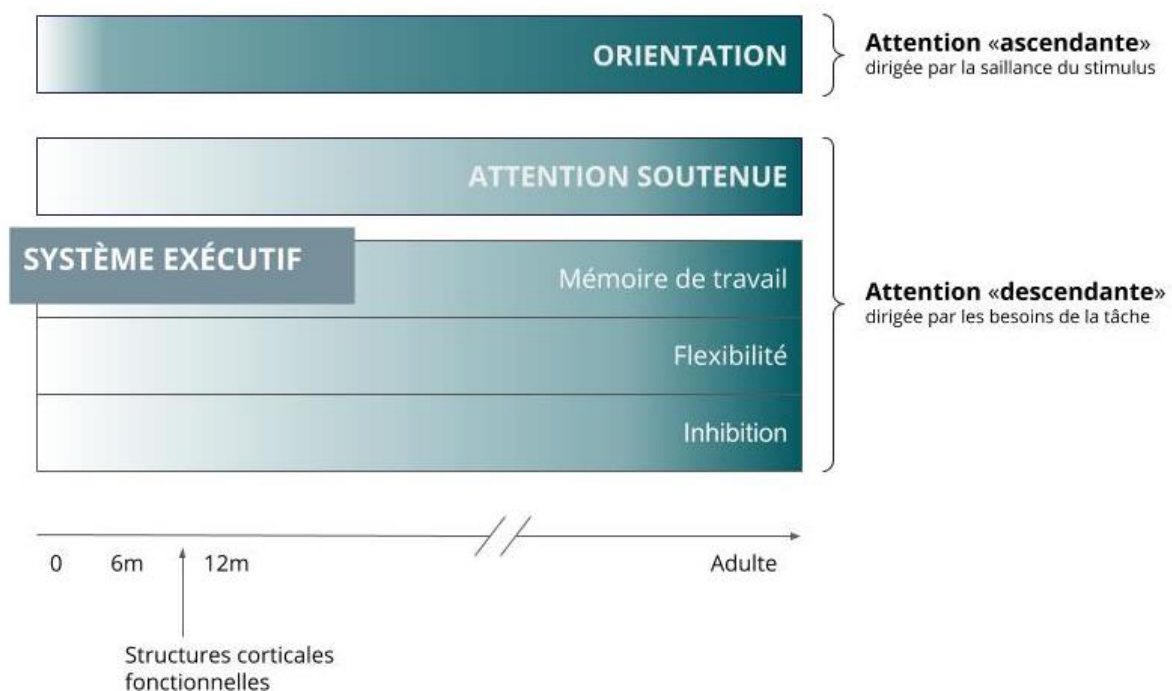


Figure 2 : Représentation schématique de la maturation des différents systèmes attentionnels de la naissance jusqu'à l'âge adulte. Les valeurs d'âge sont présentées en mois. La maturation est représentée par un dégradé de blanc (immature) à turquoise (mature).

Pour conclure, les processus attentionnels sont nécessaires pour percevoir, agir et donner un sens aux événements. Leur efficacité est notamment modulée par des facteurs liés à l'âge et la maturation cérébrale.

Dans notre étude, nous nous focaliserons sur le rôle de l'attention visuelle dans l'apprentissage de nouveaux mots, plus précisément sur le rôle de l'attention aux visages parlants.

Chapitre 3. L'attention aux visages parlants : une stratégie gagnante pour l'apprentissage de nouveaux mots

Pour apprendre un nouveau mot, le nourrisson doit sélectionner les dimensions pertinentes dans le signal de parole. Dans son environnement naturel, il perçoit ce signal dès sa naissance en face à face avec ses partenaires sociaux. Cela signifie qu'il entend non seulement la voix (parole auditive) mais aussi qu'il peut voir le visage de son interlocuteur (parole visuelle).

De nombreux travaux ont montré l'importance de l'information auditive quant à la perception de la parole, mais également que l'information visuelle joue un rôle vital. De multiples aspects facilitent grandement l'acquisition du langage chez le nourrisson : ce sont surtout l'attention conjointe et le suivi du regard ainsi que l'attention à la bouche du locuteur qui semblent être des stratégies gagnantes afin de trouver les informations pertinentes dans le signal de parole.

1. Attention conjointe et suivi du regard : un outil pour résoudre le problème de la référence

Des études avec des populations adultes ont montré que l'information contenue dans les yeux et le regard d'un interlocuteur permet de comprendre ses émotions (Everdell et al., 2007), ainsi que ses intentions communicatives (Csibra & Gergely, 2011) pour les visages non parlants.

Dans le cadre de l'apprentissage d'un nouveau mot, suivre le regard de son interlocuteur vers l'objet nommé est un outil puissant car il permet au récepteur du message de diriger son attention visuelle vers l'objet pertinent. Cela permet de réduire drastiquement la quantité d'informations à considérer. Or, une préférence visuelle générale pour les yeux des visages non-parlants de ses interlocuteurs est observée dès la naissance et perdure dans la suite de son développement (Senju & Johnson, 2009). Dès l'âge de 6 mois, le nourrisson est capable de diriger son attention sur un objet en se basant sur le regard d'autrui (Senju & Csibra, 2008), ce qui est appelé « attention conjointe ». La capacité de cette attention conjointe peut être expliquée par le fait de la saillance perceptuelle des yeux humains : ils représentent une zone de fort contraste par rapport au reste du visage avec la pupille (« noir » de l'œil) et la sclère (« blanc » de l'œil) (Senju & Johnson, 2009). Elle est assurée par le système d'attention ascendante (voir chapitre précédent).

Cette capacité précoce à contrôler son attention visuelle en s'appuyant sur le regard de l'adulte comme référence se perfectionne au cours du développement. Elle va faciliter la formation d'association mot-objet dans des tâches d'apprentissage de mots avec des visages non parlants à des étapes ultérieures du développement (Gliga et al., 2012). Dans leur étude avec des enfants, âgés de trois ans, Gliga et ses collègues ont montré que l'attention conjointe est, entre autres, un atout nécessaire pour apprendre la nouvelle association mot-objet.

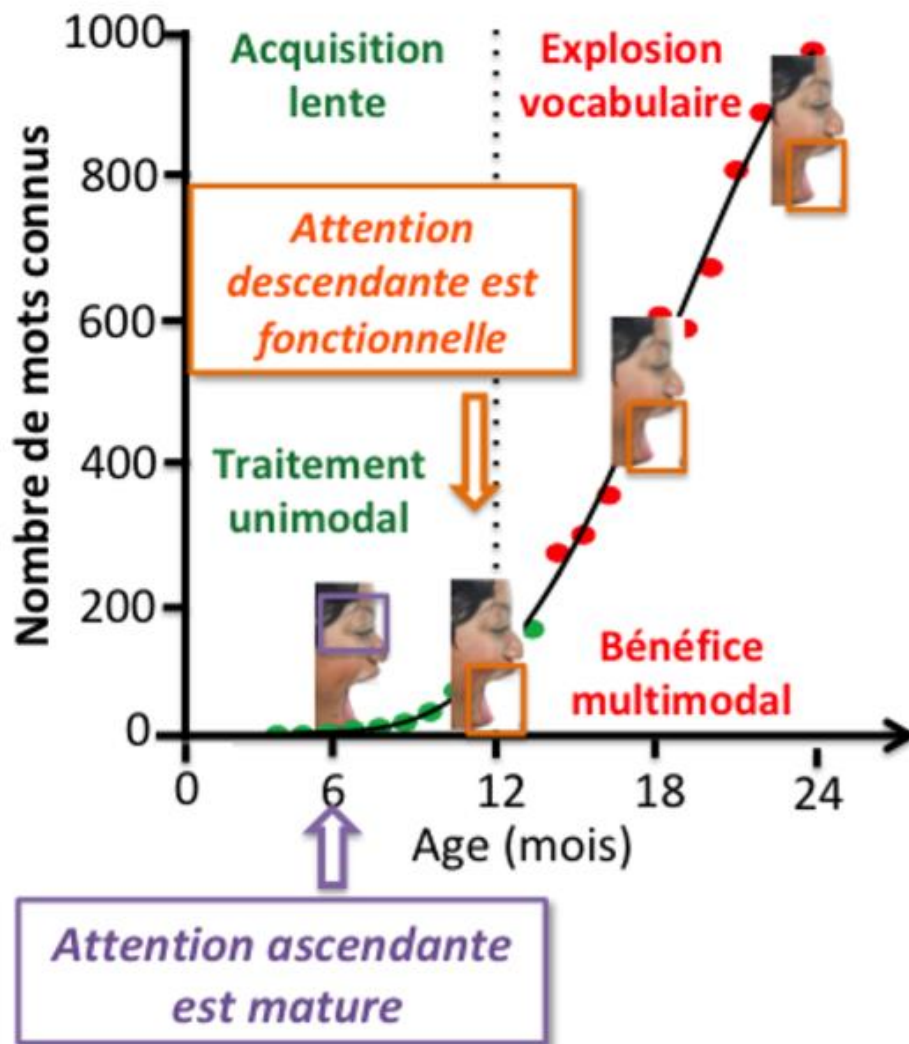
2. Attention à la bouche du locuteur : un outil pour résoudre le problème de la représentation

Se focaliser sur la bouche parlante de son interlocuteur permet à l'adulte d'améliorer l'intelligibilité d'un signal de parole audiovisuel, notamment lorsque ce signal n'est pas facile à comprendre (e.g., une langue inconnue) (Barenholtz et al., 2016). De même chez les nourrissons, la parole visuelle constitue une information redondante avec l'information acoustique, facilitant la mémorisation de syllabes (Gogate & Bahrick, 2001) et la compréhension de mots (Gogate et al., 2001).

Or, la préférence pour la bouche des visages parlants n'est pas présente dans le début du développement. En effet, jusqu'à l'âge de 7 mois, les nourrissons ne peuvent s'empêcher de fixer les yeux du locuteur, alors même que la bouche produit de la parole et donc produit des mouvements par rapport aux yeux qui restent fixes. Cette préférence visuelle initiale pour les yeux de l'interlocuteur peut s'expliquer par le fait que le système attentionnel descendant du nourrisson ne commence pas à être fonctionnel avant l'âge de 8 mois (Méary et al., 2018). C'est pour cette raison que la capacité de focalisation sur la bouche d'un visage dynamique parlant ne peut s'établir qu'à partir de ce moment-là (Csibra, 2010).

Des travaux récents montrent que les nourrissons dirigent leur attention vers la bouche du visage à partir du moment du début de leur « explosion du vocabulaire » (autour de 8 à 12 mois) (voir figure 3). La bouche fournit des informations auditives et visuo-articulatoires pertinentes. De ce fait, regarder la bouche parlante de ses partenaires sociaux permet au nourrisson de mieux encoder les phonèmes des mots de sa langue maternelle (Lewkowicz & Hansen-Tift, 2012) et aide ainsi à résoudre le problème de la représentation.

Figure 3 : Représentation schématique de la maturation de l'attention ascendante et descendante en relation avec le développement lexical chez le nourrisson avec l'âge en mois (abscisse) et nombre de mots connus (ordonnée).



De leur naissance jusqu'à la fin de la première année de vie le développement lexical est lent. Il repose sur un traitement du signal de parole majoritairement unimodal et reste largement dominé par le système d'attention ascendante (sensible à la saillance perceptive des stimuli). Dès lors que le système attentionnel descendant commence à être fonctionnel (à partir de 8 mois), le nourrisson oriente majoritairement son attention vers la bouche des visages parlants. Ce n'est qu'à partir de ce moment qu'il leur est possible de bénéficier pleinement d'un signal de parole multimodal, leur permettant d'accélérer drastiquement le rythme auquel ils acquièrent les mots de leur langue (explosion du vocabulaire).

De plus, cette préférence est d'autant plus marquée chez le nourrisson grandissant dans un environnement bilingue (Pons et al., 2015). Alternativement, un déficit de regard à la

bouche des locuteurs est retrouvé chez des enfants avec troubles spécifiques du langage (Pons et al., 2018). Cela suggère que se focaliser sur cette information visible et redondante, additionné à la parole auditive, facilite l'apprentissage du langage, et ce jusqu'à des stades avancés de l'enfance (Irwin et al., 2017).

Chapitre 4. Hypothèses pour notre étude

Notre expérience vise à mieux comprendre les contraintes liées à l'immaturation du système visuo-attentionnel et des mécanismes de mémoire du travail sur l'acquisition du langage. Plus précisément, nous nous intéressons au rôle que la multimodalité (e.g., regarder la bouche du locuteur) joue dans ce processus.

Des études antérieures ont montré qu'il existe un lien entre les stratégies générales d'attention visuelle (e.g., regarder un objet lorsqu'il est nommé) et l'apprentissage de nouveaux mots. Néanmoins, peu de travaux ont pu directement mettre en lien les explorations attentionnelles des visages avec les performances en apprentissage de mots. L'étude de Tenenbaum et collègues (2015) a pu démontrer l'existence d'une corrélation entre la préférence visuelle pour le visage d'un locuteur et la taille du vocabulaire (mesurée à l'aide d'un questionnaire parental). L'attention portée sur la bouche et le suivi du regard à l'âge de 12 mois seraient des prédicteurs pour le futur vocabulaire productif (18 et 24 mois) du nourrisson.

Dans un premier temps, notre étude vise à mettre en lien les explorations attentionnelles des visages, et plus précisément le regard porté sur la bouche, avec l'apprentissage de nouveaux mots. Le nourrisson apprend sa langue en face à face avec des partenaires sociaux. Les visages sont des stimuli très complexes, qui ne fournissent pas que des informations linguistiques. Par exemple, les yeux donnent des informations sur l'identité, le genre et l'état émotionnel du locuteur (Senju & Johnson, 2009). En revanche, la bouche fournit des informations auditives et visuo-articulatoires pertinentes. Regarder la bouche peut ainsi être une bonne stratégie pour améliorer la rétention des mots et des phonèmes qui les composent. En ce sens, pour une première hypothèse, nous supposons que le fait de regarder la bouche d'un visage parlant pendant la prononciation de nouveaux mots contribue à un meilleur apprentissage de ces mots.

En lien avec cette question, notre étude vise à mieux comprendre les contraintes liées

à l'immaturation des mécanismes cognitifs (système attentionnel et mémoire de travail) sur l'acquisition du langage. Afin de sélectionner des dimensions pertinentes dans le signal de parole (auditif et visuel) les mécanismes attentionnels sont nécessaires. Ces mécanismes permettent au nourrisson d'orienter son attention sur les yeux et la bouche du locuteur à des moments pertinents : la désignation de l'objet à apprendre (yeux) et la prononciation du mot cible à mémoriser (bouche).

Outre les mécanismes attentionnels, ce sont les mécanismes de mémoire de travail qui jouent un rôle important, permettant au nourrisson de maintenir les nouveaux mots en mémoire durant le temps de l'apprentissage. Si un nourrisson peut profiter lors de son apprentissage de nouveaux mots des informations multimodales fournies par le visage, cela dépend de la maturité de son système d'attention descendante. Comme nous l'avons abordé dans les chapitres précédents, ce système commence à être fonctionnel qu'à partir de l'âge de 8 mois (Méary et al., 2018). Étant donné que les mécanismes cognitifs (attention et mémoire de travail) sont primordiaux quant à l'apprentissage d'un nouveau mot, notre deuxième hypothèse est que l'immaturation de ces mécanismes cognitifs peut altérer ce processus.

Pour étudier ces hypothèses, des nourrissons entre 20 et 28 mois seront testés dans une tâche d'apprentissage de nouveaux mots. La technique de l'oculométrie nous permettra d'étudier leurs explorations attentionnelles du visage parlant pendant cette tâche d'apprentissage. Cela nous permettra également d'évaluer les compétences cognitives du nourrisson dans cette tranche d'âge. Plus précisément, nous allons mesurer le temps du regard porté à la bouche du locuteur lui apprenant de nouveaux mots et le mettre en rapport avec ses résultats dans une phase de test immédiate ainsi qu'avec ses capacités langagières. Les capacités langagières (la taille du vocabulaire) seront mesurées par un questionnaire parental.

Pour pouvoir mettre les analyses des explorations attentionnelles du visage en lien (indirect) avec l'immaturation des mécanismes cognitifs, nous allons intégrer des paramètres pour tester la mémoire de travail. En effet, Vlach & Johnson (2013) ont mis en évidence un rôle important de la mémoire de travail dans l'apprentissage de nouveaux mots. Leurs travaux montrent que des co-occurrences mots-objets sont apprises par des nourrissons de 16 et 20 mois lorsque cette co-occurrence est présentée en succession immédiate, essai après essai, à l'intérieur du même bloc (apprentissage massé). Cependant, seuls les nourrissons les plus âgés étaient capables d'apprendre ces co-occurrences lorsqu'elles étaient séparées par d'autres

essais, distribuées parmi différents blocs (apprentissage distribué). Ces résultats suggèrent donc que la maturation des capacités de mémoire de travail, permettant de maintenir et manipuler des co-occurrences mot-objet en mémoire, contraint les apprentissages langagiers. Pour inclure cette idée dans notre expérience, différents niveaux de complexité des matériaux linguistiques nous serviront à comprendre l'impact des capacités pouvant être stockées en mémoire de travail sur l'apprentissage. Le matériel linguistique se composera ainsi de différentes longueurs de mots : mono- et trisyllabique. Concernant notre troisième hypothèse, nous supposons que les mots courts (monosyllabiques) seront plus faciles à apprendre que les mots plus longs (trisyllabiques) car ils représentent un poids moins important en mémoire de travail.

Quant à la maturité du système visuo-attentionnel, nous supposons qu'un système plus mature (chez des nourrissons plus âgés) permet une meilleure exploration des informations visuelles fournies par le visage parlant (bouche). Cela se reflétera dans un meilleur apprentissage des mots en général mais aussi dans un meilleur apprentissage de mots plus complexes (trisyllabiques).

Partie 2

-

Méthodologie

Chapitre 5. Participants et éthique

Seize nourrissons de langue maternelle française ont participé à notre étude, dont cinq ont été exclus pour notre analyse à cause de difficultés techniques liées au script Matlab ($n = 4$) ou d'agitation ($n = 1$). Il restait par conséquent 11 sujets âgés de 605 à 863 jours ($M = 752$ jours, $SD = 85,64$ jours) dans l'échantillon final pour notre analyse (voir tableau 2).

Tableau 2 : Informations démographiques des participants, y compris l'identifiant du participant (ID), l'âge (jours et mois) et le sexe.

ID	Âge jours (mois)	Sexe (masculin, féminin)
3383	782 (25)	Féminin
3384	782 (25)	Féminin
3385	623 (20)	Féminin
3386	779 (26)	Masculin
3387	863 (28)	Masculin
3390	761 (25)	Féminin
3395	826 (27)	Masculin
3397	629 (21)	Féminin
3398	605 (20)	Masculin
3402	821 (27)	Masculin
3404	* ¹	Féminin
Moyenne (SD)	747 (25) 88 (3)	5 masculins

Initialement, pour tester ces hypothèses, nous avons prévu de réaliser une expérimentation avec deux groupes de nourrissons (18 à 20 mois et 24 à 28 mois). En raison de la pandémie mondiale causée par la diffusion du COVID-19 qui a entraîné un confinement en France, nous avons eu un nombre de participants plus réduit. Notre étude

¹ Information manquante

comporte donc 11 participants âgés de 20 à 28 mois, ce qui réduit notre échantillon à un groupe de sujets.

Les participants résidaient tous à Grenoble ou dans ses environs et ont été élevés dans des familles de statut socio-économique moyen ou moyen-supérieur. Les nourrissons étaient tous exposés la plupart du temps à la langue française dans leur milieu familial (>70%). Deux nourrissons entendaient également le wolof (parlé au Sénégal et en Mauritanie) à la maison ; un troisième participant, le portugais ; et un quatrième participant, la langue des signes.

Tous les nourrissons étaient en bonne santé et sans retard connu dans leur développement, ni troubles visuels ou auditifs. Les participants ont été recrutés à travers le système de recrutement du Babylab du Laboratoire de Psychologie et Neuro-Cognition (LPNC) de Grenoble. Le Babylab est un centre d'étude du nourrisson et du jeune enfant, spécialisé dans l'étude du développement des capacités sensorielles et motrices ; il teste plus de 400 nourrissons par an (<http://www.babylab-grenoble.fr>).

Pour l'ensemble des nourrissons, un accord parental a été exigé pour que l'enfant puisse participer à l'étude. Avant le début de l'étude, une demande d'accord éthique a été effectuée auprès du CERGA (Comité d'Éthique pour les Recherches Grenoble Alpes).

Chapitre 6. Questionnaire parental

Le vocabulaire perceptif et productif de chaque nourrisson a été évalué par un questionnaire parental. L'instrument évaluant la taille du vocabulaire des nourrissons consistait en un questionnaire parental intitulé « inventaires français du développement communicatif (IFDC) ». Ces inventaires sont basés sur des questionnaires parentaux retraçant le développement gestuel et langagier du jeune enfant. Les questionnaires permettent une évaluation rapide des aspects quantitatifs du développement communicatif aux âges de 12, 18 et 24 mois en prenant en compte le nombre de gestes réalisés, de mots produits et/ou compris, et la longueur moyenne des énoncés (Bovet et al., 2005).

Dans notre étude, le questionnaire de 18 mois ainsi que le questionnaire de 24 mois ont été appliqués. Celui de 18 mois présente une liste de 97 mots en compréhension et en production, et une évaluation de la combinaison des mots. Le questionnaire de vingt-quatre mois présente une liste de 100 mots en production, une évaluation de la combinaison des mots, et de la longueur moyenne des énoncés (LME) calculée en nombre

de mots d'après les trois énoncés les plus longs de l'enfant. L'exploitation du questionnaire prend quelques minutes et repose sur une feuille d'évaluation avec des repères en percentiles (du 10e au 90e) ou en pourcentages (Bovet et al., 2005). Les questionnaires ainsi que la feuille d'évaluation sont présentés en annexe.

L'interprétation des performances par la feuille d'évaluation n'a révélé aucun signe d'alerte et confirmé que tous les participants présentaient un développement typique du langage. Les tableaux suivants (voir tableau 3.A et 3.B) montrent les informations remplies par les parents sur la production et la compréhension de mots par leurs enfants. Les questionnaires de 24 mois ont été appliqués aux participants âgés de 25 à 28 mois, et les questionnaires de 18 mois aux participant âgés de 20 à 21 mois, respectivement.

Tableau 3: Évaluation des réponses aux questionnaires. A : Questionnaires pour les 18 mois indiquant le nombre de mots compris et dits, le nombre de mots compris et la combinaison des mots. B : Questionnaires pour les 24 mois indiquant le nombre de mots produits, la combinaison des mots et la longueur des énoncés (LME). En gris : la moyenne et l'écart-type des différents groupes.

A.

IFCD_Questionnaire 18 mois			
ID	Nombre de mots compris et dits (sur 97)	Nombre de mots compris (sur 97)	Combinaison des mots (pas encore/quelqu efois/souvent)
3385	67	28	souvent
3397	22	85	quelquefois
3398	39	89	quelquefois
Moyenne (SD)	43 (19)	67 (28)	-

B.

IFCD_Questionnaire 24 mois			
ID	Nombre de mots produits (sur 100)	Combinaison des mots (pas encore/quelqu efois/souvent)	Longueur moyenne des énoncés (LME)
3383	83	souvent	3
3384	83	souvent	3
3386	38	quelquefois	3
3387	92	souvent	5
3390	91	souvent	4
3395	72	souvent	7
3402	88	souvent	2,67
3404	91	souvent	6
Moyenne (SD)	80 (17)	-	4,21 (1,52)

Chapitre 7. Stimuli**1. Pseudo-mots**

Un corpus de pseudo-mots à consonance française a été créé. Celui-ci est constitué de trois paires de pseudo-mots monosyllabiques de type CVC (consonne-voyelle-consonne), et de trois paires de pseudo-mots trisyllabiques de type CVCVCV :

Tableau 4 : Les pseudo-mots. En haut : les pseudo-mots monosyllabiques de type CVC (consonne-voyelle-consonne) ; en bas : les pseudo-mots trisyllabiques de type CVCVCV :

Mots monosyllabiques		
Paire 1	/bym/	/vak/
Paire 2	/rip/	/fal/
Paire 3	/fuz/	/ker/

Mot trisyllabiques		
Paire 4	/bymuno/	/vakerã/
Paire 5	/riperẽ /	/falõga/
Paire 6	/fuzoly/	/kerita/

Les stimuli monosyllabiques ont été choisis en fonction de la distance phonétique de leurs phonèmes. Cela signifie que les critères articulatoires des voyelles et des consonnes ont été pris en compte afin d'assurer un contraste auditif et visuel entre les membres de chaque paire (voir Jackson et al. (1976) pour des matrices de confusion) qui soit facile à distinguer pour les nourrissons (Patterson & Werker, 2003). Les stimuli ont été créés en s'appuyant d'une part sur les stimuli de type CV utilisés dans l'étude de Havy et Zesiger (2017) et d'autre part sur les stimuli de type CVCV mobilisés dans l'étude de Nazzi (2005), deux études qui ont fait apprendre des nouveaux mots à consonance française à des nourrissons francophones.

Les stimuli ont été contrastés selon différentes caractéristiques : pour les voyelles, les contrastes peuvent impliquer l'oralité ou la nasalité (i.e., voile du palais relevé ou abaissé), l'arrondissement des lèvres, l'ouverture ou la fermeture de la cavité buccale et la position de la langue (i.e., antérieure ou postérieure). Pour les consonnes les contrastes peuvent impliquer le lieu d'articulation, le mode d'articulation (i.e., type d'obstruction au passage de l'aire ; e.g., occlusive : blocage complet de l'écoulement de l'aire au niveau de la bouche, du pharynx ou de la glotte, suivi d'un relâchement soudain de ce blocage) et la vocalisation (i.e., sourde, sonore, liquide).

A partir des mots monosyllabiques - avec leurs phonèmes consonantiques et vocaliques induisant des changements importants au niveau articulatoire - des mot trisyllabiques ont été créés tout en veillant à ce que leurs probabilités phonotactiques soient égales.

2. Enregistrement des stimuli

Les enregistrements ont été effectués dans une chambre sourde au laboratoire GIPSA (Grenoble images parole signal automatique) en collaboration avec Christophe Savariaux. L'ensemble des pseudo-mots a été enregistré à l'aide d'un microphone AKG C1000S et d'une caméra vidéo tri-CCD SONY DXC-990P à 50 images par seconde.

Dans un premier temps, une vidéo a été enregistrée pour la phase de familiarisation et la phase d'apprentissage. Des pseudo-mots ont été intégrés dans des phrases courtes, le tout prononcé par une locutrice de langue maternelle française. La locutrice était placée devant un fond bleu, et seuls sa tête et le haut de son cou étaient visibles (voir figure 5). Elle avait pour consigne d'initier les phrases présentées sur l'écran en face d'elle dans le style particulier de la parole adressée aux enfants (souriante, avec des variations importantes de l'intonation par rapport à la parole adressée à l'adulte ; Fernald & Kuhl, 1987) afin d'attirer l'attention des nourrissons et les motiver.

Les stimuli ont été numérisés et compressés à l'aide du logiciel Dps Reality v 3.1.9 et du codec Cinepak pour obtenir des fichiers vidéo au format .avis (32 bits). Dans un deuxième temps, les pistes audios pour les phases de test ont été enregistrées à l'aide du même microphone AKG C1000S et d'un enregistreur numérique Marantz PMD 670 pour obtenir des fichiers audios au format WAV. Ces enregistrements ont été ensuite traités à l'aide du logiciel Audacity (Audacity, version 2.2.2). Au total, les fichiers audios des phases de test ont une longueur totale de 6 secondes ; le début du mot cible étant toujours présenté à 2,5 secondes.

Le montage des vidéos (visage de la locutrice, sans objets) a été exécuté grâce au logiciel Adobe Premiers Elements (Adobe Premiers Elements, version 7.0). La longueur de la vidéo est de 7 secondes pour la phase de familiarisation et entre 34 et 38 secondes - en fonction de la longueur du mot - pour la phase d'apprentissage. Pendant la présentation, les vidéos ont été positionnées au milieu de l'écran.

Les exemples des stimuli pour les différentes phases sont disponibles sur le lien suivant : <https://cloud.univ-grenoble-alpes.fr/index.php/s/Sbwwx2ckdPP3YeY>.

3. Objets

Quatorze objets ont été créés initialement à l'aide du logiciel Paint3D (Windows, 2019) ; elles se différencient par leur couleur, leur forme, et sont faciles à discriminer. Les objets ont été enregistrés avec un fond noir en format .jpeg avec une taille de 200x200 pixels (5,29 x 5,29 cm).

Vingt-six adultes ont évalué l'attractivité des objets dans une échelle de Likert (quatre niveaux, de « pas du tout attrayant » à « très attrayant ») à l'aide d'un sondage

Googleforms. En se basant sur les résultats de ce sondage, nous avons choisi 12 objets sur la base de leurs scores d'attractivité et avons créé six paires d'objets (voir figure 4).

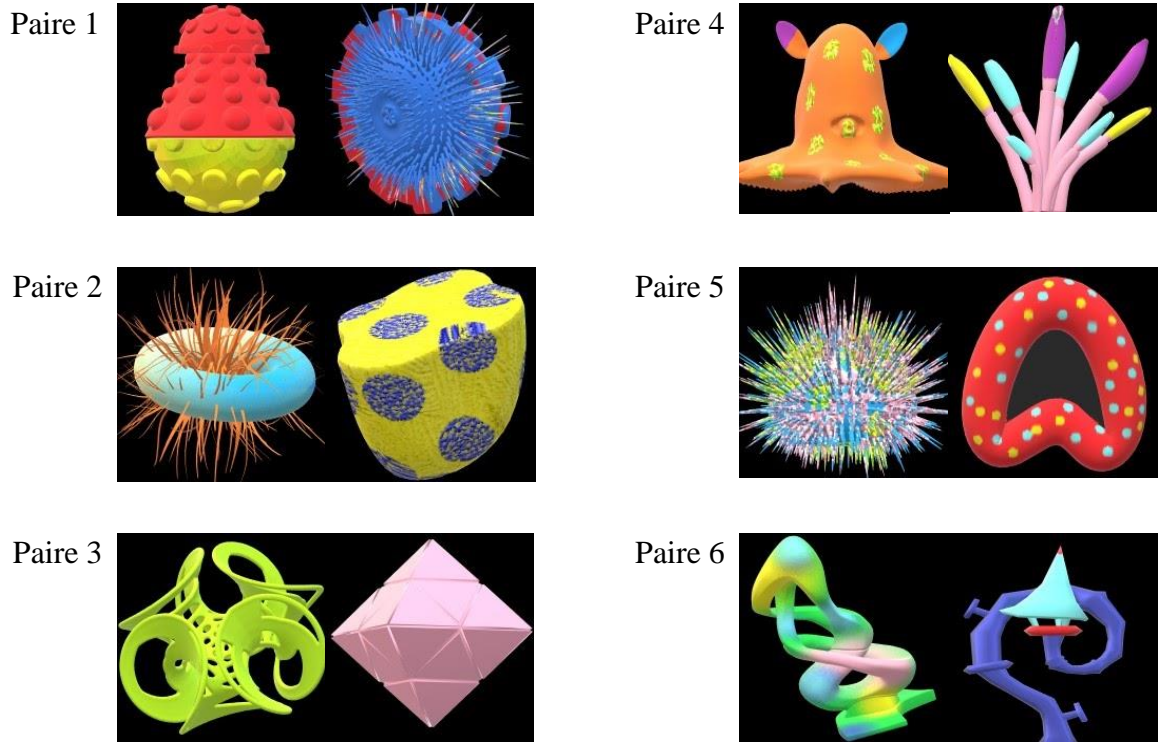


Figure 4 : Les six paires d'objets.

Pendant la phase d'apprentissage, les objets étaient présentés des deux côtés du visage de la locutrice (tous ces stimuli étant centrés dans le sens de la hauteur) et dans son point de vue ; pendant les phases de test, ils apparaissaient au même endroit, mais seuls sur un fond noir (voir figure 5).

4. Contrebalancement

Chaque paire d'objet a été ensuite répartie avec une paire de pseudo-mots en contrebalançant entre participants les éléments suivants : le type de mots (mono-/trisyllabique), l'ordre des mots (premier objet nommé situé à gauche sur l'écran et deuxième à droite, ou inversement) ainsi que l'association du label avec l'objet (e.g., paire d'objet 1 avec paire de mot 1 ; paire d'objet 1 avec paire d'objet 5).



Figure 5: Exemples de la présentation des objets. En haut : Présentation des objets pendant la phase d'apprentissage. En bas : Présentation des objets pendant les phases de test.

Chapitre 8. Dispositif

La passation de l'étude a été réalisée dans une chambre sourde au sein du Babylab du LPNC. La vidéo d'apprentissage a été présentée sur un écran de 19-pouces avec une résolution de 1204 x 768 pixels par pouce et avec un taux de rafraîchissement de 80 Hz. Les stimuli auditifs ont été présentés à travers des haut-parleurs gauches et droits, à un niveau de conversation (60 dB).

Les mouvements des yeux des nourrissons ont été enregistrés au moyen de l'oculomètre EyeLinkR 1000 (SR Research Ltd., Mississauga, Ontario, Canada) avec une fréquence d'échantillonnage de 500 Hz. Le dispositif expérimental est présenté dans la figure 6. La présentation des stimuli et l'acquisition de l'oculomètre ont été contrôlé en utilisant MatLab (R2007b) et l'EyeLink toolbox from the PsychToolbox (Cornelissen et al., 2002).



Figure 6 : Dispositif expérimental : À gauche : L'oculomètre EyeLinkR 1000 avec la présentation de la vidéo sur l'écran (<https://www.abc.net.au/news/2017-08-30/max-davis-sitting-in-a-chair/8852306>). A droite : Table de travail pour l'oculomètre. La caméra enregistre l'endroit que le nourrisson regarde sur l'écran (<https://www.abc.net.au/news/2017-08-30/researchers-use-eye-tracking-to-understand-baby-development/8852330>).

En suivant automatiquement le regard des sujets, la méthode oculométrique (en anglais « Eye-Tracking ») a servi à mettre en évidence les déplacements effectués par les yeux sur l'écran et à faire apparaître les zones où se concentre le regard. Cette méthode a ainsi permis de mesurer de façon fiable (dans une résolution de l'ordre de la milliseconde) et non intrusive les mouvements oculaires des enfants de tout âge.

Chapitre 9. Procédure

Chaque nourrisson a été testé séparément, assis sur les genoux de ses parents, à une distance de 55-60 centimètres de l'écran et de l'oculomètre. Pour faciliter la mise en place, une petite vidéo animée a été présentée afin d'attirer l'attention du nourrisson sur l'écran. Une vignette avec une cible a été placée sur le front du nourrisson (contraste noir sur fond blanc) pour permettre au traqueur oculaire de déterminer la position de la tête du participant. S'en est suivi une calibration en cinq points en présentant un cercle noir au centre de l'écran. Une fois que le nourrisson se focalise sur le cercle, celui-ci se déplace

vers une autre partie de l'écran (en haut, en bas, à droite, à gauche). Après la calibration, l'examineur a lancé le script matlab et la vidéo s'est affichée sur l'écran en suivant la procédure suivante. L'ensemble de la procédure est présenté dans la figure 7.

L'ensemble de la tâche d'apprentissage de mots était composé de plusieurs phases et a duré environ 2,5 minutes. Cette tâche a commencé avec une phase de familiarisation dans laquelle les nourrissons voyaient sur l'écran le visage de la locutrice prononçant « *Coucou ! Tu vas bien ? Viens, je vais te montrer mes jouets.* ». Elle s'est poursuivie par la phase d'apprentissage : Le visage de la locutrice apparaît, prononçant « *Regarde !* » suivi de la présentation des deux objets à l'écran sur un fond noir accompagnés d'un son (« bing »), afin de familiariser les nourrissons aux objets. Ensuite, l'attention du nourrisson a été attiré sur l'objet par le regard de la locutrice sur ce même objet, lequel se trouvait aléatoirement soit à gauche, soit à droite de la locutrice.

Ensuite, la locutrice a tourné la tête vers un objet qu'elle a nommé et ensuite vers l'autre côté de l'écran où se trouvait l'autre objet qu'elle n'a pas nommé (e.g., « *C'est un bume, un bume. Tu le vois le bume ? Il est joli ce bume !* » (objet nommé) et ensuite « *Oh !, regarde ! Il te plaît ? Moi, je l'aime bien.* » (objet non-nommé)). Cet objet correspond à une condition d'« exclusivité mutuelle » : l'idée est que lorsqu'on présente deux objets à un nourrisson, l'un familier, l'autre inconnu, le nourrisson arrive à associer un nouveau mot présenté à l'oral au nouvel objet (pour lequel il n'a pas encore de nom) plutôt qu'à l'objet familier (pour lequel il vient d'apprendre le nom). Les études de Yurovsky et Frank (2017) ont montré qu'à partir de 18 mois les nourrissons sont capables de réussir cette tâche. Le test d'exclusivité mutuelle permet de tester l'apprentissage des associations mot-objet puisque, pour avoir une bonne performance dans les deux conditions (test de l'objet nommé et test de l'objet d'exclusivité mutuelle), le nourrisson doit avoir retenu l'objet nommé lors de la phase d'apprentissage. En ce sens, l'objet d'exclusivité mutuelle nous permettra de vérifier si l'objet nommé a été bien appris et que des bons résultats ne résultent pas d'une simple familiarisation ou préférence pour l'objet.

Une phase d'apprentissage était systématiquement suivie de deux phases de test : la phase de test 1 pour la paire de mot-objet à apprendre (ou nommé) et la phase de test 2 pour l'objet d'exclusivité mutuelle (non-nommé). Lors de chaque essai, les nourrissons ont vu deux objets, un de chaque côté de l'écran, et ont écouté un court clip audio de la

locutrice leur demandant de regarder l'objet cible (e.g., test 1 (l'objet nommé) : « *Il est où le bume ? Regarde le bume !* » ; test 2 (l'objet d'exclusivité mutuelle) : « *Il est où le vak ? Regarde le vak !* »). Chaque phase de test a duré 6 secondes. La présentation du mot cible démarrait toujours à 2,5 secondes.

Lors de la phase d'apprentissage, deux nouvelles associations mot-objet à apprendre ont été présentées à chaque essai (e.g., « bume » + objet 1a ; « vak » + objet 1b, etc.). Chaque nourrisson a complété trois essais et, par conséquent, effectué l'apprentissage de trois paires de mots ou six nouvelles associations mot-objet. Une partie des nourrissons était assignée à l'apprentissage des mots **monosyllabiques** (n = 13) et une partie à des mots **trisyllabiques** (n = 3).

Des capteurs d'attention ont été intégrés au début de chaque phase pour garder l'attention du nourrisson sur l'écran.

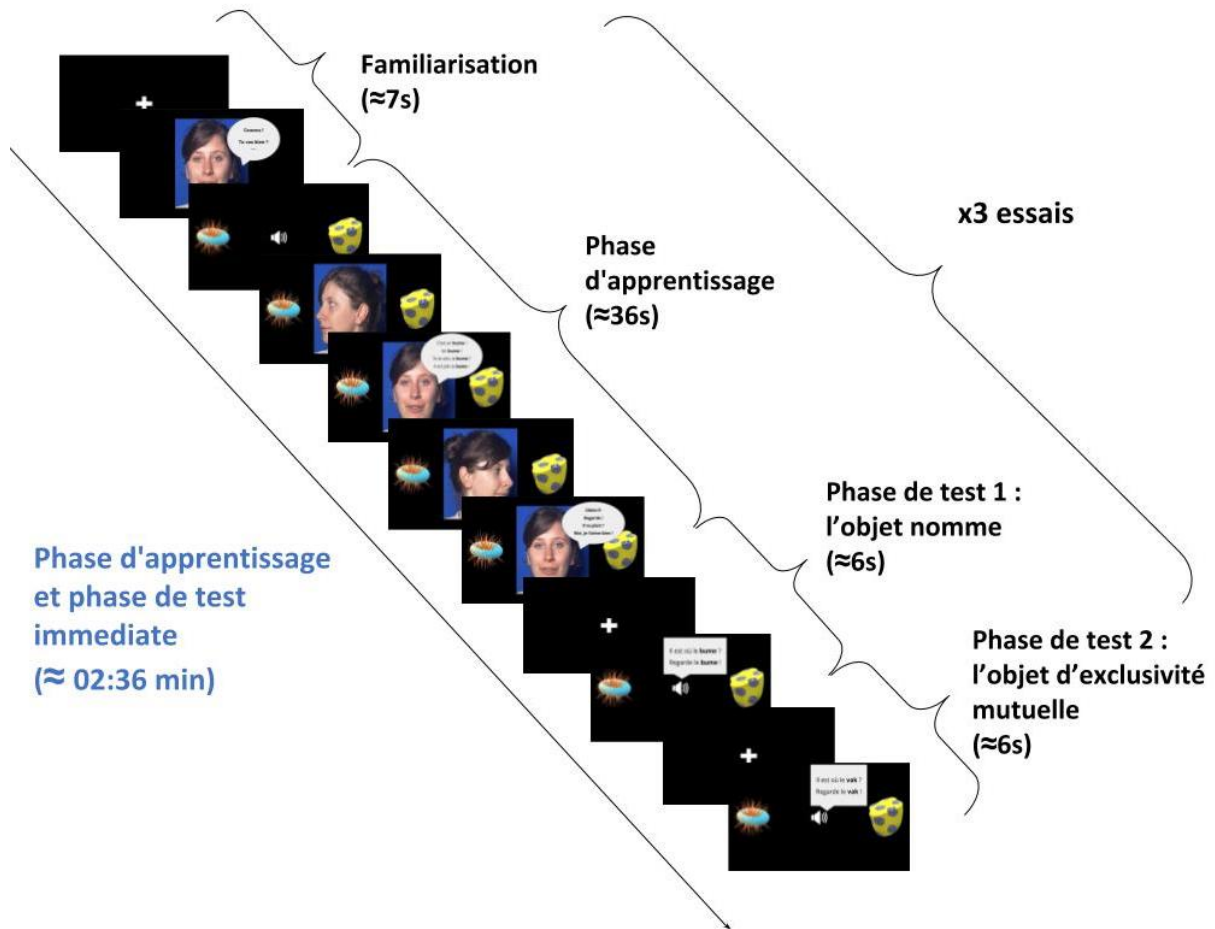


Figure 7 : Représentation schématique de l'ensemble de la procédure utilisée pour la tâche d'apprentissage de mots. Chaque séance débute avec une phase de familiarisation permettant d'habituer les nourrissons à la locutrice. Vient ensuite la phase d'apprentissage du nom d'un des objets grâce au regard de la locutrice qui se dirige vers cet objet pendant qu'il est nommé, puis vers l'autre objet qui n'est pas nommé (l'objet d'exclusivité mutuelle). Dans les phases de test, les deux objets sont présentés et un des objets est nommé : dans la phase de test 1 c'est l'objet nommé ; dans la phase de test 2 c'est l'objet d'exclusivité mutuelle, dont les nourrissons entendent le nom pour la première fois. La phase d'apprentissage et les phases de test immédiates sont répétées trois fois pour chaque nourrisson (soit trois associations mot-objet apprises par enfant). Le regard que le nourrisson porte sur l'objet correct après qu'il soit nommé sera considéré comme un indice de cet apprentissage.

Chapitre 10. Prédictions expérimentales

Nous supposons que les performances des nourrissons en apprentissage de nouveaux mots dépendent d'une grande partie du développement des mécanismes cognitifs, notamment des mécanismes de l'attention sélective et des mécanismes de la mémoire de travail.

Pendant la phase d'apprentissage, les nourrissons devraient suivre du regard les yeux de la locutrice vers l'objet et devraient à l'inverse fixer la bouche lorsque la locutrice

parle et produit le mot-cible à apprendre. Nous nous attendions à ce que les nourrissons regardent plus la bouche de l'interlocutrice que ses yeux pendant l'apprentissage de nouveaux mots.

La capacité à orienter son attention à la bouche au moment de la prononciation du mot cible dans la phase d'apprentissage devrait être positivement corrélée avec les performances dans les phases de test.

De même, ces performances d'apprentissage de nouveaux mots devraient être positivement corrélées à la taille du vocabulaire productif (donnée par les résultats du questionnaire parental).

La reconnaissance de mots appris dans la phase de test devrait être meilleure avec l'avancement en âge : Les nourrissons plus âgés (24 à 28 mois) devraient apprendre plus d'associations mot-objet (surtout les mots trisyllabiques), car leurs mécanismes attentionnels et leur mémoire de travail sont plus développés que ceux des nourrissons plus jeunes (18 à 20 mois).

Les mots courts (monosyllabiques) devraient être globalement plus faciles à apprendre que les mots plus longs (trisyllabiques) car ils représentent un poids moins important en mémoire de travail. Cet effet devrait se refléter dans des meilleures performances pour les mots monosyllabiques dans les phases de test.

Partie 3

-

Analyses et résultats

Chapitre 11. Analyses

1. *Méthode analytique*

Les données qui ont été enregistrées par l'oculomètre nous ont permis d'évaluer les stratégies exploratoires attentionnelles ainsi que le lien avec les performances en apprentissage de nouveaux mots des nourrissons.

Initialement, nous avons prévu de recruter deux groupes de nourrissons se différenciant par leur âge (un groupe de 18 à 20 mois et un groupe de 24 à 28 mois) afin d'évaluer l'effet de la maturation des mécanismes cognitifs sur cet apprentissage. En raison de la pandémie COVID-19, nous n'avons pas pu former ces deux groupes. La variable « âge » n'a donc pas pu être testée.

Trois des participants qui ont été testés ($M = 619$ jours (20,63 mois), $SD = 12,49$, tranche d'âge = 605-629 jours) étaient initialement assignés au groupe de nourrissons plus jeunes. En raison du faible nombre de participants ($N = 11$), nous avons rassemblé les données de tous les participants pour notre analyse. Au préalable, toutes les analyses ont été testées séparément pour les deux groupes. Ces résultats n'ont montré aucune différence significative entre les deux groupes.

Pour l'analyse de l'apprentissage de nouveaux mots, nous avons utilisé des modèles linéaires à effets mixtes (Magezi, 2015). Un premier avantage de ces modèles est la prise en compte de la variabilité liée à chaque individu. Dans notre étude, les données montaient jusqu'à trois essais pour chaque nourrisson. Un second avantage de ces modèles est qu'ils gèrent des données manquantes. En effet, quelques essais ont été exclus de nos analyses (voir « Nettoyage de données »). Nous avons utilisé le logiciel R (R Core Team, 2013) avec le paquet lme4 (Bates et al., 2007). La significativité a été calculée en utilisant le paquet lmerTest (Kuznetsova et al., 2017). Ce paquet applique la méthode Satterwaite pour estimer les degrés de liberté et qui génère les valeurs p pour le modèle à effets mixtes. À côté de ces modèles mixtes, des modèles linéaires simples ont été utilisés pour analyser la phase de familiarisation comme il n'y avait qu'un seul point de données par participant.

2. *Prétraitement des données et définition des aires d'intérêt (AOIs)*

Notre analyse a consisté en premier lieu en une analyse en Aires d'Intérêts (AOI ; en anglais : « Area of Interest ») à des moments clés. Les AOIs font référence à des sous-régions d'un stimulus sélectionnées préalablement. Dans notre étude, nous avons défini quatre AOIs. D'abord, deux AOIs rectangulaires dans la partie du visage le divisant horizontalement en deux parties à hauteur du milieu du nez. La partie du haut représentait la région des yeux et la partie basse la région de la bouche. Étant donné que la position de la locutrice restait constante pour tout le clip vidéo, l'endroit de ces AOIs a été maintenu pour tout le clip vidéo.

Le même paramétrage a été utilisé pour la phase d'apprentissage en ajoutant deux AOIs quadrangulaires autour des objets présentés sur les deux côtés. Pendant les phases de test, seules les AOIs des objets ont été présentés.

A l'aide du programme MATLAB, nous avons ensuite vérifié pour chaque point de données collectées s'il se présentait à l'intérieur ou à l'extérieur des AOIs définies.

Nous avons calculé ensuite les proportions du temps total de regard (PTTR ; en anglais « Proportion of total looking time (PTLT) ») que les nourrissons ont passé à regarder chaque AOI pendant chaque essai et chaque phase. Pour cela, le temps total de regard à l'AOI a été divisé par la somme du temps total de regard de toutes les AOIs définies sur l'écran selon les différentes phases.

Pour la phase de familiarisation, le temps de regard (TR) à l'AOI en question a été divisé par les temps de regard aux deux AOIs du visage (e.g., $PTTR_{bouche} = TR_{bouche} / (TR_{bouche} + TR_{yeux})$). Pour la phase d'apprentissage, le temps de regard à l'AOI en question a été divisé par le temps de regard à toutes les quatre AOIs (e.g., $PTTR_{obj.nommé} = TR_{obj.nommé} / (TR_{bouche} + TR_{yeux} + TR_{obj.nommé} + TR_{obj.exclu.mut})$). Pour les phases de test, le temps de regard à l'objet cible (qui a été testé) a été divisé par les temps de regard aux deux objets (e.g., $PTTR_{obj.exclu.mut} = TR_{obj.exclu.mut} / (TR_{obj.nommé} + TR_{obj.exclu.mut})$). Par exemple, la figure 8 montre les AOIs définies de la phase d'apprentissage.

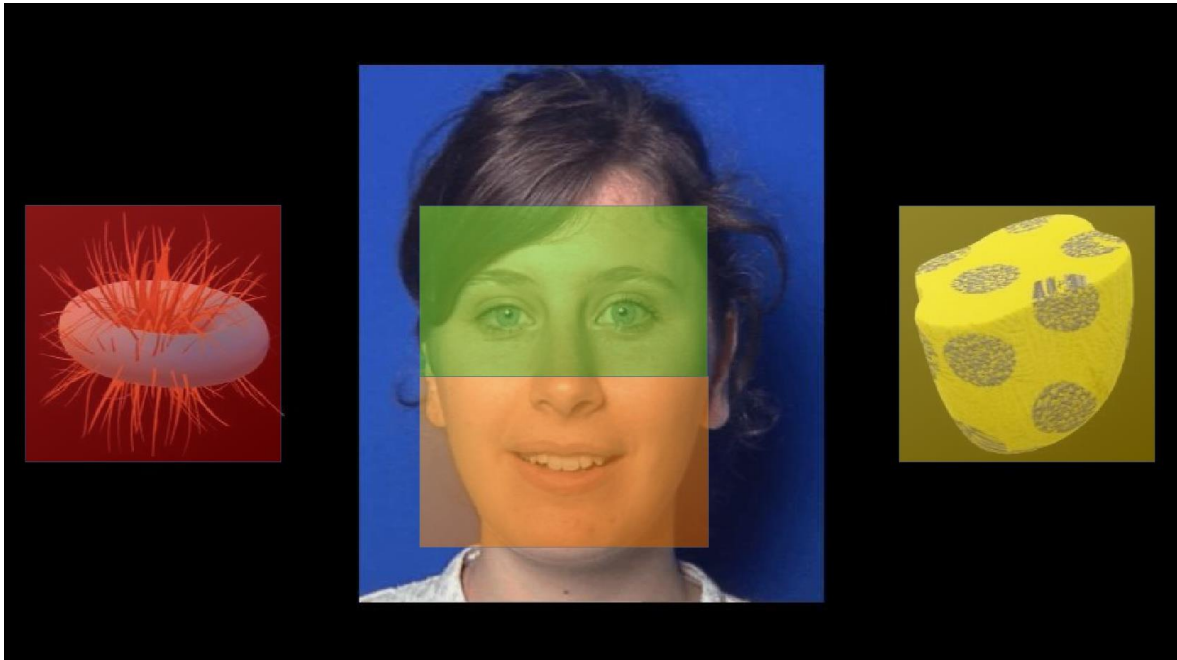


Figure 8 : Illustration des AOIs définies pour la phase d'apprentissage. En vert : l'AOI « yeux », en orange : l'AOI « bouche », les AOIs des objets sont présentées en rouge et jaune. La phase de familiarisation a inclus les AOIs « bouche » et « yeux » seuls. Les phases de test ont inclus seuls les AOIs des deux objets.

3. *Nettoyage de données*

Le nettoyage des données a consisté en une série de trois filtres successivement appliqués à l'ensemble initial de données. Pour les 11 participants, l'ensemble de données a inclus initialement 11 mesures pour la phase de familiarisation et 33 essais pour la phase d'apprentissage et les phases de test (trois essais par sujet).

Dans un premier temps, tous les essais où les nourrissons ont regardé plus de 75 % hors des AOIs prédéfinies ont été exclus (un essai de la phase de familiarisation et un essai pour la phase d'apprentissage ainsi que les essais des tests associés).

Un essai de la phase d'apprentissage ainsi que les essais associés pour les phases de test ont été exclus car le participant n'a pas exploré les deux objets pendant la phase d'apprentissage.

Comme dans Havy et Zesiger (2017), tous les essais des phases de test pendant lesquels les sujets n'ont pas exploré les deux objets avant l'objet a été nommé ont été exclus (trois essais pour le test de l'objet nommé et six essais pour le test de d'exclusivité mutuelle) afin de contrôler un éventuel biais d'objets.

Au total, les données d'un sujet pour la phase de familiarisation et 10 sur les 33 essais pour la phase d'apprentissage et les phases de test ont été retirés de l'ensemble de données initial.

Chapitre 12. Résultats

1. Phase de familiarisation

Pendant la phase de familiarisation, seul le visage de la locutrice a été présenté sur l'écran. La phase permet au nourrisson de prendre connaissance et de s'habituer à la locutrice. L'analyse de cette phase nous a permis d'explorer les patterns attentionnels spontanés des nourrissons quand ils sont confrontés à des visages parlants, en dehors d'une situation d'apprentissage.

Les résultats sont présentés dans la figure 9. Étant donné que les résultats pour cette phase comportaient un seul essai, chaque participant a été représenté par un seul point de données. Nous avons donc utilisé une régression linéaire simple pour analyser la possibilité d'une différence entre la proportion du temps total de regard à la bouche (PTTRbouche) et la proportion du temps total de regard aux yeux (PTTRyeux). Pour cette phase de familiarisation la variable AOI a eu deux valeurs : les yeux et la bouche. L'intercept (AOIbouche) a été significativement différent de zéro. Les résultats pour l'aire d'intérêt des yeux (AOIyeux) n'a pas atteint une significativité statistique ($\beta = -0,08$, $SE = 0,1$, $t = -0,82$, $p = 0,422$). Ce résultat indique que les nourrissons n'ont pas une préférence pour l'une ou l'autre de deux parties du visage (bouche et yeux).

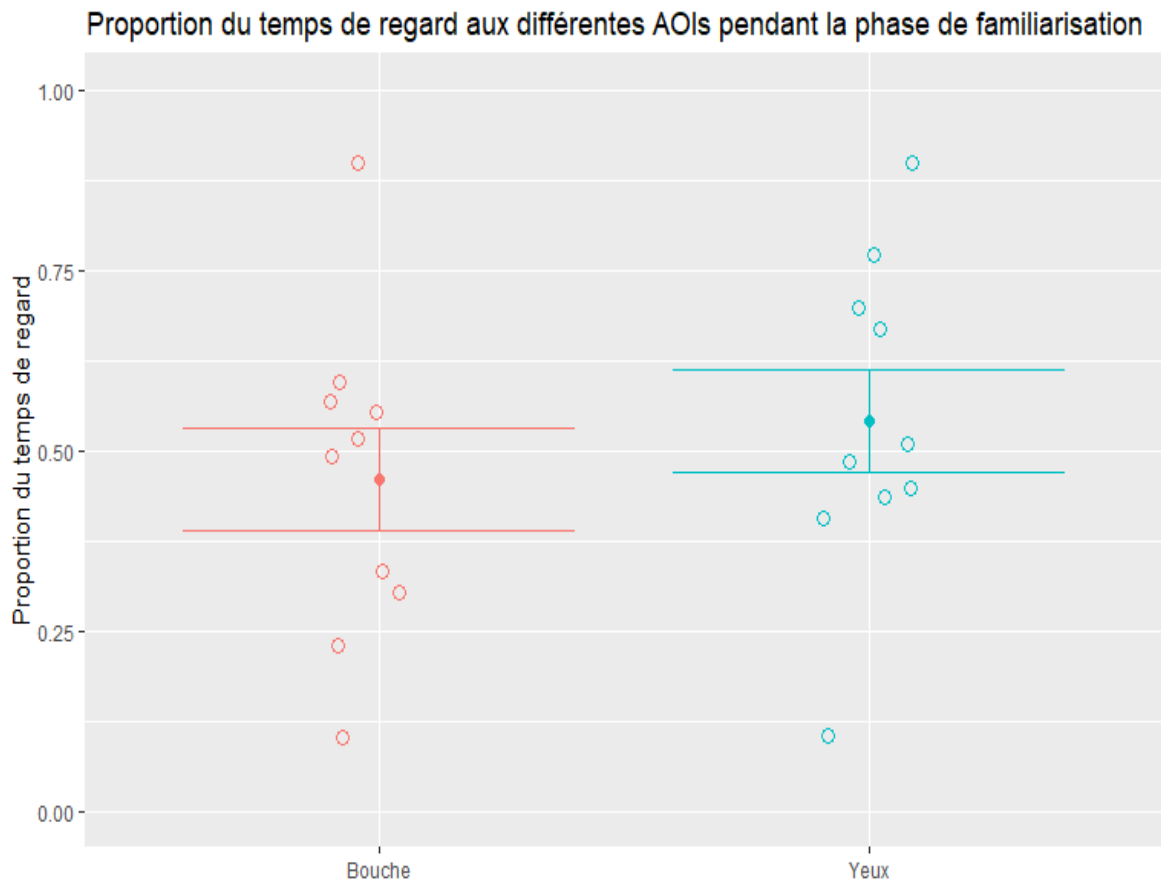


Figure 9 : Proportion du temps de regard pour l'AOI de la bouche (rouge) et pour l'AOI des yeux (bleu) pendant la phase de familiarisation. Chaque point représente un essai et un participant. Les barres d'erreurs représentent l'erreur type de la moyenne.

2. Phase d'apprentissage

Dans la phase d'apprentissage, chaque participant a suivi trois essais. Pour analyser les données de cette phase, tout en prenant en compte les variabilités interindividuelles, nous avons utilisé un modèle linéaire à effets mixtes avec la proportion du temps total de regard (PTTR) comme variable dépendante et l'AOI comme effet fixe. La variable AOI a eu quatre niveaux : les yeux, la bouche, l'objet appris (l'objet nommé) et l'objet d'exclusivité mutuelle. Nous avons également pris en compte les effets aléatoires par participant. Le modèle a été spécifié par la formule suivante : $PTTR \sim AOI + (1|participant)$.

Les résultats sont présentés dans la figure 10. L'intercept a été positif et significatif ($\beta = 0.40$, $ET = 0.02$, $t(120) = 16,93$, $p < 0,0001$) indiquant que l'intercept (la bouche) est significativement différent de zéro. Les résultats pour l'AOI de l'objet nommé ($\beta = -0,26$,

ET = 0.03, $t(120) = -7,92$, $p < 0,0001$) et pour l'AOI de l'objet d'exclusivité mutuelle ($\beta = -0,26$, ET = 0.03, $t(120) = -7,87$, $p < 0,0001$) sont négatifs et significatifs ce qui montre que les nourrissons passent plus de temps regarder la bouche (l'intercept) que les objets. De plus, l'AOI des yeux atteint également la significativité ($\beta = -0,08$, ET = 0.03, $t(120) = -2,37$, $p = 0,02$) indiquant que les nourrissons passent significativement plus de temps à regarder la région autour de la bouche de l'interlocutrice que la région autour des yeux pendant la phase d'apprentissage (la dénomination de nouveaux mots).

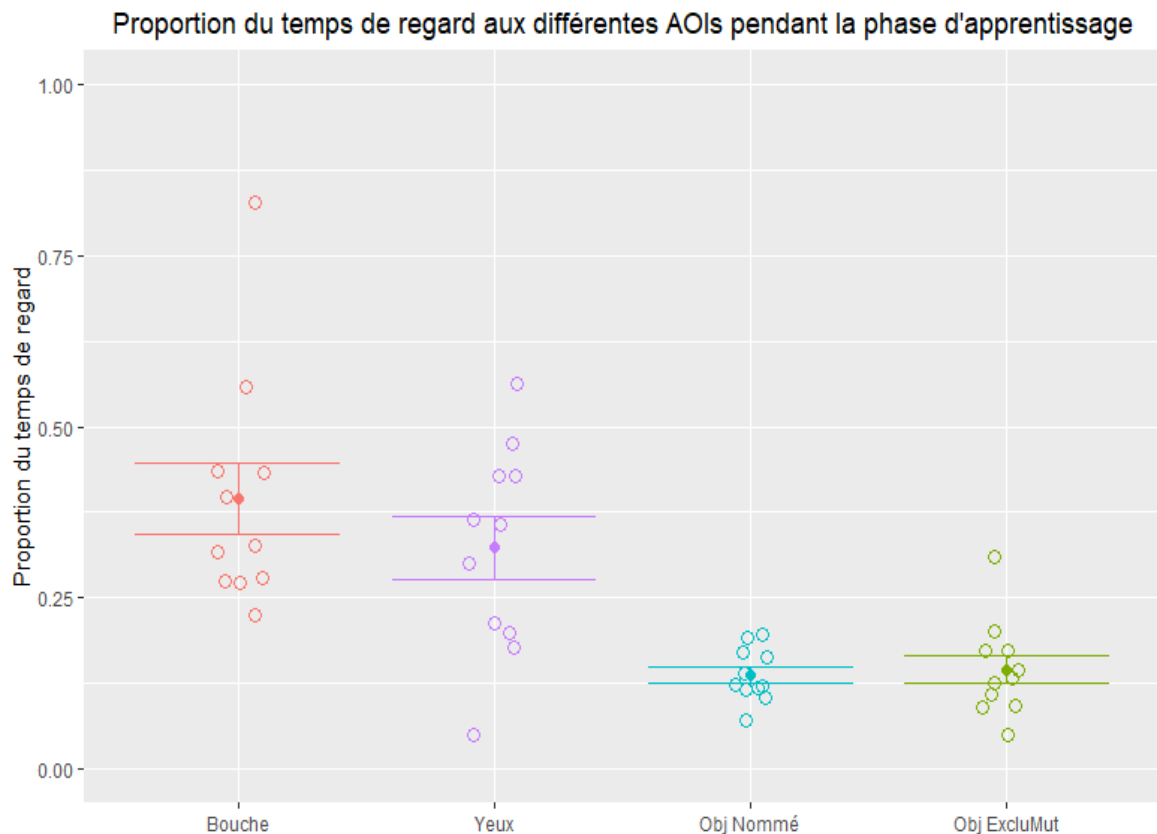


Figure 10 : Proportion du temps de regard pour les AOIs de la bouche (rouge), les yeux (violet), l'objet nommé (bleu) et l'objet d'exclusivité mutuelle (vert) pendant la phase de familiarisation. Chaque point représente les proportions du temps total de regard moyenné par participant à travers les essais. Les barres d'erreurs représentent l'erreur type de la moyenne.

Afin de contrôler un biais potentiel pour un des objets, nous avons effectué des analyses a posteriori : les comparaisons multiples des moyennes avec des contrastes Tukey en utilisant le paquet « multcomp » (Hothorn et al., 2008).

Il n'y a pas eu de différence significative entre les différentes PTTR (proportion du temps total de regard) pour l'objet d'exclusivité mutuelle ($M_{\text{ObjNommé-ObjExcluMut}} = -0.002$, SE = 0.03, $z = -0.06$, $p = 1,00$) indiquant que les participants ont passé autant de temps à regarder l'objet d'exclusivité mutuelle que l'objet nommé (ou appris).

3. *Phases de test*

Pour les deux tests, nous avons considéré le gain en proportion du temps total de regard après la dénomination comme indicateur d'apprentissage de nouveau mot. Dans chaque phase de test, seulement un des objets a été nommé. Le temps de regard à l'objet correct après qu'il a été nommé nous a servi d'indicateur permettant de vérifier l'apprentissage de ce mot.

Premièrement, la proportion du regard pour chaque objet a été calculée en divisant sa valeur respective du temps de regard par la somme du temps de regard pour les deux objets (e.g., $PTTR_Obj.ExcluMut = TR_Obj.ExcluMut / (TR_Obj.ExcluMut + TR_Obj.Nommé)$). Cette proportion a été calculée pour deux périodes : **avant** (pre-nomination ; PropNpre, durée = 2,5 secondes) et **après** (post-nomination ; PropNpost, durée = 3,5 secondes) la dénomination de l'objet cible qui a été toujours à 2,5 secondes), e.g., $PropNpre = PTTRnamed.obj_pre / (PTTRnamed.obj_pre + PTTRmutexclu.obj_pre)$.

Nous avons calculé ensuite la différence de la proportion du temps total du regard à l'objet cible en soustrayant la proportion du temps de regard de la période après la dénomination de l'objet cible (PropNpost) de celle de la période avant la dénomination (PropNpre) : $Gain = PropNpost - PropNpre$. Cela nous a permis de comparer le gain en proportion du temps total de regard avec 0 (e.g., zéro = proportion égale pour l'objet cible pour les deux périodes avant et après la dénomination de l'objet cible ; > 0 = proportion supérieure pour l'objet cible pour la période après la dénomination indiquant un apprentissage du mot. Le gain en proportion du temps total de regard à l'objet cible (différence de PropNpost-PropNpre) a été la variable dépendante dans les deux tests.

Test de l'objet nommé (Test 1)

Une analyse des effets linéaires mixtes a été effectuée afin de mieux comprendre la relation entre la différence de PropNpost-PropNpre de la préférence visuelle pour l'objet cible et le type de mot. Le type de mot est représenté par l'effet fixe à deux niveaux : la condition de mots monosyllabiques et la condition de mots trisyllabiques. La condition de mots monosyllabiques a été le niveau de référence. Les effets aléatoires par participant ont été également pris en compte. Le modèle a été défini par $Gain_{post-pre} \sim TypeMot + (1|participant)$.

Les résultats sont présentés dans la figure 11. L'intercept n'a pas été significativement différent de zéro ($\beta = 0.05$, $SE = 0.07$, $t(11,14) = 0.79$, $p = 0.505$) indiquant qu'il n'y avait pas de gain significatif de la proportion du temps de regard à l'objet cible après la dénomination dans la condition monosyllabique. L'effet du type de mot n'as pas été significativement différent de zéro non plus ($\beta = -0.12$, $SE = 0.12$, $t(13,38) = -1.12$, $p = 0.283$) indiquant qu'il n'y avait pas de différence en terme de gain de proportion du regard (PropNpost-PropNpre) entre la condition de mots monosyllabiques et la condition de mots trisyllabiques.

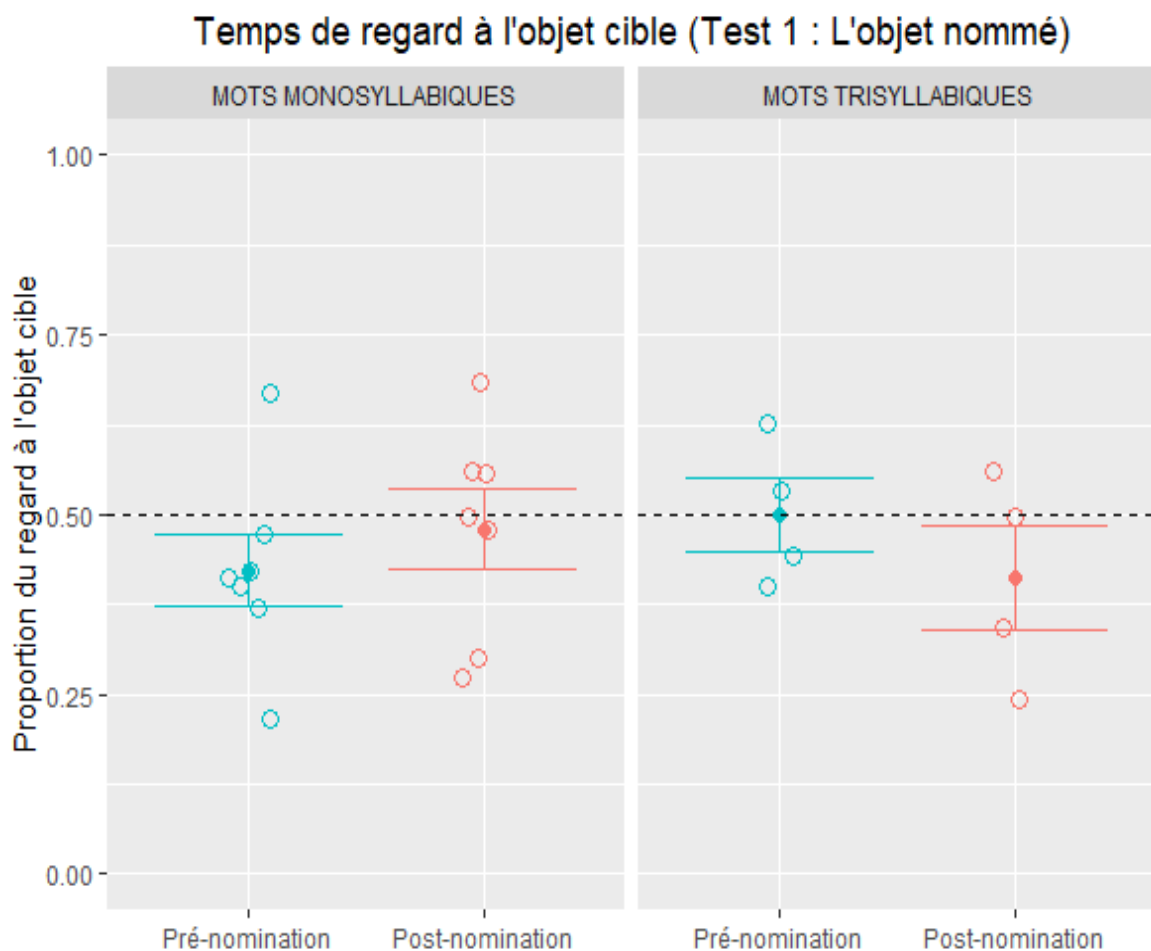


Figure 11 : Proportion du temps de regard à l'objet cible pour les périodes avant (Pré-nomination) et après (Post-nomination) la dénomination dans la phase de test de l'objet appris (ou nommé). À gauche : résultats pour la condition monosyllabique. À droite : résultats pour la condition trisyllabique. Chaque point représente les proportions du temps de regard moyennées par participant à travers les essais. Le niveau qui définit le hasard est marqué à 0,5. Les barres d'erreurs représentent l'erreur type de la moyenne.

Ces résultats montrent que le fait de nommer l'objet n'implique pas un changement significatif de la proportion du regard à l'objet cible. Ceci indique que la dénomination

n'influence pas la proportion du regard à l'objet cible que cela soit pour les mots monosyllabiques ou pour les mots trisyllabiques.

L'hypothèse que les nourrissons regardent significativement plus l'objet cible après sa dénomination n'a pas pu être confirmée. Ni l'effet du type de mot.

Test de l'objet d'exclusivité mutuelle (Test 2)

L'analyse pour le test de l'objet d'exclusivité mutuelle a été effectuée de la même manière que pour le test de l'objet nommé : $\text{Gain}_{\text{post-pre}} \sim \text{TypeMot} + (1|\text{participant})$.

Les résultats sont présentés dans la figure 12.

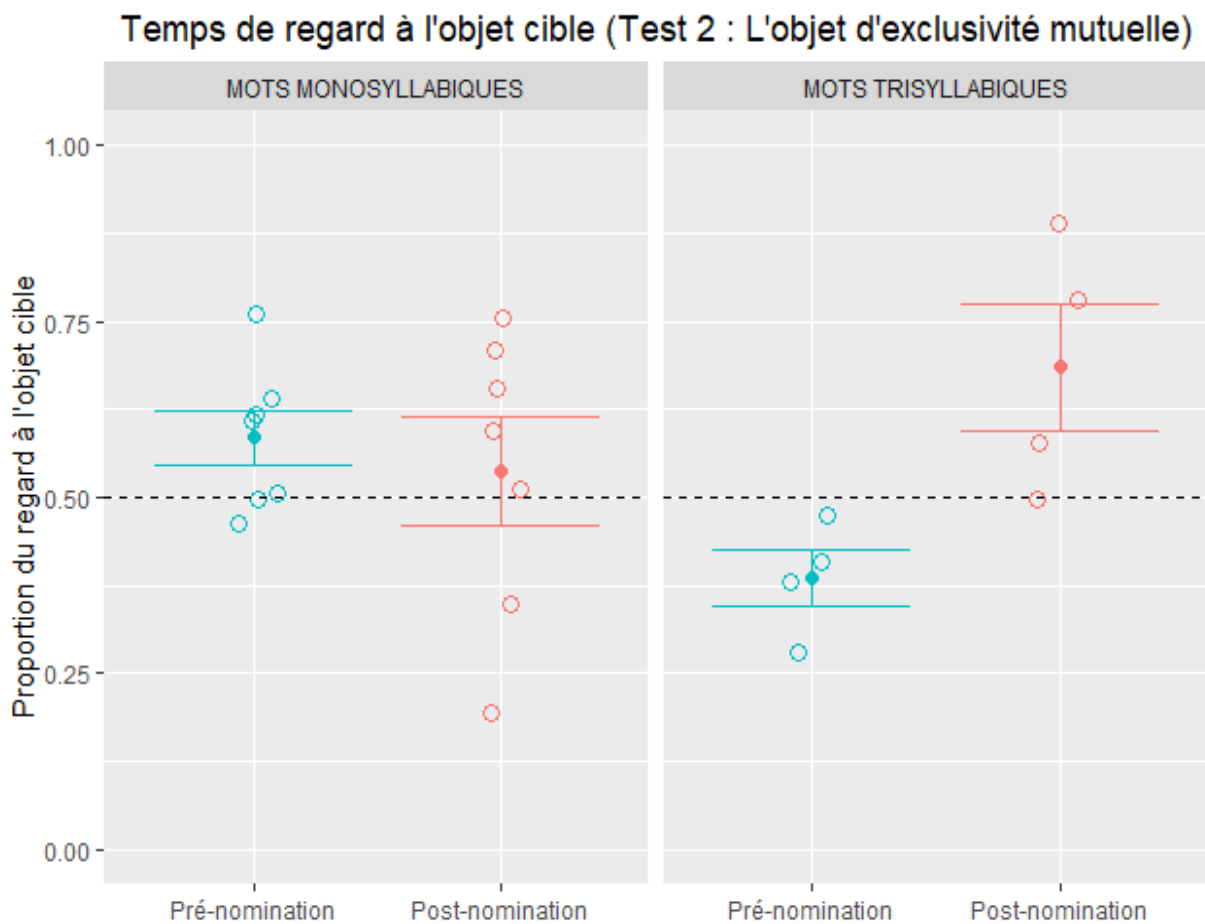


Figure 12 : Proportion du temps de regard à l'objet cible pour les périodes avant (Pré-nomination) et après (Post-nomination) la dénomination dans la phase de test de l'objet d'exclusivité mutuelle. À gauche : résultats pour la condition monosyllabique. À droite : résultats pour la condition trisyllabique. Chaque point représente les proportions du temps de regard moyenné par participant à travers les essais. Le niveau qui définit le hasard est marqué à 0,5. Les barres d'erreurs représentent l'erreur type de la moyenne.

L'intercept n'a pas été significativement différent à zéro ($\beta = 0,01$, $SE = 0.06$, $t(8,4) = 0,19$, $p = 0,855$) indiquant qu'il n'y a pas de gain significatif de la proportion du temps de regard pour l'objet cible après la dénomination dans la condition monosyllabique. Le résultat du groupe monosyllabique concorde ainsi avec le résultat du test de l'objet nommé.

L'effet de type de mot a été significatif et positif ($\beta = 0,29$, $SE = 0.10$, $t(7,8) = 2,82$, $p = 0,023$) indiquant que les nourrissons du groupe trisyllabique ont présenté un gain significatif en terme de proportion du regard pour l'objet cible après sa dénomination.

4. Préférence pour la bouche et apprentissage

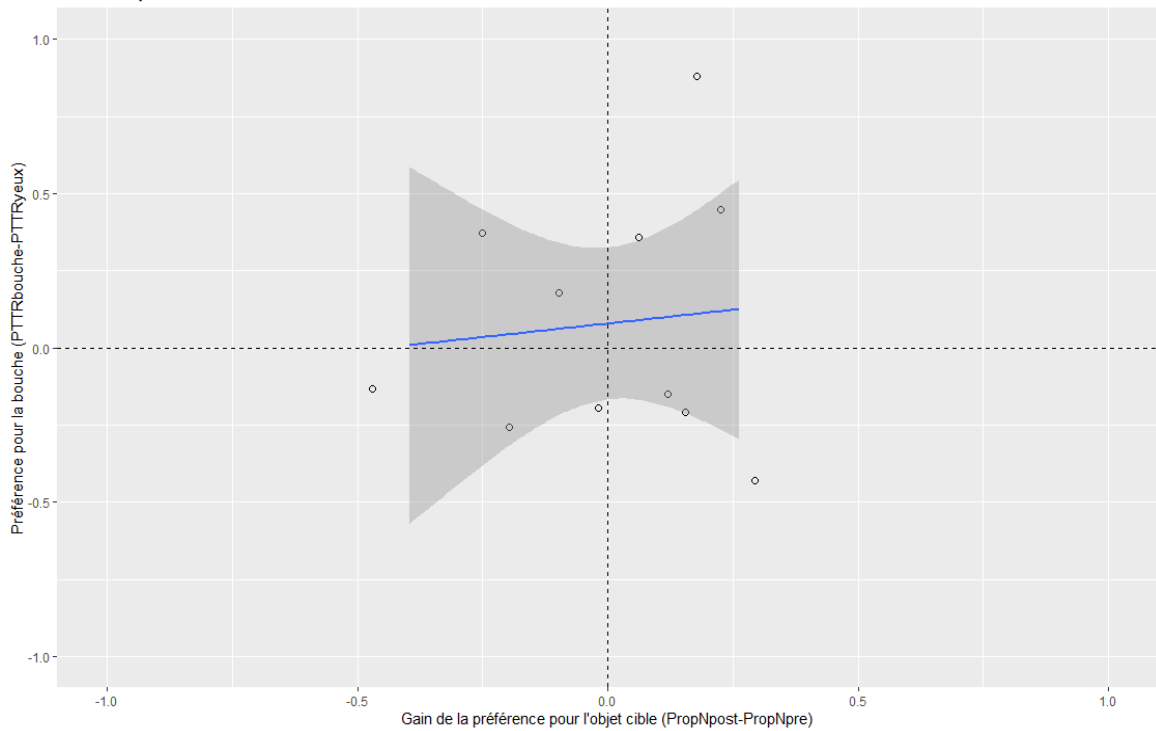
Pendant la phase d'apprentissage, la proportion du temps de regard à la bouche a été significativement différente de celle du temps de regard aux yeux. Pour vérifier notre hypothèse, à savoir qu'une préférence pour la bouche pendant la phase d'apprentissage prédit des meilleurs performances (gain de proportion de regard pour l'objet cible après dénomination) dans les deux tests, nous avons effectué une analyse utilisant des modèles linéaires à effets mixtes.

Le gain de la proportion du regard à l'objet cible après la dénomination ($\text{PropNpost} - \text{PropNpre}$) a constitué la variable dépendante. La préférence pour la bouche pendant la phase d'apprentissage a représenté l'effet fixe. La préférence pour la bouche pendant la phase d'apprentissage a été calculée en soustrayant la proportion du temps de regard aux yeux de la proportion du temps de regard à la bouche (e.g., $\text{Prefbouche} = \text{TRbouche} / (\text{TRbouche} + \text{TRyeux}) - \text{TRyeux} / (\text{TRbouche} + \text{TRyeux})$). Un résultat positif correspond à une préférence relative pour la bouche. Les effets aléatoires par participant ont été prises en compte. Le modèle a été défini pour les deux tests de la manière suivante : $\text{Gain}_{\text{post-pre}} \sim \text{Prefbouche} + (1|\text{participant})$. Les résultats sont présentés dans la figure 13.

Pour le test de l'objet nommé, l'intercept ($\beta = 0,003$, $SE = 0.06$, $t(11,67) = 0,05$, $p = 0,959$) n'a pas été significativement différent à zéro indiquant qu'il n'y avait pas de gain significative de la proportion du temps de regard vers l'objet cible après sa dénomination.

La préférence pour la bouche n'a pas été significativement différent à zéro non plus ($\beta = 0,003$, $SE = 0.15$, $t(16,46) = 0,02$, $p = 0,985$) indiquant que l'effet de la préférence pour la bouche ne prédit pas la performance dans le test ($\text{PropNpost} - \text{PropNpre}$).

Corrélation entre la performance (PropNpost-PropNpre) dans le test et la préférence pour la bouche pendant la phase d'apprentissage
Test 1 : L'objet nommé



Corrélation entre la performance (PropNpost-PropNpre) dans le test et la préférence pour la bouche pendant la phase d'apprentissage
Test 2 : L'objet d'exclusivité mutuelle

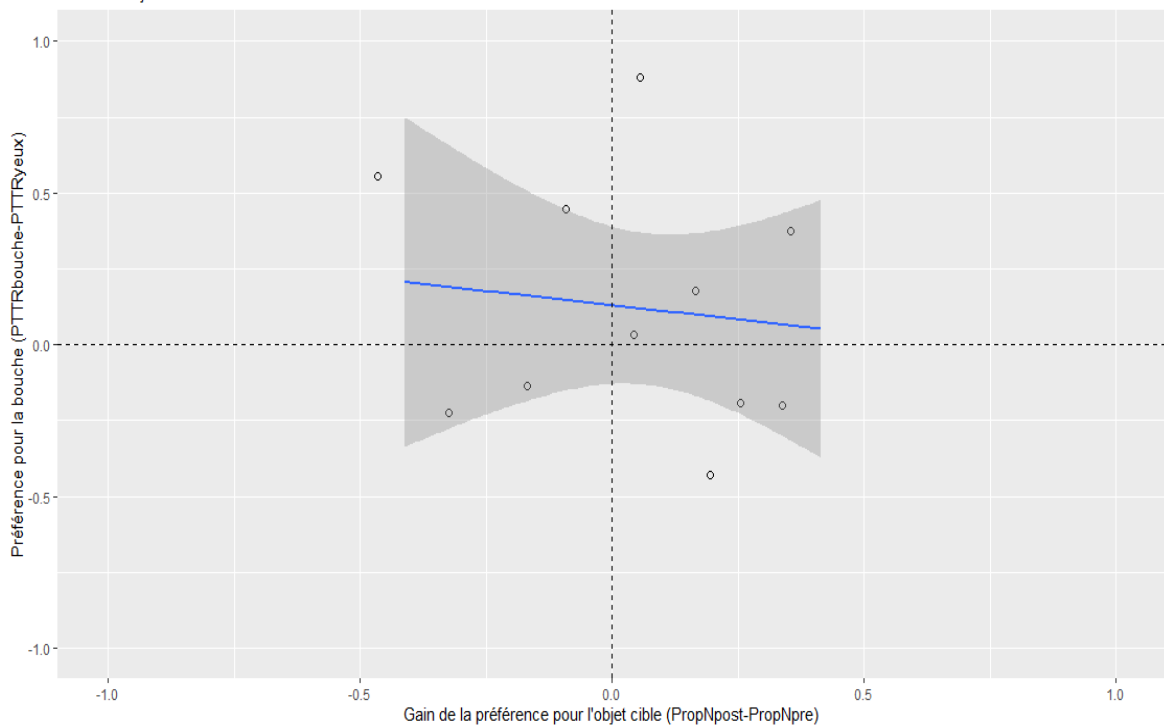


Figure 13 : Préférence pour l'objet cible dans la période de post-nomination (PropNpost-PropNpre) dans les tests (abscisse) prédit par la préférence pour la bouche (PTTRbouche-PTTRyeux) pendant la phase d'apprentissage (ordonnée). En haut : test de l'objet nommé. En bas : test de l'objet d'exclusivité mutuelle. Le hasard présenté à 0 par la ligne pointillée. Un point représente les résultats moyennés par participant.

Pour le test de l'objet d'exclusivité mutuelle, l'intercept n'a pas été significativement différent de zéro ($\beta = 0,107$, $SE = 0,07$, $t(8,03) = 1,53$, $p = 0,165$) indiquant qu'il n'y avait pas de gain significative de la proportion du temps de regard vers l'objet cible après sa dénomination. La préférence pour la bouche n'a pas été significativement différent de zéro non plus ($\beta = -0,025$, $SE = 0,16$, $t(11,46) = -0,15$, $p = 0,880$) indiquant que la préférence pour la bouche ne prédit pas la performance dans le test (PropNpost-PropNpre).

Ces résultats indiquent que de meilleures performances en apprentissage de nouveaux mots ne peuvent pas être prédits par une préférence pour la bouche de l'interlocutrice pendant l'apprentissage de nouveaux mots. L'hypothèse que regarder la bouche pendant l'apprentissage de nouveaux mots améliore l'apprentissage de ces mots n'a pas pu être confirmée.

5. Meilleures performances et vocabulaire productif

Pour vérifier notre hypothèse qui considère que le vocabulaire productif est lié à des meilleures performances (gain de proportion du regard pour l'objet cible après dénomination) dans les deux tests, nous avons effectué une analyse utilisant des modèles linéaires à effets mixtes. Le gain de la proportion du regard à l'objet cible après la dénomination (PropNpost-PropNpre) a constitué la variable dépendante. Le vocabulaire productif (résultats de questionnaire parental IFDC) a représenté l'effet fixe. Les effets aléatoires par participant ont été prises en compte. Le modèle a été défini pour les deux tests de la manière suivante : $\text{Gain}_{\text{post-pre}} \sim \text{NombreMots} + (1|\text{participant})$. Les résultats sont présentés dans la figure 14.

Il est à noter que deux questionnaires différents ont été utilisés : le questionnaire de 18 mois pour les trois participants plus jeunes (18 à 20 mois) et le questionnaire de 24 mois pour les des participants plus âgés (24 à 28 mois). Ces questionnaires étaient différents car ils utilisaient des mots différents et présentaient un nombre de mots total différent (97 pour le questionnaire de 18 mois et 100 mots pour le questionnaire de 24 mois). Par conséquent, le pourcentage de mots produits (nombre de mots sur 97 pour et nombre de mots sur 100, respectivement) a été prise en compte pour notre analyse.

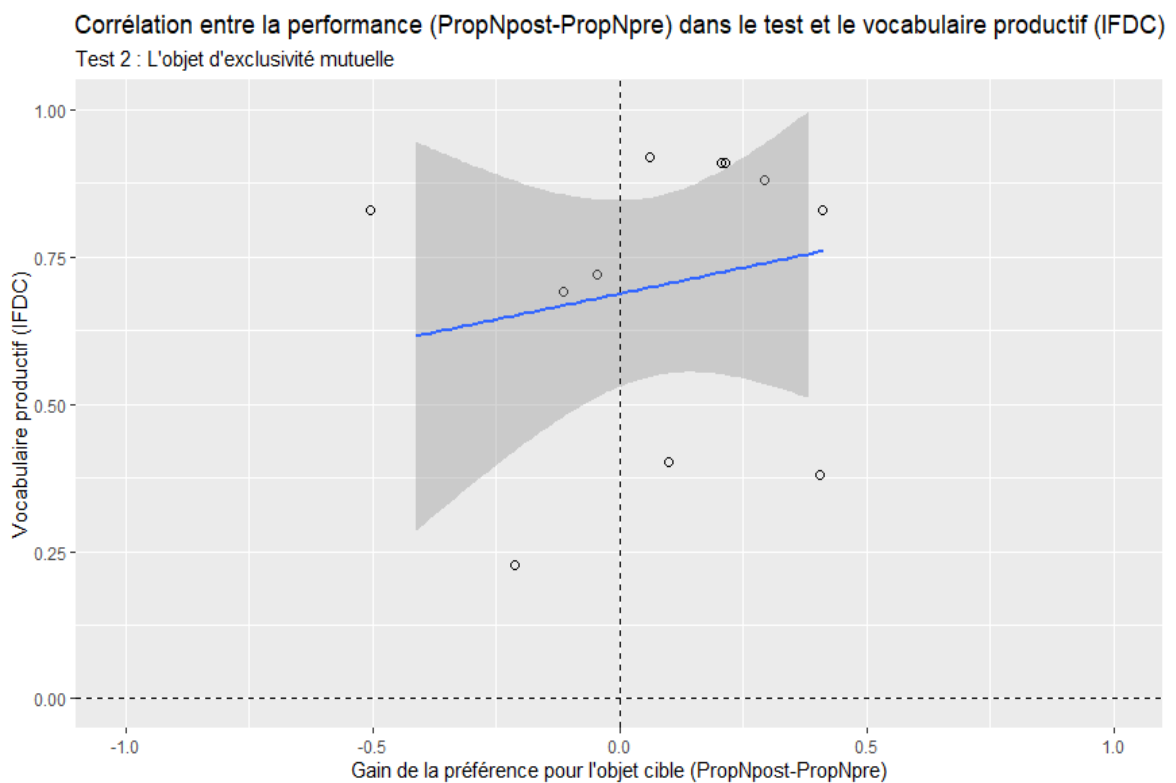
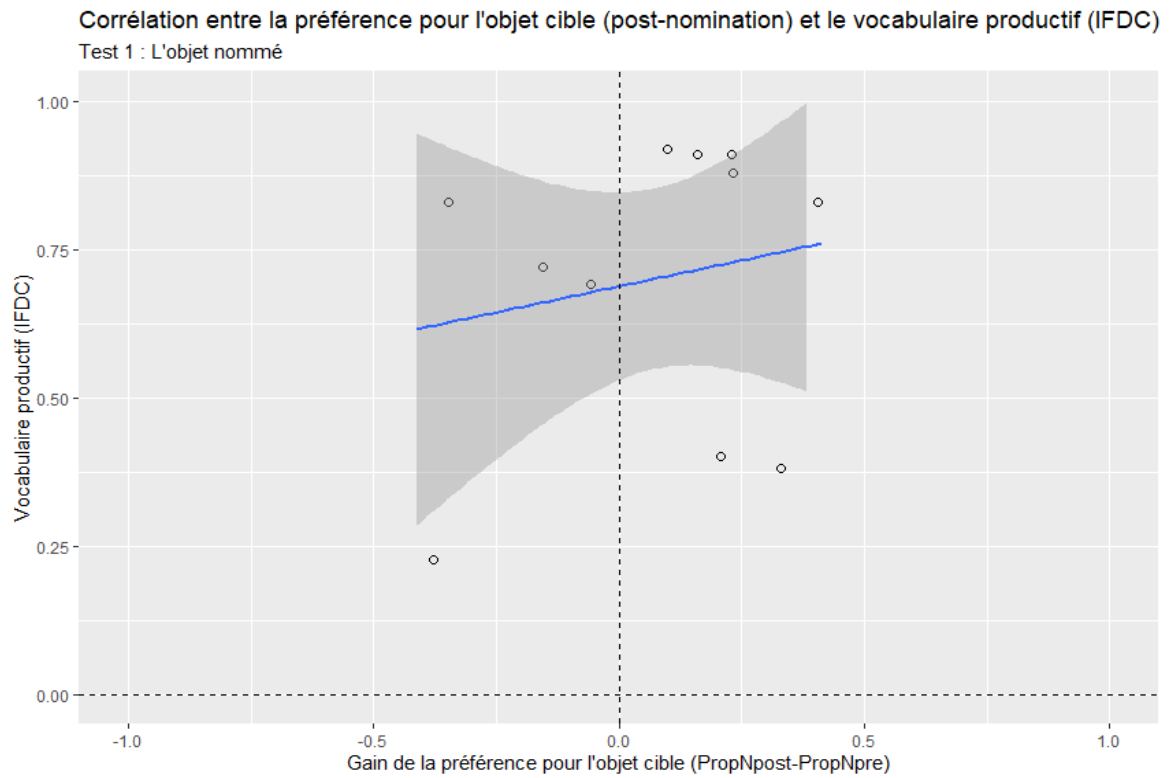


Figure 14 : Préférence pour l'objet cible dans la période de post-nomination (PropNpost-PropNpre) dans les tests (abscisse) prédit par le vocabulaire productif (IFCD) (ordonnée). En haut : test de l'objet nommé. En bas : test de l'objet d'exclusivité mutuelle. Le hasard présenté à 0 par la ligne pointillée. Un point représente les résultats moyennés par participant.

Pour le test de l'objet nommé, l'intercept ($\beta = -0,209$, $SE = 0,17$, $t(10,95) = -1,25$, $p = 0,238$) n'a pas été significativement différent de zéro indiquant qu'il n'y avait pas de gain significative de la proportion du temps de regard vers l'objet cible après sa dénomination. Le nombre de mots n'a pas été significativement différent de zéro non plus ($\beta = 0,307$, $SE = 0,23$, $t(11,28) = 1,34$, $p = 0,205$) indiquant que le vocabulaire productif ne prédit pas la performance dans ce test (PropNpost-PropNpre).

Pour le test de l'objet d'exclusivité mutuelle, l'intercept n'a pas été significativement différent de zéro ($\beta = -0,034$, $SE = 0,28$, $t(7,79) = -0,17$, $p = 0,868$) indiquant qu'il n'y avait pas de gain significatif de la proportion du temps de regard vers l'objet cible après sa dénomination. Le nombre de mots n'a pas été significativement différent de zéro non plus ($\beta = 0,206$, $SE = 0,28$, $t(8,33) = 0,75$, $p = 0,475$) indiquant que le vocabulaire productif ne prédit pas la performance dans ce test (PropNpost-PropNpre).

Ces résultats indiquent que de meilleures performances en apprentissage de nouveaux mots ne peuvent pas être prédits par le vocabulaire productif des participants. Notre hypothèse qui considère que le test de vocabulaire productif (questionnaire parental) est positivement corrélé avec la performance de l'apprentissage n'a pas pu être confirmée.

6. Conclusion de la partie résultats

En résumé, l'hypothèse que le fait de regarder la bouche de l'interlocuteur lui apprenant de nouveaux mots facilite l'apprentissage de ces mots n'a pas pu être constatée. Il n'y avait pas de corrélation significative entre les performances dans les tests et une préférence pour la bouche pendant l'apprentissage des mots testés. Il n'y avait pas de corrélation significative entre les performances dans les tests et le vocabulaire productif.

Pendant la phase de familiarisation (hors tâche d'apprentissage) les nourrissons regardent autant la bouche que les yeux, cependant pendant la phase d'apprentissage ils regardent significativement plus la bouche que les yeux et les autres AOIs. La corrélation entre l'âge et les performances n'a pas pu être testée.

L'apprentissage de mots plus courts (monosyllabiques) semble tout aussi difficile que pour les mots plus longs (trisyllabiques). Contrairement à nos attentes, les analyses suggèrent que les mots trisyllabiques ont été appris dans le test de l'objet d'exclusivité mutuelle.

Les résumés des analyses statistiques sont présentés en annexe.

Discussion

Comme nous l'avons abordé dans le cadre théorique, l'apprentissage des nouvelles associations mot-objet est un processus très complexe qui demande de recourir à de multiples compétences cognitives. Dans la vie quotidienne, pour parvenir à déduire le sens d'un mot concret, le nourrisson doit sélectionner des informations pertinentes dans le signal de parole auditif et visuel. Des mécanismes attentionnels lui permettent de se focaliser sur ces informations tout en ignorant d'autres informations qui surviennent simultanément. De plus, les mécanismes de la mémoire de travail lui permettent de maintenir ces informations le temps de l'apprentissage de ce mot.

Dans le cadre de notre étude, nous mettons l'accent sur le rôle de la bouche du locuteur pour apprendre un mot. Étant donné que la bouche porte des informations visuo-articulatoires pertinentes, regarder la bouche de son interlocuteur pourrait améliorer la mémorisation des mots et des phonèmes qui les composent. Cette stratégie pourrait ainsi être une stratégie primordiale quant à l'apprentissage de nouvelles associations mot-objet. Toutefois, traiter ces informations linguistiques visuelles et auditives demande des mécanismes attentionnels descendants ainsi que les mécanismes de mémoire de travail suffisamment développés. Chez le nourrisson, ces mécanismes cognitifs sont limités car ils se trouvent encore en plein développement (Davidson et al., 2006; Reynolds & Richards, 2009).

Nous avons formulé, dans notre étude, l'hypothèse que l'immaturation de ces mécanismes cognitifs peut contraindre le processus de l'apprentissage de nouveaux mots. Pour étudier l'influence de l'immaturation des mécanismes attentionnels et de la mémoire de travail sur l'acquisition du langage dans le contexte audiovisuel, nous avons prévu de tester deux groupes de tranches d'âge différentes : 18 à 20 mois et 24 à 28 mois. Nous supposons que les participants plus jeunes apprennent moins de nouvelles associations mot-objet que les participants plus âgés, en particulier lorsque la mémoire de travail est sollicitée (mots monosyllabiques vs. mots trisyllabiques). Vu le faible nombre des participants (N = 16) en raison de la pandémie mondiale COVID-19, nous n'avons pas pu tester cette hypothèse. Ainsi, nous avons effectué nos analyses sur un seul groupe (N = 11) d'une tranche d'âge de 20 à 28 mois. Ce groupe était divisé en deux groupes de conditions de test différentes : un groupe apprenant des mots monosyllabiques (n = 7) et un groupe apprenant des mots trisyllabiques (n = 4).

L'influence de l'immaturité des mécanismes cognitifs sur l'apprentissage de nouvelles associations mot-objet

En ce qui concerne l'ensemble du groupe de test, nous nous attendions à ce que les nourrissons aient globalement de meilleures performances en apprentissage de mots monosyllabiques qu'en apprentissage de mots trisyllabiques car les mécanismes cognitifs nécessaires au traitement du langage, notamment les mécanismes de la mémoire de travail, sont encore en cours de développement. Cette hypothèse n'a pas pu être confirmée.

Dans notre étude, la valeur du « gain », représentant une augmentation du temps de regard à l'objet cible, a été considéré comme la mesure d'un apprentissage réussi. Concernant les résultats des données de participants dans la condition de mots monosyllabiques, cette valeur de gain n'a pas été significative, ni pour le premier test de l'objet nommé, ni pour le deuxième test de l'objet d'exclusivité mutuelle. Ces résultats montrent que les participants dans cette condition n'ont pas appris le mot de l'objet nommé. Ils ne sont pas parvenus non plus à déduire le mot d'exclusivité mutuelle.

Les résultats des données des participants dans la condition de mots trisyllabiques ont attiré plus d'attention. Les participants dans cette condition n'ont pas montré des résultats positifs concernant le premier test de l'objet nommé, allant ainsi dans le même sens que les résultats de la condition de mots monosyllabiques. De manière surprenante, les résultats pour le deuxième test de l'exclusivité mutuelle étaient positifs et significatifs. Des tels résultats signifient que les nourrissons dans cette condition semblent avoir pu, bien qu'ils n'aient pas appris le mot de l'objet nommé du premier test, déduire le mot d'exclusivité mutuelle qu'ils n'avaient jamais entendu auparavant. Ces résultats sont assez surprenants car ils ne correspondent pas à nos hypothèses ou à la littérature existante. Il est très peu probable que les participants aient sélectionné avec succès l'objet cible dans le test d'exclusivité mutuelle s'ils n'ont pas appris le mot dans le test précédent, celui de l'objet nommé.

Utilisant la même procédure dans leur étude, Yurovsky et Frank (2017) montrent que les nourrissons réussissent à apprendre un nouveau mot (test de l'objet nommé) à partir de l'âge de 12 mois. C'est seulement à partir de l'âge de 18 mois que les nourrissons parviennent à déduire le sens des mots des objets d'exclusivité mutuelle. Dans ce sens, ces résultats indiquent que le test pour l'objet nommé semble être plus facile que le test pour l'objet d'exclusivité mutuelle. Par ailleurs, dans leur étude, les objets ont été nommés que

trois fois indiquant que les nourrissons encore plus jeunes (âgés de 18 mois) que ceux participant à notre étude réussissent à apprendre une nouvelle association mot-objet qu'ils n'ont entendu que quelques fois. Par conséquent, les participants de notre étude, âgés entre 20 et 28 mois auraient dû être en mesure de réussir dans le test de l'objet nommé aussi bien que dans le test de l'objet d'exclusivité mutuelle.

En revanche, bien que la conception de leur étude coïncide en grande partie avec la nôtre, les phases d'apprentissage se distinguaient en un point : la locutrice dans leur étude ne regardait pas les deux objets qui ont été présentés. La direction du regard ne s'orientait que vers l'objet nommé. On pourrait imaginer que la concentration du regard sur un seul objet permet au nourrisson de mieux se focaliser sur l'objet qui doit être appris. Par conséquent, le fait que la locutrice tourne la tête vers les deux objets (objet nommé et objet d'exclusivité mutuelle) dans notre étude augmente les informations à traiter et ainsi la charge attentionnelle du nourrisson. Il est possible que le contexte d'apprentissage que nous avons proposé dans notre étude soit plus complexe que celui de Yurovsky et Frank (2017) : le fait que l'attention des nourrissons soit dirigée vers les deux objets a pu les conduire à devoir traiter trop d'informations en même temps les empêchant d'apprendre les associations mot-objets.

De même, il est assez surprenant que les participants de notre étude soient en mesure d'apprendre les mots trisyllabiques mais pas les mots monosyllabiques (en ce qui concerne le test de l'objet d'exclusivité mutuelle). Compte tenu du fait que les capacités de mémoire de travail nécessaires pour apprendre des mots trisyllabiques sont plus importantes que celles nécessaires pour apprendre les mots monosyllabiques, ce résultat semble assez difficile à interpréter au premier abord. Néanmoins, il apparaît peu probable que les nourrissons de notre étude aient besoin de mots trisyllabiques (composés de six phonèmes différents) plutôt que de mots monosyllabiques (composés de trois phonèmes différents) pour arriver à différencier les paires de mots utilisés dans cette étude. En effet, des travaux ultérieurs (Nazzi, 2006) ont étudié chez les nourrissons l'usage de l'information phonétique lors de l'acquisition de nouveaux mots. Ils ont montré qu'à l'âge de 20 mois, les nourrissons réussissent à distinguer des mots qui diffèrent seulement par l'une de leurs consonnes, indépendamment de leur position dans le mot. En revanche, ils échouent lorsque les deux mots diffèrent seulement par l'une de leurs voyelles. Les nourrissons sont ainsi sensibles à l'ensemble des contrastes phonologiques des consonnes au niveau perceptif et au niveau lexical à cet âge.

De plus, la création des pseudo-mots dans notre recherche s'est basée sur l'étude de Havy et Zesiger (2017) qui utilise des mots monosyllabiques à consonance française. Ces mots se distinguaient phonétiquement et articulatoirement et ont été faciles à distinguer pour les nourrissons participant à l'expérience. Étant donné que les pseudo-mots que nous avons utilisé se sont basés sur ceux de Havy et Zesiger et se différencient dans leurs consonnes (Nazzi, 2006), ils auraient dû être faciles à distinguer pour les nourrissons participant à notre étude.

Il est envisageable que le choix des objets a joué un rôle quant à un éventuel biais des objets. En ce qui concerne le choix des objets, un possible biais de préférence a été contrôlé auparavant : toutes les données avec un biais visuel pour l'objet cible pendant la phase d'apprentissage ont été exclu de nos analyses. De plus, en analysant les proportions du temps de regard aux différentes AOIs pendant la phase d'apprentissage, aucune différence significative entre le temps de regard à l'objet d'exclusivité mutuelle et le temps de regard à l'objet nommé n'a pas pu être constatée. Les nourrissons ne regardaient donc pas significativement plus un objet ou l'autre dans la phase d'apprentissage. De même, tous les essais des phases de test pendant lesquels les sujets n'ont pas exploré les deux objets avant l'objet a été nommé ont été également exclus des analyses afin de contrôler un éventuel biais d'objets.

Après avoir discuté de plusieurs possibilités qui pourraient expliquer les résultats obtenus, il reste à relever deux points importants : la méthodologie de calcul d'un apprentissage dans les tests ainsi que le pouvoir statistique.

L'indicateur qui décrit l'apprentissage dans notre expérience a été le gain dans la proportion du regard après la dénomination de l'objet cible (PropNpost-PropNpre). Cette comparaison de temps de regard était basée sur les moyennes de deux fenêtres temporelles (pré- et post-nomination). Par conséquent, le calcul des proportions ne considère pas la dimension temporelle. Il ne peut donc fournir qu'un résumé général de l'exploration visuelle des objets. Le temps après que l'objet ait été nommé était de 3,5 secondes. Si nous supposons que les nourrissons ont exploré juste pour un court moment l'objet nommé directement après sa dénomination, une telle tendance n'aurait pas pu être détectable avec notre méthode d'examen vu qu'elle a pris en compte les valeurs moyennes de la durée du regard sur un AOI.

Pour évaluer si les nourrissons regardaient potentiellement le bon objet dans la condition de test de l'objet nommé mais de manière plus faible, nous devrions effectuer une analyse de décours temporel. C'est une analyse plus fine et approfondie des données de suivi oculaire collectées. Dans leur étude, Yurovsky et Frank (2017) ont utilisé cette méthode pour traiter leurs données oculométriques d'une tâche d'apprentissage de nouveaux mots. Ainsi, le premier regard vers l'objet cible après sa dénomination indique l'apprentissage de ce mot.

Par conséquent, analyser nos données de la même façon, permettrait de prendre en considération l'aspect temporel de l'exploration visuelle. Une telle analyse n'était pas possible dans le cadre de notre travail. Toutefois, les données de suivi oculaire que nous avons collectées pourraient être utilisées pour une telle analyse complémentaire.

Finalement, il est important de souligner le faible pouvoir statistique de notre étude comme explication à nos résultats non significatifs. Pour savoir si cet effet de résultats non significatifs des tests (à l'exception du test de l'objet d'exclusivité mutuelle de la condition de mots trisyllabiques) continue, il faudrait tester un plus grand nombre de participants. Comme la taille de l'échantillon dans notre étude était très petite (condition de mots monosyllabiques : $n = 7$; condition de mots trisyllabiques : $n = 4$), elle devra être poursuivie avec un plus grand nombre de participants (e.g., 30 participants par groupe d'âge et par condition).

L'attention à la bouche pendant la phase de familiarisation et la phase de l'apprentissage de nouveaux mots

Grâce aux capacités des mécanismes cognitifs commençant à être fonctionnelles, nous supposons que le nourrisson devrait réussir à se détacher des yeux du locuteur à des moments précis pour bénéficier des informations visuo-articulatoires fournies par la bouche de l'interlocuteur. Il pourrait ainsi mieux identifier et traiter les mots et les phonèmes qui les composent. Pour étudier comment les nourrissons explorent le visage d'un interlocuteur pendant une tâche d'apprentissage de nouveaux mots, nous avons analysé les proportions du temps de regard vers les différentes Aires d'Intérêt (AOIs).

La phase de familiarisation est la phase dans laquelle les nourrissons sont présentés pour la première fois à l'interlocutrice. Cette phase sert à ce que les nourrissons peuvent s'habituer à la personne. Elle a une durée de 7 secondes. Nos analyses ne montrent pas une différence significative entre le temps de regard vers l'AOI de la bouche et l'AOI

des yeux. Contrairement à ce qui a été observé dans la littérature, cela signifie que les nourrissons regardaient autant de temps la bouche que les yeux, mais pas plus souvent la bouche. En effet, il est reporté de plus en plus que les nourrissons profitent encore plus longtemps des informations visuo-articulatoires fournies par la bouche, et cela jusqu'à des stades avancés de leur enfance (Irwin et al., 2017).

Cela pourrait être dû à un biais pour les yeux au début de la phase de familiarisation. Dans la littérature, plusieurs études montrent un biais pour les yeux pendant les premières secondes étant présenté à un inconnu. En regardant le décours temporel dans une communication, les nourrissons commencent à regarder ses yeux pendant les premières secondes pour mieux saisir son interlocuteur, aussi lorsque celui-ci parle (Méary et al., 2018).

Un biais pour les yeux est également reporté par l'étude de Senju et Johnson (2009). Notre étude réduit ce biais. Il est à noter ici que la durée de cette phase de familiarisation de notre test a été très courte avec seulement 7 secondes. Un biais pour les yeux est envisageable au début de la présentation du visage parlant (par exemple, les 3 premières secondes), puis un biais vers la bouche le reste de la phase de familiarisation (par exemple, les 4 dernières secondes) mais il n'est pas vérifiable avec notre méthode d'analyse. Il faudrait faire une analyse de décours temporel permettant d'examiner le regard des nourrissons au visage parlant de la locutrice à différents moments de la phase de familiarisation.

Des études effectuées précédemment ont décrit un regard vers la bouche pendant la deuxième moitié de la première année de vie (entre 10 et 12 mois), afin de mieux identifier et apprendre les phonèmes de sa langue maternelle mais qui s'affaiblit (Lewkowicz & Hansen-Tift, 2012), il est reporté de plus en plus que les nourrissons profitent encore plus longtemps des informations visuo-articulatoires fournies par la bouche, et cela jusqu'à des stades avancés de leur enfance (Fort et al., 2018).

En accord avec cette idée, l'analyse de nos données a montré un biais pour l'AOI de la bouche de l'interlocutrice pendant la phase d'apprentissage. Cela indique que les nourrissons passent significativement plus du temps à regarder l'AOI de la bouche que l'AOI des yeux ou les AOIs des objets après la phase de familiarisation. Cela pourrait signifier que les nourrissons ont été capables de se détacher des yeux de l'interlocutrice

après s'être s'habituer à elle. Cependant, d'autres études seraient nécessaires afin de confirmer cette hypothèse.

Regarder la bouche d'un visage parlant pendant la prononciation de nouveaux mots et un meilleur apprentissage de ces mots

Le biais pour la bouche que nous avons trouvé grâce à notre analyse pourrait signifier que regarder la région de la bouche d'un visage parlant, lui apprenant des nouveaux mots, est une stratégie importante des nourrissons vers l'âge de deux ans pour identifier et traiter ces mots et ainsi mieux les apprendre. Regarder les informations pertinentes pourrait améliorer la saillance du signal de parole entier facilitant l'encodage lexicale dans la mémoire (Bahrick & Lickliter, 2014).

Nous avons analysé si de meilleures performances en apprentissage de nouveaux mots dans les tests peuvent être prédites par une préférence pour la bouche de l'interlocutrice pendant la phase d'apprentissage de nouveaux mots. Nos analyses n'ont pas montré de corrélation entre une préférence pour la bouche pendant la phase d'apprentissage. Notre hypothèse donc n'a pas pu être confirmée.

Cela peut également être dû au fait que la force statistique du test était très faible. De plus, il est possible qu'il n'y ait pas eu de corrélation positive à cause des résultats non significatifs dans les deux tests (à l'exception du test de l'objet d'exclusivité mutuelle dans la condition trisyllabique). Il est également envisageable qu'ils regardent plus la bouche mais cela ne les aide pas suffisamment. Ceci démontre que leurs mécanismes cognitifs sont toujours en train de se développer.

Lien entre le vocabulaire productif et la performance dans les tests

Des études longitudinales de Tenenbaum et collègues (2015) ainsi que Young et collègues (2009) mettent en évidence une corrélation positive entre la préférence pour la bouche de l'interlocuteur avec leur développement du vocabulaire productif ultérieur.

Nous avons supposé qu'un meilleur apprentissage de mots est positivement corrélé avec le développement des mécanismes cognitifs et ainsi avec le vocabulaire productif des nourrissons. De ce fait, nous avons analysé un possible lien entre le vocabulaire productif et la performance dans les tests. Les résultats n'ont pas montré une telle corrélation.

D'un côté, c'est parce que la performance dans les tests a été non significatif dans la plupart des cas, mais les résultats du questionnaire ont été extrêmement positifs pour tous les participants. Il est à noter que nous utilisons deux questionnaires différents (voir partie méthodologie) : Un questionnaire de 18 mois ($n = 3$) et un questionnaire de 24 mois ($n = 8$). Ces questionnaires proposent des mots différents et semble être mieux adaptés pour vérifier un développement langagier à travers l'ensemble des questionnaires (étude longitudinale). En perspective, il serait judicieux d'utiliser des questionnaires identiques entre les âges pour évaluer le vocabulaire productif des participants.

Conclusion

Avec cette étude, nous avons tenté de mieux comprendre les mécanismes cognitifs sous-jacents quant à l'acquisition du langage. De ce fait, nous avons supposé que des mécanismes attentionnels ainsi que des mécanismes de mémoire de travail, supplémentaires aux mécanismes d'apprentissage statistique, seraient nécessaires pour sélectionner et maintenir dans la mémoire les informations importantes qui se trouvent dans le signal de parole pour apprendre une nouvelle association mot-objet. Plus précisément, nous nous sommes intéressés au rôle de l'attention visuelle, dans ce processus, notamment le regard sur la bouche des locuteurs. Nous avons également supposé que le fait de pouvoir profiter de ces informations pertinentes fournies par la bouche est contraint par l'immaturation des mécanismes attentionnels et de la mémoire de travail.

La crise sanitaire mondiale nous a empêché de collecter assez de données et de tester toutes nos hypothèses. Au début, nous avions prévu de tester deux groupes de nourrissons de tranches d'âge différents afin d'évaluer le rôle du développement cognitif. Finalement, nous avons effectué nos analyses sur les données d'une seule tranche d'âge (20-28 mois) et une très faible fréquence d'échantillonnage ($N = 11$). Par conséquent, avec notre étude nous ne pouvons pas tirer de conclusions générales et justifiées, la plupart de nos résultats n'ayant pas été significative. Les questions auxquelles nous avons tenté de répondre dans ce travail de recherche nécessitent, de fait, un échantillon plus grand.

Notre travail a relevé des inconvénients par rapport à une analyse de données d'oculomètre à travers les données d'exploration visuelle moyennées quant à l'évaluation d'apprentissage de nouveaux mots. Même si nous n'avons pas été capables de valider la plupart de nos hypothèses, nous avons réussi à créer un paradigme d'une tâche d'apprentissage de nouveaux mots français qui est adaptable et utilisable pour d'autres études futures à ce sujet.

Bibliographie

- Babineau, M., & Peperkamp, S. (2017). L'acquisition du langage chez le bébé. *L'éducation à La Lumière Des Neurosciences. Actes Du 12ème Colloque Petite Enfance*, 55–62.
- Bahrick, L. E., & Lickliter, R. (2014). Learning to Attend Selectively: The Dual Role of Intersensory Redundancy. *Current Directions in Psychological Science*, 23(6), 414–420. <https://doi.org/10.1177/0963721414549187>
- Barenholtz, E., Mavica, L., & Lewkowicz, D. J. (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition*, 147, 100–105.
- Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R Package Version*, 2(1), 74.
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, 109(9), 3253–3258.
- Bovet, F., Danjou, G., Langue, J., Moretto, M., Tockert, E., & Kern, S. (2005). Les inventaires français du développement communicatif (IFCD): un nouvel outil pour évaluer le développement communicatif du nourrisson. *Médecine et Enfance*, 25(6), 327–332.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, 52(1), 337–367.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, 34(4), 613–617.
- Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25(2), 141–168.
- Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1149–1157.
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037–2078.
- de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, 7, 44.
- Everdell, I. T., Marsh, H., Yurick, M. D., Munhall, K. G., & Paré, M. (2007). Gaze behaviour in audiovisual speech perception: Asymmetrical distribution of face-directed fixations. *Perception*, 36(10), 1535–1545.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10(3), 279–293.
- Fort, M., Ayneto-Gimeno, A., Escrichs, A., & Sebastian-Galles, N. (2018). Impact of bilingualism on infants' ability to learn from talking and nontalking faces. *Language Learning*, 68, 31–57.
- Fort, M., Kandel, S., Chipot, J., Savariaux, C., Granjon, L., & Spinelli, E. (2013). Seeing the initial articulatory gestures of a word triggers lexical access. *Language and Cognitive Processes*, 28(8), 1207–1223.
- Fort, M., Spinelli, E., Savariaux, C., & Kandel, S. (2010). The word superiority effect in audiovisual speech perception. *Speech Communication*, 52(6), 525–532.

- Gillet, P., & Barthélémy, C. (2011). Développement de l'attention chez le petit enfant: implications pour les troubles autistiques. *Développements*, 3, 17–25.
- Gliga, T., Elsabbagh, M., Hudry, K., Charman, T., Johnson, M. H., & team, B. (2012). Gaze following, gaze reading, and word learning in children at risk for autism. *Child Development*, 83(3), 926–938.
- Gogate, L. J., & Bahrick, L. E. (2001). Intersensory redundancy and 7-month-old infants' memory for arbitrary syllable-object relations. *Infancy*, 2(2), 219–231.
- Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). The intersensory origins of word-comprehension: an ecological–dynamic systems view. *Developmental Science*, 4(1), 1–18.
- Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70(2), 109–135.
- Havy, M., & Zesiger, P. (2017). Learning spoken words via the ears and eyes: evidence from 30-month-old children. *Frontiers in Psychology*, 8, 2122.
- Hothorn, T., Bretz, F., Westfall, P., & Heiberger, R. M. (2008). Multcomp: simultaneous inference for general linear hypotheses. *R Package Version*, 0–1.
- Irwin, J., Brancazio, L., & Volpe, N. (2017). The development of gaze to a speaking face. *The Journal of the Acoustical Society of America*, 141(5), 3145–3150.
- Jackson, P. L., Montgomery, A. A., & Binnie, C. A. (1976). Perceptual dimensions underlying vowel lipreading performance. *Journal of Speech and Hearing Research*, 19(4), 796–812.
- Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, 13(2), 339–345.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, 97(22), 11850–11857.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Lewkowicz, D. J., & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences*, 109(5), 1431–1436.
- Magezi, D. A. (2015). Linear mixed-effects models for within-participant psychology experiments: an introductory tutorial and free, graphical user interface (LMMgui). *Frontiers in Psychology*, 6, 2.
- Méary, D., Jaggie, C., & Pascalis, O. (2018). Multisensory representation of gender in infants: An eye-tracking study. *Language Learning*, 68, 14–30.
- Mintz, T. H. (1996). *The roles of linguistic input and innate mechanisms in children's acquisition of grammatical categories*. University of Rochester Rochester, NY.
- Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: a balancing act across the lifespan. *Trends in Neurosciences*, 39(8), 567–579.
- Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, 98(1), 13–30.
- Nazzi, T. (2006). No Title. *Recherches Linguistiques de Vincennes*, 35(1), 2. https://www.cairn-int.info/load_pdf.php?ID_ARTICLE=E_RLV_035_0002
- Newman, R. S. (2005). The cocktail party effect in infants revisited: listening to one's name in noise. *Developmental Psychology*, 41(2), 352.
- Patterson, M. L., & Werker, J. F. (2003). Two-month-old infants match phonetic information in lips and voice. *Developmental Science*, 6(2), 191–196.
- Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35, 73–89.

- Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science*, 26(4), 490–498.
- Pons, F., Sanz-Torrent, M., Ferinu, L., Birulés, J., & Andreu, L. (2018). Children with SLI can exhibit reduced attention to a talker's mouth. *Language Learning*, 68, 180–192.
- Quine, W. V. O. (1961). Word and Object. *Philos. Phenomenol. Res.*, 22(1), 115.
- Reynolds, G. D., & Richards, J. E. (2009). Cortical source localization of infant cognition. *Developmental Neuropsychology*, 34(3), 312–329.
- Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science*, 19(1), 41–46.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928.
- Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), 668–671.
- Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, 13(3), 127–134.
- Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development*, 9(1), 25–49.
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. <https://doi.org/10.1016/j.cognition.2007.06.010>
- Team, R. C. (2013). *R: A language and environment for statistical computing*. Vienna, Austria.
- Tenenbaum, E. J., Sobel, D. M., Sheinkopf, S. J., Malle, B. F., & Morgan, J. L. (2015). Attention to the mouth and gaze following in infancy predict language development. *Journal of Child Language*, 42(6), 1173–1190.
- Vihman, M. M. (1996). *Phonological Development: The Origins of Language in the Child (Applied Language Studies)*.
- Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, 127(3), 375–382.
- Young, G. S., Merin, N., Rogers, S. J., & Ozonoff, S. (2009). Gaze behavior and affect at 6 months: predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Developmental Science*, 12(5), 798–814.
- Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244–262.
- Yu, C., & Smith, L. B. (2016). The social origins of sustained attention in one-year-old human infants. *Current Biology*, 26(9), 1235–1240.
- Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: salience and social cues in early word learning. *Developmental Science*, 20(2), 10.1111/desc.12349. <https://doi.org/10.1111/desc.12349>

Table des illustrations

Figure 1 : Situations d'apprentissage de nouvelles associations mot-objet	12
Figure 2 : Développement des différents systèmes attentionnels.....	16
Figure 3 : Maturation de l'attention ascendante et descendante.....	19
Figure 4 : Les six paires d'objets.....	30
Figure 5: Deux exemples de présentation d'objets.....	31
Figure 6 : Dispositif expérimental	32
Figure 7 : Représentation schématique de l'ensemble de la procédure.....	35
Figure 8 : Illustration des AOIs définies pour la phase d'apprentissage	40
Figure 9 : Proportion du temps de regard pour les AOIs pendant la familiarisation.....	42
Figure 10 : Proportion du temps de regard pour les AOIs pendant l'apprentissage.....	43
Figure 11 : Proportion du temps de regard à l'objet cible (Test 1)	45
Figure 12 : Proportion du temps de regard à l'objet cible (Test 2)	46
Figure 13 : Corrélation entre la préférence portée à la bouche et les performances	48
Figure 14 : Corrélation entre le vocabulaire productif et les performances	50
Tableau 1 : Illustration du déroulé de l'expérience de Saffran et al. (1996)	11
Tableau 2 : Informations démographiques des participants	24
Tableau 3 : Évaluation des réponses aux questionnaires.....	26
Tableau 4 : Les pseudo-mots	27

Table des annexes

Annexe 1 : Questionnaire de 18 mois.....	66
Annexe 2 : Questionnaire de 24 mois.....	67
Annexe 3 : Feuille d'évaluation.....	68
Annexe 4 : Résumés des analyses statistiques.....	69

Annexe 1

Questionnaire de 18 mois

QUESTIONNAIRE DE 18 MOIS (à remplir par les parents)

© S. Kern

Nom Prénom

Sexe Date de naissance

Nombre de frères et sœurs Ages des frères et sœurs

Mode de garde dominant

Langue que l'enfant entend le plus souvent

Langues parlées à la maison

Personne qui remplit le questionnaire (père, mère, les deux) Date

Cochez :

- les cases de la colonne C (compris) pour les mots que l'enfant comprend mais ne dit pas encore (vous pouvez considérer que l'enfant comprend un mot même s'il ne le comprend que dans une seule situation) ;
- les cases de la colonne CD (compris et dit) pour les mots que l'enfant comprend et qu'il utilise actuellement de manière spontanée (si sa prononciation est différente de celle des adultes, cochez tout de même le mot).

	C	CD		C	CD		C	CD
attention	<input type="checkbox"/>	<input type="checkbox"/>	cuillère	<input type="checkbox"/>	<input type="checkbox"/>	nom de l'enfant	<input type="checkbox"/>	<input type="checkbox"/>
au revoir	<input type="checkbox"/>	<input type="checkbox"/>	cuisine	<input type="checkbox"/>	<input type="checkbox"/>	non	<input type="checkbox"/>	<input type="checkbox"/>
avion	<input type="checkbox"/>	<input type="checkbox"/>	dehors	<input type="checkbox"/>	<input type="checkbox"/>	oreille	<input type="checkbox"/>	<input type="checkbox"/>
a/avoir faim	<input type="checkbox"/>	<input type="checkbox"/>	derrière	<input type="checkbox"/>	<input type="checkbox"/>	où	<input type="checkbox"/>	<input type="checkbox"/>
a/avoir soif	<input type="checkbox"/>	<input type="checkbox"/>	donne/donner	<input type="checkbox"/>	<input type="checkbox"/>	pantalon	<input type="checkbox"/>	<input type="checkbox"/>
bain	<input type="checkbox"/>	<input type="checkbox"/>	dors/dormir/faire dodo	<input type="checkbox"/>	<input type="checkbox"/>	parc (public)	<input type="checkbox"/>	<input type="checkbox"/>
balle	<input type="checkbox"/>	<input type="checkbox"/>	eau	<input type="checkbox"/>	<input type="checkbox"/>	photo	<input type="checkbox"/>	<input type="checkbox"/>
banane	<input type="checkbox"/>	<input type="checkbox"/>	encore	<input type="checkbox"/>	<input type="checkbox"/>	pied	<input type="checkbox"/>	<input type="checkbox"/>
bébé	<input type="checkbox"/>	<input type="checkbox"/>	ferme/fermer	<input type="checkbox"/>	<input type="checkbox"/>	poisson	<input type="checkbox"/>	<input type="checkbox"/>
béé béé	<input type="checkbox"/>	<input type="checkbox"/>	fourchette	<input type="checkbox"/>	<input type="checkbox"/>	porte	<input type="checkbox"/>	<input type="checkbox"/>
bien	<input type="checkbox"/>	<input type="checkbox"/>	frigo	<input type="checkbox"/>	<input type="checkbox"/>	poubelle	<input type="checkbox"/>	<input type="checkbox"/>
bois/boire	<input type="checkbox"/>	<input type="checkbox"/>	fromage	<input type="checkbox"/>	<input type="checkbox"/>	poule	<input type="checkbox"/>	<input type="checkbox"/>
bon/ne	<input type="checkbox"/>	<input type="checkbox"/>	gâteau/biscuit	<input type="checkbox"/>	<input type="checkbox"/>	poupée	<input type="checkbox"/>	<input type="checkbox"/>
bonjour	<input type="checkbox"/>	<input type="checkbox"/>	gentil/gentille	<input type="checkbox"/>	<input type="checkbox"/>	poussette/pousse-pousse	<input type="checkbox"/>	<input type="checkbox"/>
bonne nuit	<input type="checkbox"/>	<input type="checkbox"/>	jette/jeter	<input type="checkbox"/>	<input type="checkbox"/>	prends/prendre	<input type="checkbox"/>	<input type="checkbox"/>
bouteille	<input type="checkbox"/>	<input type="checkbox"/>	joue/jouer	<input type="checkbox"/>	<input type="checkbox"/>	pyjama	<input type="checkbox"/>	<input type="checkbox"/>
bras	<input type="checkbox"/>	<input type="checkbox"/>	là	<input type="checkbox"/>	<input type="checkbox"/>	regarde/regarder	<input type="checkbox"/>	<input type="checkbox"/>
bravo	<input type="checkbox"/>	<input type="checkbox"/>	lait	<input type="checkbox"/>	<input type="checkbox"/>	sale	<input type="checkbox"/>	<input type="checkbox"/>
brosse à dent	<input type="checkbox"/>	<input type="checkbox"/>	lapin	<input type="checkbox"/>	<input type="checkbox"/>	sieste	<input type="checkbox"/>	<input type="checkbox"/>
canapé/divan	<input type="checkbox"/>	<input type="checkbox"/>	lave/laver	<input type="checkbox"/>	<input type="checkbox"/>	soupe	<input type="checkbox"/>	<input type="checkbox"/>
canard	<input type="checkbox"/>	<input type="checkbox"/>	lit	<input type="checkbox"/>	<input type="checkbox"/>	sucette/lolette/tutte	<input type="checkbox"/>	<input type="checkbox"/>
chambre	<input type="checkbox"/>	<input type="checkbox"/>	livre	<input type="checkbox"/>	<input type="checkbox"/>	table	<input type="checkbox"/>	<input type="checkbox"/>
chapeau/bonnet	<input type="checkbox"/>	<input type="checkbox"/>	lunettes	<input type="checkbox"/>	<input type="checkbox"/>	tape/taper	<input type="checkbox"/>	<input type="checkbox"/>
chat	<input type="checkbox"/>	<input type="checkbox"/>	main	<input type="checkbox"/>	<input type="checkbox"/>	télécommande	<input type="checkbox"/>	<input type="checkbox"/>
chaud/chaude	<input type="checkbox"/>	<input type="checkbox"/>	maintenant	<input type="checkbox"/>	<input type="checkbox"/>	tombe/tomber	<input type="checkbox"/>	<input type="checkbox"/>
chausson/pantoufle	<input type="checkbox"/>	<input type="checkbox"/>	maman	<input type="checkbox"/>	<input type="checkbox"/>	touche/toucher	<input type="checkbox"/>	<input type="checkbox"/>
chaussure/soulier	<input type="checkbox"/>	<input type="checkbox"/>	mange/manger	<input type="checkbox"/>	<input type="checkbox"/>	va/aller	<input type="checkbox"/>	<input type="checkbox"/>
cheval	<input type="checkbox"/>	<input type="checkbox"/>	maison	<input type="checkbox"/>	<input type="checkbox"/>	vache	<input type="checkbox"/>	<input type="checkbox"/>
cheveux	<input type="checkbox"/>	<input type="checkbox"/>	merci	<input type="checkbox"/>	<input type="checkbox"/>	voiture/auto	<input type="checkbox"/>	<input type="checkbox"/>
chut	<input type="checkbox"/>	<input type="checkbox"/>	meuh	<input type="checkbox"/>	<input type="checkbox"/>	vroum	<input type="checkbox"/>	<input type="checkbox"/>
cocorico	<input type="checkbox"/>	<input type="checkbox"/>	miam-miam	<input type="checkbox"/>	<input type="checkbox"/>	yeux	<input type="checkbox"/>	<input type="checkbox"/>
coin-coin	<input type="checkbox"/>	<input type="checkbox"/>	moi	<input type="checkbox"/>	<input type="checkbox"/>			
couche/lange	<input type="checkbox"/>	<input type="checkbox"/>	musique	<input type="checkbox"/>	<input type="checkbox"/>			

Est-ce que l'enfant a déjà commencé à combiner des mots
comme par exemple «gâteau encore» ou «papa pati» ? pas encore quelquefois souvent

Annexe 2

Questionnaire de 24 mois

QUESTIONNAIRE DE 24 MOIS (à remplir par les parents)

© S. Kern

Nom Prénom

Sexe Date de naissance

Nombre de frères et sœurs Ages des frères et sœurs

Mode de garde dominant

Langue que l'enfant entend le plus souvent

Langues parlées à la maison

Personne qui remplit le questionnaire (père, mère, les deux) Date

Cochez les cases correspondant aux mots que l'enfant produit actuellement de manière spontanée (hors imitation). Si sa prononciation est différente de celle des adultes, cochez tout de même le mot.

aïe	<input type="checkbox"/>	cochon	<input type="checkbox"/>	monsieur	<input type="checkbox"/>
allô	<input type="checkbox"/>	coin-coin	<input type="checkbox"/>	moto	<input type="checkbox"/>
assiette	<input type="checkbox"/>	compote	<input type="checkbox"/>	musique	<input type="checkbox"/>
attention	<input type="checkbox"/>	couche/lange	<input type="checkbox"/>	nez	<input type="checkbox"/>
au revoir	<input type="checkbox"/>	coucou	<input type="checkbox"/>	nom de l'enfant	<input type="checkbox"/>
a/avoir peur	<input type="checkbox"/>	cuillère	<input type="checkbox"/>	oreille	<input type="checkbox"/>
ballon	<input type="checkbox"/>	dame	<input type="checkbox"/>	où	<input type="checkbox"/>
bateau	<input type="checkbox"/>	dehors	<input type="checkbox"/>	ouaf-ouaf	<input type="checkbox"/>
beau/belle	<input type="checkbox"/>	eau	<input type="checkbox"/>	pain	<input type="checkbox"/>
bébé	<input type="checkbox"/>	école/crèche	<input type="checkbox"/>	pantalon	<input type="checkbox"/>
bée bée	<input type="checkbox"/>	écrit/écrire	<input type="checkbox"/>	papa	<input type="checkbox"/>
biberon/bibi	<input type="checkbox"/>	éléphant	<input type="checkbox"/>	pars/partir/parti	<input type="checkbox"/>
bois/boire	<input type="checkbox"/>	encore	<input type="checkbox"/>	pas	<input type="checkbox"/>
bon/bonne	<input type="checkbox"/>	fais/faire un bisou	<input type="checkbox"/>	pâtes	<input type="checkbox"/>
bonbons	<input type="checkbox"/>	ferme/fermer	<input type="checkbox"/>	pleure/pleurer	<input type="checkbox"/>
bonjour	<input type="checkbox"/>	fleur	<input type="checkbox"/>	pluie	<input type="checkbox"/>
bottes	<input type="checkbox"/>	froid/froide	<input type="checkbox"/>	poisson	<input type="checkbox"/>
bouche	<input type="checkbox"/>	fromage	<input type="checkbox"/>	pomme	<input type="checkbox"/>
bras	<input type="checkbox"/>	ici	<input type="checkbox"/>	porte	<input type="checkbox"/>
ça	<input type="checkbox"/>	là	<input type="checkbox"/>	pot	<input type="checkbox"/>
cache/cacher	<input type="checkbox"/>	lait	<input type="checkbox"/>	poubelle	<input type="checkbox"/>
cadeau	<input type="checkbox"/>	lapin	<input type="checkbox"/>	pyjama	<input type="checkbox"/>
caillou	<input type="checkbox"/>	lit	<input type="checkbox"/>	quoi	<input type="checkbox"/>
canard	<input type="checkbox"/>	livre	<input type="checkbox"/>	sale	<input type="checkbox"/>
casse/casser	<input type="checkbox"/>	lumière	<input type="checkbox"/>	s'il te plaît	<input type="checkbox"/>
chaise	<input type="checkbox"/>	lune	<input type="checkbox"/>	télé	<input type="checkbox"/>
chat	<input type="checkbox"/>	main	<input type="checkbox"/>	tombe/tomber	<input type="checkbox"/>
chaud/chaude	<input type="checkbox"/>	maison	<input type="checkbox"/>	verre	<input type="checkbox"/>
chaussure/soulier	<input type="checkbox"/>	maman	<input type="checkbox"/>	voiture/auto	<input type="checkbox"/>
cheval	<input type="checkbox"/>	mange/manger	<input type="checkbox"/>	vroum	<input type="checkbox"/>
cheveux	<input type="checkbox"/>	merci	<input type="checkbox"/>	yaourt/yogourt	<input type="checkbox"/>
chien/toutou	<input type="checkbox"/>	meuh	<input type="checkbox"/>	yeux	<input type="checkbox"/>
chocolat	<input type="checkbox"/>	miaou	<input type="checkbox"/>		
chut	<input type="checkbox"/>	moi	<input type="checkbox"/>		

– Est-ce que l'enfant a déjà commencé à combiner des mots comme par exemple «gâteau encore» ou «papa pati»? pas encore quelquefois souvent

– Si vous avez répondu de façon positive à la question précédente, indiquez les trois phrases les plus longues que l'enfant produit de manière spontanée actuellement :

1

2

3

Annexe 3

Feuille d'évaluation

FEUILLE D'ÉVALUATION DES RÉPONSES AUX QUESTIONNAIRES

© S. Kern

Il est recommandé de faire remplir les questionnaires courts à 12, 18 et 24 mois, de façon à décrire un itinéraire de développement communicatif de l'enfant.
 Pour une raison de simplicité, l'interprétation des résultats est la même chez les filles et chez les garçons, bien que dans tous les domaines les performances des filles soient supérieures à celles des garçons.
 Deux résultats inférieurs ou égaux au 10^e percentile sur le même questionnaire ou deux questionnaires d'âges différents représentent des signes d'alerte. L'absence de babillage canonique à 12 mois et l'absence de combinaison de mots à 24 mois sont assimilées à des résultats inférieurs au 10^e percentile.
 Chaque feuille est documentée individuellement pour être gardée dans le dossier de l'enfant ou remise aux parents.

(Mettre une croix dans la case correspondant aux réalisations de l'enfant)

Enfant (nom, prénom, date de naissance) :

12 mois, date :

Babillage canonique	non acquis		acquis (quelquefois ou souvent)			
Babillage diversifié	non acquis		acquis (quelquefois ou souvent)			
Gestes réalisés (nombre)	≤ 9	10-11	12-14	15-16	17-18	≥ 19
Compréhension (nombre de mots compris)	≤ 11	12-19	20-29	30-44	45-54	≥ 55
Production (nombre de mots dits)	0	1	2-3	4-6	7-10	≥ 11
percentiles		10 ^e	25 ^e	50 ^e	75 ^e	90 ^e

18 mois, date :

Compréhension (nombre de mots compris)	≤ 51	52-68	69-83	84-93	94-97	≥ 98
Production (nombre de mots dits)	≤ 4	5-9	10-18	19-28	29-51	≥ 52
Combinaison de mots	non acquise		acquise			
percentiles		10 ^e	25 ^e	50 ^e	75 ^e	90 ^e

24 mois, date :

Production (nombre de mots dits)	≤ 28	29-46	47-74	75-91	92-98	≥ 99	
Longueur moyenne des énoncés*	combinaison non acquise		2,66	3,66	4,66	6	≥ 7
percentiles		10 ^e	25 ^e	50 ^e	75 ^e	90 ^e	

* La longueur moyenne des énoncés est calculée en divisant par 3 le nombre total des mots contenus dans les 3 phrases les plus longues.

Annexe 4

Résumés des analyses statistiques

Tableau 5 : Régression linéaire pour la phase de familiarisation

Paramètres	Estimation (SE)	Test de t	Valeur p
Intercept	0,46 (0,07)	6,45	< 0,001
AOI yeux	0,08 (0,10)	0,82	0,422

Le niveau de référence est l'AOI bouche. SE = Erreur Standard. $R^2 = 0,036$. R ajusté = -0,174.

Tableau 6 : Régression linéaire à effets mixtes du temps de regard pendant la phase d'apprentissage prédit par les AOIs

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	0,40 (0,02)	16,93 (120)	< 0,001
AOI objet ME	-0,26 (0,03)	-7,87 (120)	< 0,001
AOI objet Nommé	-0,27 (0,03)	-7,92 (120)	< 0,001
AOI yeux	-0,07 ((0,03)	-2,37 (120)	0,0195
COMPARAISONS MULTIPLES			
Objet ME - Bouche	0.26 (0.03)	-7.87	< 0,001
Objet Nommé - Bouche	-0.27 (0.03)	-7.93	< 0,001
Yeux - Bouche	-0.08 (0.03)	-2.37.	0,083
Objet Nommé–Objet ME	-0.00 (0.03)	-0.06	0,999
Yeux – Objet ME	0.19 (0.03)	5.50	< 0,001
Yeux – Objet Nommé	0.19 (0.03)	5.56	< 0,001

Le niveau de référence est l'AOI bouche. SE = Erreur Standard. ddl = degrés de liberté. ME = exclusivité mutuelle. La comparaison multiple des moyennes a été fait utilisant la méthode de Tukey.

Tableau 7 : Régression linéaire à effets mixtes du gain de proportion du temps de regard à l'objet cible pendant la phase de Test 1 (objet nommé) prédit par le type de mot

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	0,05 (0,07)	0,67 (11,12)	0,505
Type de mot (Trisyll)	-0,13 (0,12)	-1,12 (13,38)	0,238

Le niveau de référence est la condition de mots monosyllabiques. SE = Erreur Standard. ddl = degrés de liberté.

Tableau 8 : Régression linéaire à effets mixtes du gain de proportion du temps de regard à l'objet cible pendant la phase de Test 2 (objet d'exclusivité mutuelle) prédit par le type de mot

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	0,01 (0,06)	0,19 (8,40)	0,855
Type de mot (Trisyll)	0,29 (0,10)	2,82 (7,80)	0,023

Le niveau de référence est la condition de mots monosyllabiques. SE = Erreur Standard. ddl = degrés de liberté.

Tableau 9 : Modèle linéaire à effets mixtes de la corrélation entre la préférence pour la bouche pendant la phase d'apprentissage et la performance dans le Test 1 (objet nommé)

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	0,00 (0,06)	0,05 (11,67)	0,959
Pref bouche	0,00 (0,15)	0,02 (16,46)	0,985

SE = Erreur Standard. ddl = degrés de liberté. Pref bouche = préférence relative pour la bouche (proportion).

Tableau 10 : Modèle linéaire à effets mixtes de la corrélation entre la préférence pour la bouche pendant la phase d'apprentissage et la performance dans le Test 2 (objet d'exclusivité mutuelle)

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	0,12 (0,07)	1,52 (8,04)	0,165
Pref bouche	-0,03 (0,16)	-0,15 (11,46)	0,880

SE = Erreur Standard. ddl = degrés de liberté. Pref bouche = préférence relative pour la bouche (proportion).

Tableau 11 : Modèle linéaire à effets mixtes de la corrélation entre le vocabulaire productif et la performance dans le Test 1 (objet nommé)

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	-0,21 (0,17)	-1,25 (10,95)	0,238
Nombre de mots	0,31 (0,23)	1,34 (11,28)	0,205

SE = Erreur Standard. ddl = degrés de liberté. Pref bouche = préférence relative pour la bouche (proportion).

Tableau 12 : Modèle linéaire à effets mixtes de l'effet de la préférence pour la bouche pendant la phase d'apprentissage sur la performance dans le Test 2 (objet d'exclusivité mutuelle)

Paramètres	Estimation (SE)	Test de t (ddl)	Valeur p
EFFETS FIXES			
Intercept	-0,03 (0,14)	-0,17 (7,79)	0,868
Nombre de mots	0,21 (0,28)	0,75 (8,33)	0,475

SE = Erreur Standard. ddl = degrés de liberté. Pref bouche = préférence relative pour la bouche (proportion).

MOTS-CLÉS : apprentissage de nouveaux mots, attention sélective, mémoire de travail, développement typique, oculométrie

RÉSUMÉ

Ce travail a tenté de mieux comprendre les mécanismes cognitifs sous-jacents de l'acquisition du langage. Nous avons supposé qu'à côté des mécanismes d'apprentissage statistique, des mécanismes attentionnels et de mémoire de travail sont nécessaires pour qu'un nourrisson puisse sélectionner et maintenir dans la mémoire les informations pertinentes du signal de parole. L'attention visuelle, notamment le regard sur la bouche, pourrait jouer un rôle important quant à l'apprentissage de nouveaux mots. Nous avons supposé que le fait de pouvoir profiter de ces informations pertinentes fournies par la bouche est contraint par l'immaturation des mécanismes attentionnels et de la mémoire de travail. Des nourrissons français (20 à 28 mois) ont été testés sur une tâche d'apprentissage de nouvelles associations mots-objet. L'exploration visuelle d'un visage parlant ainsi que les performances dans les tests ont été évaluées à l'aide de la technique de l'oculométrie. Toutes nos hypothèses n'ont pas pu être vérifiées au vu de notre échantillon relativement faible. La plupart des résultats n'a pas atteint la significativité. Des études complémentaires avec plus de participants seront nécessaires afin de vérifier ces résultats.

KEYWORDS: word learning, selective attention, working memory, typical development, eye tracking

ABSTRACT

This work sought to better understand the cognitive mechanisms underlying language acquisition. We hypothesized that alongside statistical learning mechanisms, attentional and working memory mechanisms are necessary for an infant to select and maintain the relevant information in memory which are presented in the speech signal. Visual attention, especially looking at the mouth of the speaker, could play an important role in word learning. We assumed taking advantage of the important information provided by the mouth is limited by the immaturity of attentional mechanisms and working memory. French infants (20-28 months) were tested in how well they learned new word-object associations. The eye-tracking technique was used to monitor the infant's visual exploration of a talking face as well as their performance in the tests. Not all of our hypotheses could be verified given our relatively small sample size. Most of the results were not statistically significant. Further studies with a bigger sample size are recommended to re-examine the results.