



**HAL**  
open science

# Traitement de la prosodie émotionnelle dans le trouble du spectre de l'autisme : approche comportementale et neurophysiologique

Annabelle Merchie

## ► To cite this version:

Annabelle Merchie. Traitement de la prosodie émotionnelle dans le trouble du spectre de l'autisme : approche comportementale et neurophysiologique. Sciences cognitives. Université de Tours, 2024. Français. <NNT : >. <tel-05082089>

**HAL Id: tel-05082089**

**<https://hal.science/tel-05082089v1>**

Submitted on 23 May 2025

**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.



HAL Authorization

# UNIVERSIT  DE TOURS

** COLE DOCTORALE : Sant , Sciences Biologiques et Chimie du  
Vivant (SSBCV)**

**UMR1253 iBrain –  quipe 1 Autisme et Neurod veloppement**

**TH SE** pr sent e par :  
**Annabelle Merchie**

Soutenue le : **10 d cembre 2024**

Pour obtenir le grade de : **Docteur de l'universit  de Tours**  
Discipline/ Sp cialit  : **Neurosciences**

**TRAITEMENT DE LA PROSODIE  
EMOTIONNELLE DANS LE TROUBLE  
DU SPECTRE DE L'AUTISME**  
*Approche comportementale et  
neurophysiologique*

**TH SE dirig e par :**  
Dr. GOMOT Marie

CR, INSERM, iBrain, Tours

**RAPPORTEURES :**  
Pr. AMESTOY Anouck  
Dr. CACLIN Anne

PU-PH, CNRS, INCIA, Bordeaux  
DR, INSERM, CRNL, Lyon

**PR SIDENTE DU JURY :**  
Pr. FERR  Sandrine

Pr., INSERM, iBrain, Tours

---

**JURY : (directeurs, rapporteurs, examinateurs)**

Pr. AMESTOY Anouck  
Dr. AUCOUTURIER Jean-Julien  
Dr. CACLIN Anne  
Pr. DESMIDT Thomas  
Pr. FERR  Sandrine  
Dr. GOMOT Marie

PU-PH, CNRS, INCIA, Bordeaux  
DR, CNRS, FEMTO-ST, Besan on  
DR, INSERM, CRNL, Lyon  
PU-PH, INSERM, iBrain, Tours  
Pr., INSERM, iBrain, Tours  
CR, INSERM, iBrain, Tours



*À Élise et Alma*



## Remerciements

Tout d'abord, je tiens à remercier Marie Gomot, ma directrice de thèse. Marie, merci d'avoir été là tout au long de cette thèse, de m'avoir permis de m'épanouir dans le laboratoire et dans la recherche. Merci de m'avoir fait confiance, il y a quatre ans, pour commencer cette thèse. Tu m'as poussée dès le début à découvrir de nouvelles méthodes et techniques d'exploration, et je t'en remercie. Merci aussi d'avoir été présente et à l'écoute, tout en me laissant la liberté dont j'avais besoin. Merci de m'avoir soutenue dans les moments plus difficiles de la thèse. Finalement, je garde un bon souvenir de Stockholm !

Je tiens aussi à remercier Anouck Amestoy, Jean-Julien Aucouturier, Anne Caclin, Thomas Desmidt et Sandrine Ferré d'avoir accepté de faire partie de ce jury de thèse ainsi que d'évaluer ce travail. Je remercie tout particulièrement Jean-Julien pour son aide et ses conseils tout au long de la thèse, qui m'ont permis de mieux appréhender les méthodes et les concepts que ce travail impliquait.

Merci à toute l'équipe de recherche « Autisme et Troubles du Neurodéveloppement » ; ces quatre années ont été enrichissantes tant sur le plan scientifique que personnel grâce à vous tous-tes. Merci à Frédérique Bonnet-Brilhault pour son accueil dans l'équipe et pour ses mots encourageants quand il le fallait. Merci à Marianne et Frédéric, toujours là quand on a besoin d'eux avec un mot gentil ou un avis scientifique. Je tenais aussi à remercier Claire et Nadia qui m'ont fait découvrir le monde de la pupille et qui ont toujours eu un regard avisé et bienveillant sur mon travail. Enfin, merci aux linguistes de l'équipe de m'avoir fait découvrir un nouveau pan de la recherche. Vous savoir dans l'équipe est un privilège nous permettant de prendre du recul sur ce que nous faisons au laboratoire. Louise, tu es arrivée plus récemment mais merci pour nos discussions plus techniques, ton soutien et tes encouragements.

Merci à la Fondation Européenne pour l'Avancement des Neurosciences d'avoir financé cette dernière année de thèse, me permettant de finaliser ce travail dans les meilleures conditions.

Merci à toutes les infirmières que j'ai eu la chance de côtoyer aux explorations fonctionnelles Hélène, Céline, Chloé, Marie, Lucille et Juliette pour votre gentillesse à l'égard des doctorant·e·s et de prendre toujours cinq minutes pour discuter avec nous et nous changer les idées.

Je remercie les équipes du Département Universitaire Psychiatrie de l'Enfant et l'Adolescent et du Centre Ressources Autisme pour leur investissement et leur aide inestimable dans ce travail. Merci à Bénédicte, Clotilde, Aude-Marie, Hélène, Élise, Romuald, Aurore, Céline, Océane et Shasha pour votre accueil et votre gentillesse dans le laboratoire et le service. Merci Rémy de répondre toujours présent quand nous avons

## Remerciements

besoin d'aide. Sarah et Virginie, merci pour votre patience quand je venais vous demander plusieurs fois par jour au secrétariat du CRA « est-ce que le patient a accepté la recherche ? », je vous souhaite le meilleur pour la suite ici ou ailleurs ! Je remercie tout particulièrement le docteur Emmanuelle Houy-Durand pour l'aide dans le recrutement des participants autistes.

Un grand merci à toutes les personnes autistes et neurotypiques qui ont accepté de participer aux différentes études menées dans ce travail. Sans vous ce travail n'aurait jamais vu le jour. Je remercie aussi tous les parents et les enfants autistes que j'ai eu la chance de rencontrer avec Zoé, vous donnez un sens à tout ce que nous faisons ici.

Merci à Pablo et à Aynaz, rencontrés plus récemment mais qui m'ont tous les deux permis d'avoir une nouvelle vision de ce que je faisais. Merci pour vos conseils et votre aide dans le traitement et l'analyse des données.

*Cositas*, merci d'avoir été là tout au long de cette thèse. Rosa, merci de m'avoir aidée tant de fois en statistiques et sur le traitement de données. Te savoir dans le laboratoire était la garantie de passer un bon moment. Camille, merci de m'avoir aussi bien accueillie dans le bureau quand le masque et le COVID nous empêchaient de nous intégrer facilement. Loyal et Lisa, merci d'avoir été là quand la thèse était parfois plus difficile avec toujours un mot réconfortant. Marco, merci d'avoir été là les jeudis pour nous rappeler que la thèse était un beau projet mais que ça ne devait pas être toute notre vie ! Et enfin, Marta, je suis heureuse d'avoir commencé la thèse avec toi, tu as été un super binôme de début de thèse. Tu peux être fière du travail que tu as accompli et te voilà maintenant maîtresse de conférences.

Merci aux doctorant·e·s arrivés par la suite : Marine, Raphaël, Vivien, Mohammad, Zoé et Camille, je vous souhaite le meilleur pour la suite. Camille et Zoé, je suis heureuse que vous ayez rejoint le laboratoire plus récemment. Zoé, ma première stagiaire, mais aujourd'hui ma collègue de thèse. Merci pour tous ces moments thé/café à parler de science (mais pas que). J'espère que la suite se passera bien pour toi au labo, je te laisse tous mes scripts et je ne serai jamais loin si tu as besoin ! Camille, merci pour ta joie de vivre et la persévérance dont tu fais preuve, ne panique pas, tu vas y arriver, la thèse va bien se passer. Merci à toutes les deux pour le temps passé à relire cette thèse à la virgule près et pour vos encouragements dans la dernière phase de rédaction. Merci de m'avoir permis de passer un été de rédaction serein, grâce à vous je ne me suis jamais sentie seule au labo, je vous revaudrai ça !

Maintenant plus personnellement, je tiens à remercier toutes les personnes qui ont été là pour moi de près ou de loin. Sarah et Agathe, le lycée est bien loin maintenant, mais je vous remercie d'avoir été présentes tout au long de ces quatre années (et bien avant). Ces quelques week-ends à vos côtés ont été de très belles parenthèses. Siham et Jenna, je ne pensais pas rencontrer de nouvelles amies si importantes pendant ce master

## Remerciements

toulousain. Merci d'avoir toujours pris de mes nouvelles et de m'avoir encouragée pendant cette thèse. Siham, merci pour ces vacances passées chez toi et ce nouvel an que je n'oublierai pas. Je serai là pour vous soutenir quand le moment sera venu pour vous !

Merci à ma famille d'avoir été là et d'avoir toujours cru en moi. Je suis fière d'appartenir à ce clan familial. Je remercie mes grands-parents, je sais que c'est une chance inouïe de vous avoir aussi présents et investis dans tout ce que j'entreprends. Je suis heureuse et reconnaissante de vous avoir à mes côtés. Sarah, merci d'être ma sœur et d'être toujours un soutien inébranlable quand le besoin s'en fait sentir. Merci d'avoir apporté dans notre famille Antonin et ta nouvelle petite famille ! Grâce à vous deux j'ai passé deux étés merveilleux qui m'ont permis de m'échapper quelques semaines de la thèse. Élise et Alma, merci d'être là, je suis fière d'être maintenant « Docteur Tata ».

Enfin, merci à mes parents d'être qui vous êtes. Vous m'avez appris à prendre confiance en moi, à croire en mes capacités tout en développant un sens critique et une ouverture d'esprit essentiels. Merci d'avoir toujours cru en moi quand je n'y arrivais plus moi-même et de m'avoir redonné envie quand c'était plus difficile. Papa, merci de m'avoir donné ta rigueur scientifique et ta sensibilité. Maman, merci de m'avoir appris à cultiver ma créativité et mon envie de découvrir. Je vous aime.

Finalement, je remercie toutes les personnes qui de près ou de loin m'ont accompagnée dans tout le cheminement de cette thèse. Ça n'a pas toujours été simple, je n'ai pas toujours pris de nouvelles aussi régulièrement que je l'aurais voulu, mais vous avez toujours répondu présents quand le besoin s'en faisait ressentir.

Ce travail de thèse est l'aboutissement d'un travail personnel mais qui n'aurait jamais été possible sans l'aide de toutes les personnes qui m'entourent. Merci !



## Résumé

Bien que le diagnostic du trouble du spectre de l'autisme (TSA) nécessite la présence conjointe de troubles de la communication sociale et de comportements répétitifs et restreints, peu d'études intègrent à la fois les domaines socio-émotionnel et perceptif. Il reste à déterminer si les difficultés rencontrées par les personnes autistes sont liées à des déficits émotionnels spécifiques, à des particularités sensorielles plus marquées pour les stimuli sociaux, ou à la combinaison des deux. La perception et le traitement des émotions vocales constituent un bon modèle pour explorer chacune des étapes de la boucle *Perception – Représentation – Action* dans l'autisme. L'objectif de cette thèse était de caractériser les différentes étapes de cette boucle chez des adultes autistes, en utilisant diverses méthodes d'exploration, afin d'identifier quelles étapes seraient atypiques et contribueraient à l'absence de réponse aux émotions observée. Plusieurs questions de recherche ont ainsi été posées afin de caractériser ces étapes : i) Comment caractériser l'adaptation neurale aux émotions vocales dans l'autisme ? ii) Quelles sont les représentations perceptuelles des émotions vocales dans l'autisme ? iii) Une résonance motrice aux émotions vocales est-elle observée dans l'autisme et quel est le lien entre cette résonance et les représentations ?

Les deux premières études de ce travail ont permis de caractériser l'adaptation neurale à la voix et aux émotions vocales chez les personnes autistes et les adultes neurotypiques. Ces études ont mis en évidence une adaptation neurale atypique dans l'autisme touchant principalement les sons non-vocaux. Nous avons ensuite montré que les représentations mentales des émotions vocales étaient similaires pour une grande partie de l'échantillon d'adultes autistes, en lien avec les mesures d'empathie. Enfin, l'étude de la résonance motrice et de la réaction pupillaire a révélé une contagion émotionnelle chez les adultes neurotypiques mais une absence totale de réponse motrice chez les adultes autistes. Ainsi, des représentations similaires des émotions vocales et une bonne reconnaissance de celles-ci ne suffisent pas à déclencher une résonance motrice chez les personnes autistes.

Nos résultats ont mis en évidence des différences, principalement dans la dernière étape de la boucle, alors que la perception et les représentations des émotions vocales semblent relativement préservées dans l'autisme. Les résultats sont discutés au regard des théories suggérant des difficultés d'incarnation des émotions dans l'autisme.

Cette approche intégrative originale a permis de dissocier les processus perceptifs et émotionnels, de déterminer à quel niveau les processus physiopathologiques peuvent intervenir, et ainsi de proposer un nouvel éclairage mécanistique sur les difficultés socio-émotionnelles dans l'autisme.



## **Abstract**

Both socio-emotional impairments and repetitive behaviors/restricted interests are required to establish a diagnosis of autism spectrum disorder (ASD). Yet, little research has attempted to integrate socio-emotional and perceptive domains to explain ASD impairments. It remains to be determined whether difficulties in ASD are related to emotion-specific deficits, to sensory peculiarities which are even more marked for social than non-social stimuli, or a combination of both. Vocal emotion perception and processing serves as a good model for exploring each stage of the *Perception – Representation – Action* loop in autism. The objective of this thesis was to characterize the different phases of this loop in autistic adults using various exploratory methods in order to identify which stages might be atypical and contribute to the observed lack of response to emotions. Several research questions were asked to achieve our goal: i) How can neural adaptation to vocal emotions be characterized in autism? ii) What are the perceptual representations of vocal emotions in autism? iii) Is motor resonance to vocal emotions possible in autism, and, if so, what is the relationship between this resonance and the representations?

Atypicalities appear to be present in the very first stage of the loop. The first two studies in this work characterized neural adaptation to voice and vocal emotions in autistic and neurotypical adults. These studies revealed an atypical neural adaptation in autism mainly for non-vocal sounds. Then, we showed that the mental representations of vocal emotions were similar for most of the sample of autistic adults and that these representations correlated with empathetic measures. Finally, the study of motor resonance and pupillary reactivity revealed the presence of emotional contagion in neurotypical adults but a total lack of motor response in autistic adults. Similar representations of vocal emotions neurotypicals and good recognition of these emotions were not sufficient to trigger motor resonance in autism.

Our results highlighted differences primarily in the last stage of the loop, while the perception and representations of vocal emotions appear relatively preserved in autism. The findings are discussed in light of theories suggesting difficulties with emotional embodiment in autism.

This original integrative approach enabled the dissociation of perceptual and emotional processes and to determine at what level of pathophysiological mechanisms may intervene, thus offering new mechanistic insights into socio-emotional difficulties in autism.



## Table des matières

<b>Remerciements</b> .....	<b>III</b>
<b>Résumé</b> .....	<b>VII</b>
<b>Abstract</b> .....	<b>IX</b>
<b>Table des matières</b> .....	<b>XI</b>
<b>Liste des figures</b> .....	<b>XV</b>
<b>Liste des tableaux</b> .....	<b>XVI</b>
<b>Liste des abréviations</b> .....	<b>XVII</b>
<b>Avant-propos</b> .....	<b>XIX</b>
<b>Cadre théorique</b> .....	<b>XXI</b>
<b>1. Chapitre 1 : L'autisme</b> .....	<b>1</b>
<b>1.1. Historique</b> .....	<b>1</b>
<b>1.2. Caractéristiques de la pathologie</b> .....	<b>1</b>
1.2.1. Diagnostic et symptomatologie .....	1
1.2.1. Intérêts spécifiques et comportements répétitifs .....	4
1.2.2. Communication et interactions sociales : traitement des émotions.....	4
<b>2. Chapitre 2 : Perception et traitement de la prosodie émotionnelle</b> .....	<b>9</b>
<b>2.1. Prosodie</b> .....	<b>9</b>
2.1.1. Définition .....	9
<b>2.2. Prosodie émotionnelle</b> .....	<b>9</b>
2.2.1. Définition et fonction de la prosodie émotionnelle .....	9
2.2.2. Indices acoustiques de la prosodie émotionnelle .....	10
2.2.3. Reconnaissance de la prosodie émotionnelle .....	11
2.2.4. Cas particulier du sourire vocal .....	12
<b>2.3. Traitement cérébral de la voix et des émotions vocales</b> .....	<b>15</b>
2.3.1. Modèle de perception et de traitement de la voix et des émotions vocales .....	16
2.3.1.1. Réponse à la voix .....	16
2.3.1.2. Réponse cérébrale aux émotions vocales.....	17
<b>3. Chapitre 3 : Perception et traitement de la prosodie émotionnelle dans l'autisme ....</b>	<b>21</b>
<b>3.1. Prosodie émotionnelle dans l'autisme</b> .....	<b>21</b>
3.1.1. Atypies de production .....	21
3.1.2. Atypies de perception et de reconnaissance des émotions vocales .....	22
3.1.1. Traitement cérébral de la voix et de la prosodie émotionnelle dans l'autisme .....	24
<b>3.1. Adaptation et encodage de la régularité des émotions vocales.</b> .....	<b>27</b>
<b>4. Chapitre 4 : Réponse aux émotions vocales et contagion émotionnelle</b> .....	<b>63</b>
<b>4.1. Contagion émotionnelle chez le sujet neurotypique</b> .....	<b>63</b>
4.1.1. Résonance motrice émotionnelle.....	64
4.1.1.1. Modèles historiques des émotions.....	65
4.1.1.2. Les neurones miroirs.....	66

## Table des matières

4.1.1.3.	Cognition incarnée .....	67
4.1.1.4.	Modèle neurocognitif de la contagion émotionnelle (NMEC) .....	68
4.1.2.	Contagion autonome et pupillométrie .....	69
4.1.3.	La contagion émotionnelle auditive .....	70
4.1.4.	Cas particulier du sourire .....	71
<b>4.2.</b>	<b>Contagion émotionnelle dans l'autisme .....</b>	<b>72</b>
4.2.1.	Résonance motrice et contagion émotionnelle .....	72
4.2.1.1.	Expressions faciales .....	73
4.2.1.2.	Contagion pupillaire .....	74
<b>5.</b>	<b>Hypothèses et objectifs .....</b>	<b>77</b>
	<b>Methodologie .....</b>	<b>81</b>
<b>1.</b>	<b>Populations .....</b>	<b>83</b>
1.1.1.	Cadre éthique .....	83
1.1.2.	Populations recrutées .....	83
<b>2.</b>	<b>Matériel et méthodes .....</b>	<b>84</b>
<b>2.1.</b>	<b>Électroencéphalographie .....</b>	<b>84</b>
2.1.1.	Principe de l'EEG .....	84
2.1.2.	Séquence de stimulation .....	86
2.1.3.	Acquisition et traitement du signal .....	87
<b>2.2.</b>	<b>Corrélation inverse .....</b>	<b>88</b>
2.2.1.	Principe de la corrélation inverse .....	88
2.2.2.	Bruit interne .....	91
2.2.3.	Création des sons et décours expérimental .....	92
<b>2.3.</b>	<b>Électromyographie faciale et pupillométrie .....</b>	<b>95</b>
2.3.1.	Principes de l'EMG et de la pupillométrie .....	95
2.3.1.1.	Électromyographie faciale (fEMG) .....	95
2.3.1.2.	Pupillométrie .....	96
2.3.2.	Séquence de stimulation .....	98
2.3.3.	Acquisition et traitement du signal .....	100
2.3.3.1.	Électromyographie faciale (fEMG) .....	100
2.3.3.2.	Pupillométrie .....	100
	<b>Résultats expérimentaux .....</b>	<b>103</b>
<b>1.</b>	<b>Adaptation neurale à la prosodie émotionnelle .....</b>	<b>105</b>
1.1.	Avant-propos .....	105
1.2.	Étude n°1 : Adaptation neurale à des sons vocaux et non-vocaux chez l'adulte neurotypique .....	106
1.3.	Étude n°2 : Effet de la prosodie émotionnelle sur l'adaptation neurale à des sons vocaux et non-vocaux dans l'autisme .....	120
1.4.	Synthèse .....	131
<b>2.</b>	<b>Représentation mentale du sourire vocal .....</b>	<b>133</b>
2.1.	Avant-propos .....	133
2.2.	Étude n°3 : Représentation perceptive du sourire vocal dans l'autisme .....	134

## Table des matières

2.3. Synthèse .....	161
<b>3. Résonance motrice et contagion émotionnelle au sourire vocal .....</b>	<b>163</b>
3.1. Avant-propos .....	163
3.2. Étude n°4 : Résonance motrice et autonome au sourire vocal chez l'adulte neurotypique 164	
3.3. Étude n°5 : Résonance motrice au sourire vocal chez l'adulte autiste .....	178
3.4. Synthèse .....	205
<b>Discussion générale.....</b>	<b>207</b>
<b>1. Discussion.....</b>	<b>209</b>
1.1. Synthèse des résultats .....	209
1.2. Réponse contrôlée versus réponse automatique : stratégies de compensation ? .....	212
1.3. Traitement de la prosodie émotionnelle dans l'autisme : cas du cerveau Bayésien ? ...	214
1.4. L'autisme : une cognition incarnée atypique ?.....	216
<b>2. Limitations.....</b>	<b>219</b>
2.1. Limites méthodologiques relatives aux protocoles.....	219
2.2. Taille de l'échantillon et reflet de l'hétérogénéité clinique .....	220
<b>3. Perspectives.....</b>	<b>222</b>
3.1. Perspectives de recherche .....	222
3.2. Applications thérapeutiques .....	225
<b>4. Conclusion .....</b>	<b>227</b>
<b>Références .....</b>	<b>229</b>
<b>Annexes .....</b>	<b>269</b>
<b>Annexe 1 : Récapitulatif des participants recrutés dans les différentes études .....</b>	<b>271</b>
<b>Annexe 2 : Étude préliminaire de l'adaptation neurale chez l'adulte neurotypique .....</b>	<b>273</b>
<b>Annexe 3 : Étude préliminaire de la résonance motrice chez l'adulte neurotypique.....</b>	<b>279</b>
<b>Annexe 4 : Données supplémentaires de l'étude n°3.....</b>	<b>283</b>
<b>Annexe 5 : Données supplémentaires de l'étude n°4.....</b>	<b>287</b>
<b>Annexe 5 : Liste des publications .....</b>	<b>295</b>



## Liste des figures

<b>Figure 1 : Sourire de Duchenne</b> .....	13
<b>Figure 2 : Modulation prosodique et relation avec les modifications anatomiques associées (Arias, Rachman, et al., 2021)</b> .....	14
<b>Figure 3: Modèle du traitement de la voix adapté de Belin et al. (2011)</b> .....	17
<b>Figure 4 : Séquence d'activation cérébrale en réponse à la prosodie émotionnelle</b> .....	18
<b>Figure 5 : Modèle de la contagion émotionnelle adapté de Prochazkova &amp; Kret (2017).</b> ...	69
<b>Figure 6 : Résonance motrice au sourire vocal à partir de Arias et al. (2018)</b> .....	72
<b>Figure 7 : Représentation des différentes études présentées dans la thèse</b> .....	79
<b>Figure 8: Caractéristiques des sons et de la séquence utilisés pour l'étude en EEG de l'adaptation neurale</b> .....	87
<b>Figure 9 : Positionnement des électrodes d'EEG dans le système 10-20</b> .....	88
<b>Figure 10 : Synthèse de la méthode d'acquisition et de pré-traitement des données EEG</b>	88
<b>Figure 11 : Exemple de paradigme de corrélation inverse en visuel</b> .....	89
<b>Figure 12 : Modèle du sourire vocal à partir de la corrélation inverse Ponsot et al. (2018)</b>	91
<b>Figure 13 : Résultats analyse convergence RevCor à partir des données de Ponsot et al. (2018)</b> .....	94
<b>Figure 14 : Synthèse de la méthode d'acquisition et de traitement des données de corrélation inverse</b> .....	94
<b>Figure 15 : Innervation et anatomie des muscles de la face</b> .....	96
<b>Figure 16: Schéma de l'innervation sympathique et parasympathique de la pupille</b> .....	97
<b>Figure 17: Décours de la tâche dans le protocole de mesure de pupillométrie et fEMG et échelle de jugement du degré de sourire projetée à l'écran</b> .....	99
<b>Figure 18 : Synthèse de la méthode d'acquisition et de pré-traitement des données de pupillométrie et fEMG</b> .....	101
<b>Figure 19 : Schéma de synthèse des études n°1 et n°2 à propos de l'adaptation neurale à la prosodie émotionnelle</b> .....	131
<b>Figure 20 : Schéma de synthèse de l'étude n°3 à propos de la représentation perceptive du sourire vocal dans l'autisme</b> .....	161

## Liste des tableaux

<b>Figure 21: Schéma de synthèse des études n°4 et 5 à propos de la résonance motrice et autonome chez l'adulte neurotypique et autiste .....</b>	<b>205</b>
<b>Figure 22 : Récapitulatif des résultats dans l'autisme des différentes études et conséquences prévisibles sur le plan clinique .....</b>	<b>210</b>
<b>Figure 23 : Profils des participants autistes ayant participé à l'ensemble des études ....</b>	<b>212</b>
<b>Figure 24 : Cerveau Bayésien et prédiction dans l'autisme (Haker et al., 2016) .....</b>	<b>216</b>
<b>Figure 25 : Protocole et résultats envisagés pour tester conjointement la perception et la production de la prosodie émotionnelle .....</b>	<b>224</b>

## Liste des tableaux

<b>Tableau 1 : Modulations acoustiques caractéristiques des émotions primaires exprimées à travers la voix.....</b>	<b>10</b>
<b>Tableau 2 : Récapitulatif des participants inclus dans les protocoles définitifs présentés dans ce travail de thèse .....</b>	<b>84</b>
<b>Tableau 3 : Récapitulatif et caractéristiques des phrases utilisées dans études n°4 et 5.</b>	<b>98</b>

## Liste des abréviations

**CRA** : Centre de Ressources Autisme

**CS** : *Corrugator supercilii*

**DSM** : *Diagnostic and Statistical Manual of mental disorders*, Manuel Diagnostique et Statistique des troubles mentaux

**EEG** : électroencéphalogramme

**F0** : fréquence fondamentale

**F1/F2/F3/F4** : fréquence des formants 1 à 4

**fEMG** : électromyographie faciale

**FTPV** : *Fronto-Temporal Positivity to Voice*, positivité fronto-temporale à la voix

**IRMf** : Imagerie par Résonance Magnétique fonctionnelle

**LC-NE** : Locus-Coeuruleus NorépinÉpinephrine

**MMN** : *MisMatch Negativity*

**MNS** : *Mirror Neuron System*, système des neurones miroirs

**NMEC** : *Neurocognitive Model of Emotional Contagion*, modèle neurocognitive de la contagion émotionnelle

**RE** : *Repetition Enhancement*

**RevCor** : *Reverse Correlation*, corrélation inverse

**RP** : *Repetition Positivity*

**RS** : *Repetition Suppression*

**SEPIA** : *Sensory and Emotional Processing In Autism spectrum disorders*

**SNA** : Système Nerveux Autonome

**TDAH** : Trouble du Déficit de l'Attention avec ou sans Hyperactivité

**TVA** : *Temporal Voice Area*, aire temporelle à la voix

**TSA** : Trouble du Spectre de l'Autisme

**ZM** : *Zygomaticus major*



## Avant-propos

Un sourire ne se voit pas au téléphone, mais il s'entend. Ce simple geste vocal peut transformer une conversation, influencer les émotions de l'interlocuteur et même modifier l'issue d'un échange. Pour provoquer une réponse congruente chez la personne qui reçoit ce sourire, plusieurs étapes sont nécessaires, définissant ainsi la boucle *Perception – Représentation – Action*. Traditionnellement décrite dans le domaine visuel dans la théorie de la vision "écologique" de Gibson (1986), cette boucle peut également s'appliquer à la réponse aux émotions vocales.

Dans l'autisme, l'absence de réponse au sourire social est un des principaux signes d'alerte observés. Par ailleurs, un autre élément caractéristique de l'autisme est une prosodie atypique. En effet, parfois décrite comme monotone ou théâtrale, il s'agit d'un des éléments évalués lors du diagnostic d'autisme. Ces deux éléments amènent à se poser la question de la réponse au sourire vocal dans l'autisme. Quelles sont la ou les étapes atypiques qui conduisent à une réponse différente aux émotions vocales dans l'autisme ? L'objectif de cette thèse est de caractériser les différentes étapes de la boucle *Perception – Représentation – Action* en réponse au sourire vocal chez des adultes autistes<sup>1</sup>. Identifier les atypies dans ces étapes permettrait d'en cibler une en particulier pour améliorer la prise en charge et la remédiation de personnes autistes afin de faciliter les interactions sociales.

La première partie du manuscrit est dédiée à la description du cadre théorique dans lequel s'inscrivent les différentes études menées au cours de la thèse. Un premier chapitre introductif présente brièvement l'historique et la symptomatologie de l'autisme. Le deuxième chapitre se concentre sur la littérature relative à la perception et au traitement de la prosodie émotionnelle au cours développement typique. Le troisième chapitre décrit ce traitement dans l'autisme, avec un accent mis sur l'adaptation à la régularité dans certains troubles du neurodéveloppement, incluant une revue de la littérature publiée. Enfin, le quatrième chapitre de l'introduction, explore la réponse aux émotions vocales et la contagion émotionnelle dans

---

<sup>1</sup> Dans cette thèse j'ai choisi de qualifier les participants de personnes autistes et neurotypiques (ou au développement typique) d'après l'étude de Geelhand (2023) sur le principe de la primeur de l'identité plutôt que du diagnostic.

## Avant-propos

l'autisme et au cours du développement typique. Les objectifs de ce travail de thèse et les hypothèses associées sont également exposés à la fin de cette introduction.

La partie expérimentale qui suit l'introduction, est structurée en plusieurs sections. Elle commence par une description des populations recrutées et des méthodes d'exploration utilisées. Chaque étude réalisée est ensuite présentée avec une description en français précédant chacun des articles issus de ce travail. Enfin, une synthèse des résultats et des études pour chacune des grandes étapes de traitement est fournie.

Les premiers résultats expérimentaux de la thèse concernent l'adaptation neurale à la prosodie émotionnelle chez l'adulte neurotypique (étude n°1) puis dans l'autisme (étude n°2), étudiée via les réponses neurophysiologiques à des sons vocaux et non-vocaux. La deuxième partie présente l'étude n°3, qui caractérise les représentations mentales du sourire vocal chez les adultes autistes à l'aide d'une tâche de corrélation inverse. Enfin, la dernière partie concerne la résonance et la contagion émotionnelle aux émotions vocales chez l'adulte neurotypique (étude n°4) puis dans l'autisme (étude n°5), évaluées à l'aide de la mesure de l'activité des muscles faciaux et des variations du diamètre pupillaire.

Enfin, la dernière section de cette thèse propose une discussion générale des cinq études réalisées, en mettant en perspective les résultats obtenus pour caractériser les profils de perception de la prosodie émotionnelle dans l'autisme. Les limitations de ces recherches ainsi que les perspectives expérimentales et cliniques pour l'amélioration de la prise en charge des personnes autistes y sont également abordées.

## **Cadre théorique**

---



## 1. Chapitre 1 : L'autisme

### 1.1. Historique

La première description de l'autisme dans un article scientifique date de 1926 et a été faite par une pédopsychiatre russe, Grunia Ssucharewa (Ssucharewa, 1926). Elle y décrit six garçons avec une « attitude autistique » présentant des particularités cognitives et émotionnelles avec une forte inflexibilité et des difficultés à s'adapter. Malheureusement ses écrits n'ont été découverts que tardivement (Andronikof & Fontan, 2016) et ce sont les descriptions de Léo Kanner et Hans Asperger, plus tardives, qui seront considérées comme princeps (Asperger, 1944; Kanner, 1943). Dans chacune de ces descriptions historiques, on observe chez les enfants une symptomatologie similaire, caractérisée par des difficultés dans les interactions sociales, une rigidité, des intérêts spécifiques et répétitifs, ainsi que des particularités dans le langage et les perceptions sensorielles.

Dans la description de Kanner en 1943, les concepts de la dyade symptomatologique de l'autisme sont déjà évoqués, sous les termes de *aloneness* pour la première dimension et de *sameness* pour la seconde. La dimension *aloneness* selon Kanner correspond au retrait sur soi et à l'absence de communication et d'interaction avec les autres. La seconde dimension, *sameness*, fait alors référence à la recherche d'immuabilité et à la résistance au changement observées dans l'autisme.

### 1.2. Caractéristiques de la pathologie

#### 1.2.1. Diagnostic et symptomatologie

Le Trouble du Spectre de l'Autisme (TSA) est un trouble du neurodéveloppement diagnostiqué en présence de difficultés de la communication et des interactions sociales, ainsi que de comportements et intérêts spécifiques et répétitifs (American Psychiatric Association, 2013). L'autisme serait la résultante de facteurs environnementaux et génétiques, dans 64 à 91% des cas des causes génétiques ont été évoquées (Tick et al., 2016). Bien que des causes biologiques aient été identifiées dans certains cas, aucune n'est présente dans tous les cas d'autisme. Le diagnostic se base toujours sur l'évaluation comportementale des personnes par plusieurs professionnels de santé, en observant les capacités langagières et de

## Chapitre 1 : L'autisme

communication, le développement sensorimoteur et le fonctionnement cognitif (Geschwind, 2011; Hyman et al., 2020; Lord et al., 2018).

Les critères diagnostiques de TSA dans le DSM-5 (Manuel Diagnostique et Statistique des Troubles Mentaux 5<sup>ème</sup> édition, American Psychiatric Association, 2013) représentent deux dimensions symptomatologiques, on parle de « dyade autistique » : (A) déficits de la communication et des interactions sociales et (B) le caractère spécifique et répétitif des comportements et des intérêts. D'après le DSM-5 pour poser un diagnostic d'autisme il faut retrouver 3 symptômes parmi la catégorie A et 2 symptômes sur 4 dans la catégorie B (voir encadré n°1). L'ADOS (*Autism Diagnostic Observation Schedule*) est un des outils utilisés en clinique pour faciliter le diagnostic d'autisme (Lord et al., 2000). Cet entretien semi-structuré est adapté à l'âge et au niveau du patient, avec quatre modules disponibles pour couvrir de la petite enfance à l'âge adulte. Il se compose d'une série d'activités et de scénarios sociaux conçus pour observer les comportements de communication, d'interaction sociale, de jeu et d'imagination de l'individu. Ce test permet d'évaluer divers aspects du développement social et communicatif, fournissant des informations essentielles pour le diagnostic et l'évaluation de l'autisme. Au cours de l'ADOS, la personne chargée de l'évaluation analysera également la prosodie (émotionnelle) du patient.

Le TSA représente entre 0,9 % et 1,2 % des naissances (Delobel-Ayoub et al., 2020), soit environ 7 500 bébés chaque année en France. La Haute Autorité de Santé estime, donc, qu'environ 100 000 jeunes de moins de 20 ans et près de 600 000 adultes sont autistes en France à ce jour (source : <https://handicap.gouv.fr/la-strategie-nationale-autisme-et-troubles-du-neurodeveloppement-2018-2022>). Par ailleurs, ces chiffres sont revus à la hausse depuis plusieurs années avec une prévalence dans certaines études de 1/44 (Maenner, 2023). Ces différences dans la prévalence sont la conséquence d'un élargissement des critères diagnostiques (spectre) mais aussi d'une plus grande connaissance de l'autisme et par conséquent, d'un meilleur accès au diagnostic. Une forme de « sur-diagnostic » a ainsi été retrouvée aux États-Unis (Fombonne, 2018) mais aussi en France (Delobel-Ayoub et al., 2020)

Dans l'autisme, il existe une différence dans le nombre de personnes diagnostiquées chez les hommes et chez les femmes avec un sex-ratio de 1 femme pour 4 hommes (Maenner, 2021). Ces différences de prévalence selon le genre sont discutées comme étant une conséquence

d'outils cliniques et diagnostiques non-adaptés pour les phénotypes féminins d'autisme (Lockwood Estrin et al., 2021). De plus, une compensation et un camouflage des comportements et des symptômes plus importants seraient retrouvés chez les femmes que chez les hommes autistes conduisant à un sous-diagnostic (Lai et al., 2017; Livingston & Happé, 2017). Une récente étude norvégienne, menée sur un vaste échantillon de deux millions d'enfants et d'adultes, autistes et non autistes, révèle un ratio hommes/femmes dans l'autisme de 3,67/1 chez les enfants et de 2,57/1 chez les adultes. Ces résultats suggèrent que les filles reçoivent un diagnostic d'autisme à un âge plus tardif que les garçons (Posserud et al., 2021).

**Critères d'un diagnostic d'autisme dans le DSM-5  
American Psychiatric Association (2013)**

- A) Déficiences persistantes de la communication et des interactions sociales observés dans des contextes variés
  - Déficiences de la réciprocité sociale ou émotionnelle
  - Déficiences des comportements de communication non-verbaux utilisés au cours des interactions sociales
  - Déficiences du développement, du maintien et de la compréhension des relations
- B) Caractère restreint et répétitif des comportements, des intérêts ou des activités
  - Caractère stéréotypé ou répétitif des mouvements, de l'utilisation d'objet ou du langage
  - Intolérance au changement, adhésion indéfectible à des routines ou à des modes comportementaux verbaux ou non-verbaux ritualisés
  - Intérêts extrêmement restreints et fixes, anormaux soit dans leur intensité, soit dans leur but
  - Hypo ou hyperréactivité aux stimulations sensorielles ou intérêt inhabituel pour les aspects sensoriels de l'environnement
- C) Les symptômes doivent être présents dès les étapes précoces du développement
- D) Les symptômes occasionnent un retentissement cliniquement significatif en termes de fonctionnement actuel, social, scolaire (professionnels ou dans d'autres domaines importants).
- E) Ces troubles ne sont pas mieux expliqués par un handicap intellectuel (trouble du développement intellectuel) ou un retard global du développement. La déficience intellectuelle et le trouble du spectre de l'autisme sont fréquemment associés. Pour permettre un diagnostic de comorbidité entre un trouble du spectre de l'autisme et un handicap intellectuel, l'altération de la communication sociale doit être supérieure à ce qui serait attendu pour le niveau de développement général.

**Encadré 1 : critères diagnostiques d'un trouble du spectre de l'autisme d'après le DSM-5**

Depuis le DSM-5, l'autisme est considéré comme un trouble au « long court », qui se maintient tout au long de la vie. Cependant peu d'études s'intéressent spécifiquement à la cognition et au fonctionnement des adultes autistes, en comparaison au nombre d'études menées chez l'enfant (Mason et al., 2021). Par ailleurs, une meilleure compréhension de l'autisme à l'âge adulte est importante, car même si certains comportements peuvent être compensés à l'âge adulte, l'appréhension des phénomènes sous-jacents est essentielle dans la prise en charge et l'accompagnement des adultes autistes.

### **1.2.1. Intérêts spécifiques et comportements répétitifs**

Un des deux volets de la dyade symptomatologique concerne le caractère spécifique et répétitif des comportements, des intérêts ou des activités avec en particulier un besoin et une recherche d'immuabilité. En effet, dans l'autisme on retrouve des comportements répétés qui permettent une forme d'autostimulation (comme le *flapping*<sup>2</sup>) et dans l'objectif de préserver au maximum un environnement stable au cours du temps (Leekam et al., 2011; Turner, 1999). Ces comportements sont aussi décrits comme répondant au besoin et à la nécessité d'installation de routines chez les enfants et les adultes autistes (Crespo et al., 2013; Rodger & Umaibalan, 2011). En effet, la mise en place de routines et de rituels permet d'organiser et de planifier les journées, les rendant moins sujettes aux imprévus, difficiles à gérer par les personnes autistes. Dans l'autisme, cette recherche d'immuabilité viendrait compenser l'intolérance au changement et les difficultés à traiter les modifications non planifiées qui rendent le quotidien difficile (Gabriels et al., 2005; Murphy et al., 2005).

### **1.2.2. Communication et interactions sociales : traitement des émotions**

Les troubles de la communication et des interactions sociales constituent des caractéristiques centrales de l'autisme. Ces défis se manifestent par des difficultés à établir et à maintenir des relations sociales, à comprendre et à utiliser les codes de communication non verbale, et à ajuster ses comportements en fonction des contextes sociaux (American Psychiatric Association, 2022; Tager-Flusberg et al., 2005). Les personnes autistes rencontrent des difficultés d'ajustement émotionnel, se traduisant par une reconnaissance et une expression

---

<sup>2</sup> Le *flapping* est un comportement stéréotypé fréquemment associé aux TSA, utilisé par les individus pour gérer les stimuli sensoriels et les émotions (Moore & Goodson, 2003; Ozonoff et al., 2005)

des émotions moins adaptées. Ces obstacles impactent non seulement, les interactions sociales, mais aussi la compréhension et la gestion des émotions, contribuant à un traitement émotionnel atypique (Frith, 2003; Uljarevic & Hamilton, 2013).

Dans l'autisme, des atypies dans la réponse aux émotions sont observées, caractérisées par un manque de contagion émotionnelle et de résonance motrice (Helt et al., 2020; McIntosh et al., 2006), ainsi que par une absence de réaction au sourire social (Beall et al., 2008). Ce dernier point est particulièrement significatif, car il constitue l'un des premiers signes d'alerte chez le jeune enfant. D'ailleurs, toujours dans les descriptions princeps de l'autisme, ces signes étaient évoqués comme « mimiques parfois paradoxales : rit avec une expression faciale triste » (Andronikof & Fontan, 2016). On parle alors d'amimie dans l'autisme, une absence d'expression faciale, et notamment en réponse aux émotions d'autrui (McIntosh et al., 2006).

Les émotions sont difficilement reconnues dans l'autisme, notamment dans la modalité visuelle (Harms et al., 2010; Uljarevic & Hamilton, 2013). Toutefois la question de la reconnaissance des émotions vocales a également été largement posée (Golan et al., 2007; Hubbard et al., 2017). Les émotions produites et exprimées par les personnes autistes sont aussi plus difficilement reconnues par des personnes non-autistes (Grossman, 2015; Sheppard et al., 2016). Cette idée est d'ailleurs reprise dans le concept de la « double empathie », qui suggère que les difficultés de compréhension entre les personnes neurotypiques et autistes sont le résultat de malentendus réciproques, et non exclusivement imputables aux personnes autistes (Milton, 2012). Dans l'autisme des difficultés de partage des émotions ont été largement montrées dans la modalité visuelle (Hobson, 2019; Mundy & Sigman, 1989) mais ces atypies sont aussi présentes dans la modalité auditive au travers en particulier de la production et la perception de la voix (Paul, Augustyn, et al., 2005). Dans le domaine de la production vocale, les atypies de production de prosodie émotionnelle ont été discutées, avec des difficultés dans le placement de la voix et dans l'utilisation des différents indices acoustiques permettant d'exprimer une émotion. Une description détaillée de ces atypies de production et de perception de la prosodie émotionnelle est présentée dans le chapitre 3 de l'introduction (Perception et traitement de la prosodie émotionnelle dans l'autisme).

## Chapitre 1 : L'autisme

Afin de mieux caractériser les atypies de perception de la prosodie émotionnelle dans l'autisme, dans cette introduction sera présentée la littérature portant sur la perception et le traitement de la prosodie émotionnelle dans le développement typique et dans l'autisme. Enfin, une dernière partie sera consacrée aux travaux portant sur la réponse aux émotions et en particulier aux émotions vocales.

### En résumé

- L'autisme est un trouble du neurodéveloppement avec une prévalence de 1/100 en France, caractérisé par des difficultés dans la communication et les interactions sociales et des comportements répétitifs et spécifiques.
- Dès les premières descriptions historiques des atypies de la prosodie émotionnelle ont été évoquées.
- Une absence de réponse automatique aux émotions a aussi été montrée.
- Des difficultés dans l'adaptation et le traitement du changement ont été mises en évidence dans l'autisme.



## 2. Chapitre 2 : Perception et traitement de la prosodie émotionnelle

### 2.1. Prosodie

#### 2.1.1. Définition

Le langage est constitué de deux grandes dimensions : linguistique et paralinguistique. La linguistique correspond à l'aspect sémantique du langage qui permet de faire passer un message à travers l'utilisation de mots, alors que la dimension paralinguistique est caractérisée par la prosodie. La prosodie correspond non pas à ce que l'on dit, mais à la manière dont on le dit. Il est aussi possible de parler de « musique » ou de « mélodie » de la voix. La prosodie peut être segmentée en plusieurs fonctions (Martinet, 1982) : extralinguistique, paralinguistique et linguistique.

- La prosodie extralinguistique correspond aux caractéristiques du locuteur (sexe, âge, dialecte). Par exemple, la voix d'un enfant est plus aiguë et légère comparée à celle d'un adulte.
- La prosodie paralinguistique permet de transmettre l'état d'esprit du locuteur ou ses émotions. Par exemple, une voix enjouée et dynamique avec une intonation montante peut indiquer la joie.
- La prosodie linguistique participe à l'actualisation et la structuration des niveaux linguistiques (notamment le lexique, la syntaxe, la sémantique et la pragmatique). Par exemple, pour poser une question, l'intonation monte à la fin des phrases.

La prosodie est caractérisée par le rythme, l'accentuation et l'intonation. Ces paramètres font référence à des variations spectro-temporelles de certains paramètres acoustiques : la durée (mesurée en s) (Cummins, 2002), l'intensité sonore (en dB) (Fry, 1955) et la fréquence (en Hz) (Bolinger, 1985). La durée correspond au rythme du discours alors que l'intensité et la fréquence correspondent à l'amplitude et à la hauteur du son.

### 2.2. Prosodie émotionnelle

#### 2.2.1. Définition et fonction de la prosodie émotionnelle

D'après Mehrabian (1981), la transmission d'une émotion à travers un discours passe pour 7% par les mots, pour 38% par le ton de la voix (i.e., la prosodie émotionnelle) et pour 55% par le

langage corporel. La prosodie émotionnelle est donc plus importante que le contenu sémantique dans la transmission d'une émotion à travers le discours (Dietrich et al., 2006). Une compréhension et une perception correcte de la prosodie émotionnelle est donc essentielle dans l'interprétation, l'évaluation de l'état émotionnel et des intentions d'autrui.

### 2.2.2. Indices acoustiques de la prosodie émotionnelle

Les émotions ont chacune une signature vocale rendant leur reconnaissance possible grâce à la simple écoute de la voix. Plusieurs paramètres acoustiques permettent la reconnaissance des émotions vocales, mais le plus corrélé à cette capacité est la fréquence de la fondamentale (F0) (Banse & Scherer, 1996). Le tableau 1 ci-dessous récapitule les modifications de la voix caractéristiques de chacune des émotions primaires selon la littérature (Belin et al., 2008; Hammerschmidt & Jürgens, 2007; Laukka, 2005; Scherer, 2003; Williams & Stevens, 1972). La signature et la reconnaissance des émotions vocales semblent être assez universelles. En effet au travers de langues et de cultures différentes, il est possible de retrouver et d'identifier l'information émotionnelle transmise (Bryant & Barrett, 2008; Pell et al., 2009; Sauter et al., 2010).

	Joie	Colère	Tristesse	Dégoût	Peur
Amplitude de la fondamentale	↗↗	↗↗	↘	↗	↗
Fréquence de la fondamentale	↗↗	↗↗	↘	?	↗
Intensité	↗	↗↗	↘	↗	↗
Débit de parole	↗	↗	↘	?	↗

**Tableau 1 : Modulations acoustiques caractéristiques des émotions primaires exprimées à travers la voix**

↘ diminution ; ↗ augmentation ; ? pas de connaissance d'après la littérature

Une revue de la littérature s'est intéressée aux caractéristiques acoustiques des émotions et des sentiments positifs (bonheur, joie, euphorie, tendresse, plaisir...) en observant des vocalisations non-verbales (rire par exemple) et la prosodie. Les auteurs ont mis en évidence des similitudes entre les deux catégories de son (vocalisations et prosodie), particulièrement pour le bonheur et la joie (Kamiloglu et al., 2020). Ainsi, une augmentation de la fréquence fondamentale, de la puissance du son et des différents formants est observée. Ces

modifications sont les caractéristiques les plus claires pour l'identification d'une émotion positive dans la voix humaine (Kamiloglu et al., 2020).

### **2.2.3. Reconnaissance de la prosodie émotionnelle**

Dès l'âge de 2 mois, la prosodie utilisée pour parler à un bébé a un effet sur son attention et sa réponse. En effet, à cet âge, les bébés ont une attention plus soutenue pour du discours parlé en « parentais » (i.e., discours orienté bébé) en comparaison à un discours avec une prosodie utilisée entre adultes (Pegg et al., 1992; Thiessen et al., 2005). Chez le bébé de 7 mois, la réponse électrophysiologique est modulée par la prosodie émotionnelle, indiquant que dès les premiers mois de vie les réseaux cérébraux de réponse à ces indices ont commencé à se mettre en place (Grossmann et al., 2005). Les compétences en reconnaissance et identification des émotions vocales s'améliorent avec l'âge (Sauter et al., 2013). Chez les enfants âgés de 9 à 13 ans, cette capacité atteint un niveau où ils peuvent évaluer correctement la prosodie émotionnelle, bien que leurs performances ne soient pas encore aussi bonnes que celles des adultes (Aguert et al., 2013). Cependant, bien que la reconnaissance de la prosodie émotionnelle soit fonctionnelle chez les enfants plus âgés (9-13 ans), ils privilégient toujours le contenu sémantique plutôt que prosodique pour identifier l'état émotionnel du locuteur, contrairement aux adultes (Golan et al., 2007). Même en présence d'indices visuels, qui supportent l'émotion transmise, le contenu sémantique prend le dessus (Filippi et al., 2017). À l'adolescence, les jugements de valence et d'éveil (*arousal*) diffèrent de ceux des adultes, ce qui suggère une moins bonne capacité à discriminer les caractéristiques globales et intégrées des catégories émotionnelles. Ces propriétés holistiques incluent l'ensemble des indices vocaux, tels que le ton, l'intensité et le rythme, qui, ensemble, influencent la reconnaissance des émotions et la réponse appropriée aux signaux émotionnels dans les environnements sociaux des adolescents (Morningstar et al., 2019). Les adolescents jugeront de manière plus globale les émotions avec une intégration moins fine des indices prosodies émotionnelles, se traduisant par des jugements émotionnels moins discriminants que les adultes.

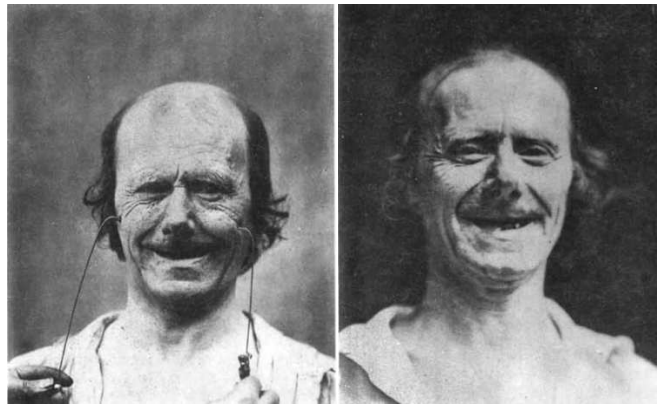
Chez les adultes, bien que la reconnaissance de la prosodie émotionnelle soit universelle à travers différentes cultures et langues (Bryant & Barrett, 2008; Pell et al., 2009; Sauter et al., 2010), des différences peuvent néanmoins apparaître. En effet, des différences en fonction du genre ont été mises en évidence avec une meilleure reconnaissance de la prosodie

émotionnelle par les femmes que par les hommes (Collignon et al., 2010; Lambrecht et al., 2014). Cependant, les émotions produites par des locuteurs sont mieux reconnues que celles produites par des locutrices (Filipe et al., 2015). Enfin, une diminution de la reconnaissance des émotions vocales avec l'avancée en âge a été observée (Lambrecht et al., 2012; Paulmann et al., 2008). Toutefois, une bonne reconnaissance des émotions vocales et faciales a été associée à un état de bien-être, plus marqué dans les populations plus âgées (Carton et al., 1999).

#### **2.2.4. Cas particulier du sourire vocal**

Le sourire correspond à une contraction bilatérale des muscles grands zygomatiques (*Zygomaticus major*, ZM). Il appartient aux unités d'action communes entre les cultures pour partager et transmettre une émotion positive (Ekman et al., 1969). Le sourire fait partie des émotions conservées dans l'évolution, partagées et exprimées par plusieurs espèces, notamment les primates (Burrows et al., 2006; Izard, 1971). Les bases neuroanatomiques supportant le sourire sont conservées au niveau phylogénétique (Burrows et al., 2006). Chez l'homme, le sourire est aussi partagé entre des individus avec des bagages culturels, des niveaux socio-culturels et des expériences sensorielles diverses, y compris chez les individus nés sourds et muets (Darwin, 2009; Eibl-Eibesfeldt, 2017; Ekman & Keltner, 1997; Izard, 1977). De même, les aveugles congénitaux sourient dans des situations sociales adaptées (Valente et al., 2018).

Le sourire a été étudié et décrit à travers l'histoire et notamment par Guillaume Duchenne de Boulogne (1806-1875) un électrophysiologiste français. L'utilisation d'électrostimulations des muscles du visage a permis de mettre en évidence leur importance dans l'expression des émotions et en particulier du sourire. Il reprend ses résultats dans un ouvrage *Mécanisme de la physiologie humaine* (Duchenne (de Boulogne), 1876). Dans ce livre, il décrit le muscle grand zygomatique comme le muscle de la « joie et de la bienveillance ». Selon lui, un sourire est considéré comme sincère lorsqu'il y a contraction des grands zygomatiques (*Zygomaticus major*) mais aussi de l'orbiculaire palpébral inférieur (*Orbicularis oculi*) (figure 1). La contraction conjointe de ces deux muscles est ainsi qualifiée de « sourire de Duchenne » (Ekman et al., 1990).



**Figure 1 : Sourire de Duchenne**

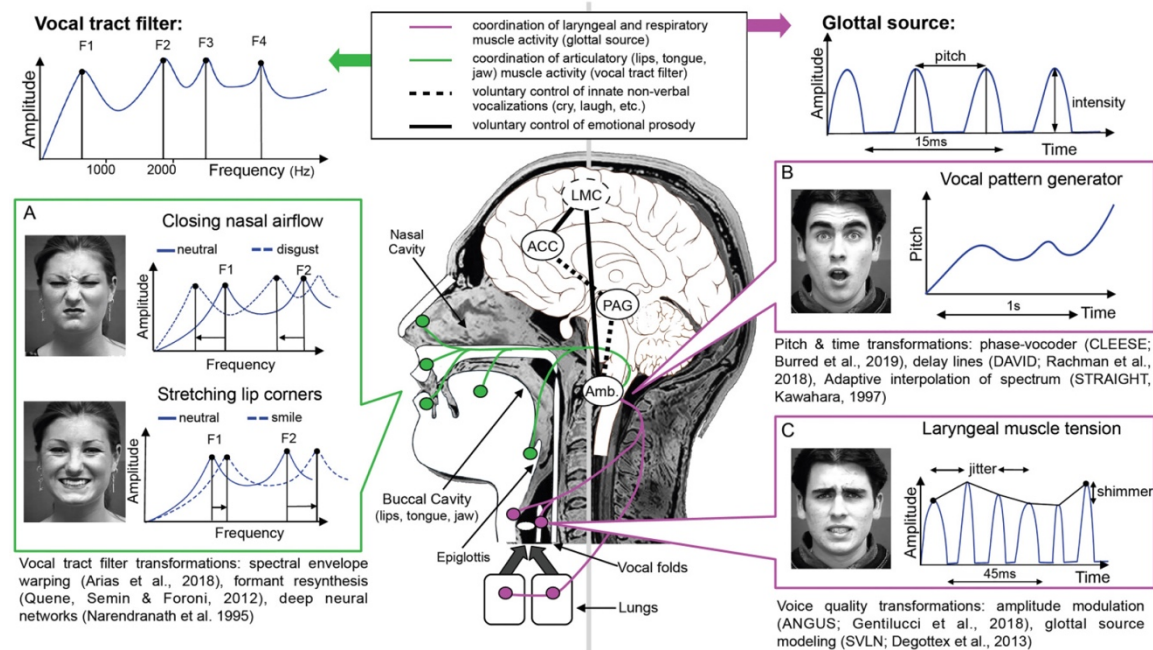
*Photographies du sourire induit par électrostimulation (à gauche) et d'un sourire sincère impliquant le grand zygomatique et l'orbiculaire inférieur (à droite) (Duchenne (de Boulogne), 1876) (depuis Wikipédia).*

Le sourire permet la transmission d'une émotion positive entre deux individus. Une personne souriante est perçue comme plus attirante, gentille et intelligente en comparaison à une personne non-souriante (U. Hess et al., 2002; Reis et al., 1990). Cela se confirme aussi dans le monde scientifique. En effet, une étude a montré que les scientifiques qui sourient le plus sur leur photo de profil, sur un réseau social scientifique, étaient plus cités (Kaczmarek et al., 2018).

Le sourire peut être utilisé pour exprimer des émotions différentes : la joie (expression d'une émotion positive), l'affiliation (expression d'une motivation sociale positive) et la domination (communication et maintien d'un statut social) (Niedenthal et al., 2010; Rychlowska et al., 2017). Certains auteurs considèrent le sourire comme une des émotions faciales les plus complexes car selon le contexte, un sourire peut avoir des significations différentes (Fischer & LaFrance, 2015; LaFrance, 2011). Ces différentes expressions de sourire impliquent la contraction de muscles faciaux différents. Cependant, en fonction du contexte et des variations de ces expressions, il est possible de reconnaître l'émotion exprimée à travers le sourire (Rychlowska et al., 2017). Un sourire de joie implique l'activation des muscles zygomatiques mais aussi un mouvement des yeux alors qu'un sourire d'affiliation est concentré uniquement au niveau de la bouche (Rychlowska et al., 2017).

Comme démontré précédemment, les émotions peuvent être reconnues uniquement à travers la prosodie en se basant sur les caractéristiques acoustiques. Le sourire ne fait pas exception (Aubergé & Cathiard, 2003; Tartter, 1980; Tartter & Braun, 1994), celui-ci peut aussi s'entendre dans la voix. En effet, la contraction des muscles zygomatiques va conduire à une modulation

du tractus vocal perceptible dans la voix (Arias, Rachman, et al., 2021) (figure 2). Le sourire vocal et ses caractéristiques acoustiques, perceptibles uniquement à travers la voix, ont été utilisés comme technique de management dans la prospection téléphonique (Basso & Oullier, 2010).



**Figure 2 : Modulation prosodique et relation avec les modifications anatomiques associées (Arias, Rachman, et al., 2021)**

Caractéristiques neurologiques, anatomiques et acoustiques de la voie de production vocale. La voix est produite lorsque le flux d'air expiratoire provenant des poumons met les plis vocaux du larynx en oscillation (à droite). Cette onde sonore, ou source glottique, est résonnée à travers le tractus vocal et filtrée par les articulateurs oraux tels que la langue et les lèvres (à gauche), puis est finalement émise par la bouche et le nez. Le contrôle neuronal à différents niveaux de cette voie implique une hiérarchie de systèmes corticaux et sous-corticaux, incluant les noyaux prémoteurs du tronc cérébral tels que le noyau ambigu (Amb.), la substance grise périaqueducatale (PAG), le cortex cingulaire antérieur (ACC) et le cortex moteur laryngé (LMC). Des techniques de transformation de la voix existent pour manipuler de manière paramétrique un signal vocal préenregistré à tous les niveaux de la voie de production de la voix : filtre du tractus vocal (A), hauteur et intonation (B), et timbre de la source glottique (C).

Des études ont examiné les indices acoustiques associés au sourire vocal et leur utilisation pour communiquer une émotion positive. Barthel et Quené (2015) ont constaté que le sourire est associé à une augmentation du deuxième formant (F2) pour les mots contenant la voyelle arrondie /o:/, ainsi qu'une augmentation de l'intensité et de la fondamentale. Dans l'étude de Podesva et al. (2015), le sourire était associé à une augmentation de F2, et dans les études de El Haddad et al. (2017; 2015), à une augmentation des fréquences des formants et de la

fondamentale. Une plus grande dispersion de F1 et F2 a également été rapportée (Drahota et al., 2008).

La littérature suggère que le sourire vocal est caractérisé par une hauteur moyenne, une intensité et des formants plus élevés (Barthel & Quené, 2015; Lasarczyk & Trouvain, 2008; Quené et al., 2012).

Une étude récente a cherché à mettre en évidence les caractéristiques acoustiques associées à la représentation perceptive d'un sourire dans une voix à l'aide d'une tâche de corrélation inverse (RevCor) (Ponsot, Arias, et al., 2018). Le principe de la méthode de la corrélation inverse est détaillé dans la partie méthodologie de ce manuscrit. Cette étude a montré dans un groupe de 10 adultes neurotypiques que la représentation perceptive (i.e., modèle interne du sourire vocal, modèle auquel une émotion est comparée pour la catégoriser, ou non, comme souriante) est commune et robuste entre les individus. Elle est caractérisée par une augmentation des fréquences des 2 premiers formants (F1 et F2) et une augmentation en énergie de F2, F3 et F4.

Le sourire est une émotion qui conduit à une synchronisation entre celui qui l'exprime et celui qui le perçoit et cela va conduire à un mimétisme facial (*facial mimicry*) (Dimberg et al., 2000). Cette réaction est présente chez le jeune enfant (Simpson et al., 2014) et dans plusieurs espèces (Burrows et al., 2006). Un Modèle Simulationniste des Sourires (SIMS, Simulation of Smiles Model, Niedenthal et al., 2010) a été proposé pour expliquer la production et la perception du sourire à l'aide de données comportementales, neuroscientifiques et physiologiques. Cependant, ce modèle spécifie que le contact visuel est essentiel dans la mise en place de ce mimétisme.

### **2.3. Traitement cérébral de la voix et des émotions vocales**

Après avoir décrit la reconnaissance des indices prosodiques, l'état des connaissances sur le traitement cérébral des émotions vocales dans le développement typique est décrit.

Le traitement de la prosodie émotionnelle repose d'abord sur celui de la voix, qui en est le vecteur principal, c'est pourquoi une brève description du traitement cérébral de la voix est également incluse dans ce chapitre. Les modèles et théories abordés seront ensuite mis en perspective dans le contexte de l'autisme dans le chapitre suivant.

### **2.3.1. Modèle de perception et de traitement de la voix et des émotions vocales**

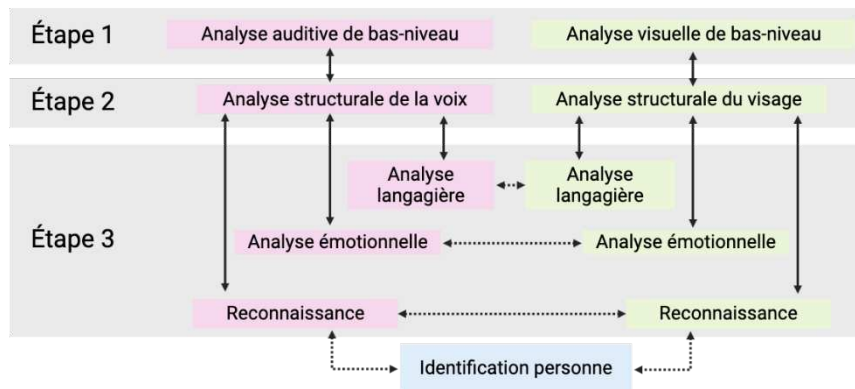
#### 2.3.1.1. Réponse à la voix

La voix est une stimulation auditive particulière pour l'être humain ; en effet nous sommes programmés pour y répondre automatiquement, spécifiquement et rapidement (Abrams et al., 2008; Belin et al., 2000, 2004; Fecteau et al., 2005). À travers la voix, c'est toute la communication entre les personnes qui est possible via les aspects linguistiques et extra linguistiques.

Une aire cérébrale spécifique, nommée *Temporal Voice Area* (TVA), située dans le Sillon Temporal Supérieur (STS), a été identifiée comme étant spécialisée dans la réponse à la voix (Belin et al., 2004). Chez les adultes neurotypiques, cette région montre une activation chez 94% des participants en réponse à des stimuli vocaux, avec une fiabilité test-retest élevée ( $r > 0,9$ ) (Pernet et al., 2015). La TVA traite la voix de manière localisée, tout comme l'aire fusiforme des visages (*Fusiform Face Area*, FFA) s'occupe du traitement des visages dans le cortex temporal inférieur (Belin et al., 2000; Kanwisher et al., 1997). Ensuite, l'information vocale est envoyée vers des voies spécifiques. Ces voies sont dédiées à l'encodage des aspects linguistiques, émotionnels ou identitaires de la voix (Belin et al., 2002; Ethofer et al., 2006; Grandjean et al., 2008; Latinus & Belin, 2011; Schall et al., 2015). La présentation d'un stimulus vocal entraîne une réponse spécifique mesurable en neurophysiologie, connue sous le nom de *Fronto-Temporal Voice Positivity* (FTVP), observée tant chez les enfants (Rogier et al., 2010) que chez les adultes (Charest et al., 2009; De Lucia et al., 2010). Cette réponse reflète l'activité des régions fronto-temporales en réaction à la voix (Belin et al., 2000; Bidet-Caulet et al., 2017). Un modèle en trois étapes, similaire à celui du traitement des visages (A. W. Young & Bruce, 2011), a été proposé pour expliquer le traitement cérébral de la voix (Belin et al., 2004, 2011) (figure 3) :

- Étape 1 : analyse de bas niveau dans les noyaux sous-corticaux et dans le cortex auditif primaire.
- Étape 2 : encodage structural pour extraire l'architecture fonctionnelle du son au niveau des régions bilatérales du STS médian et identification de la nature vocale d'un son pour un traitement distinct des sons d'autre nature.

- Étape 3: dissociation du traitement vocal en 3 voies fonctionnelles pour évaluer l'information langagière, la prosodie émotionnelle et l'identité vocale.



**Figure 3: Modèle du traitement de la voix adapté de Belin et al. (2011)**  
En rose : voie de traitement des voix ; en vert : voie de traitement des visages

Par ailleurs, les voix contenant une information émotionnelle sont traitées de manière spécifique, ce qui induit une réponse sensiblement différente.

### 2.3.1.2. Réponse cérébrale aux émotions vocales

Le traitement du contenu prosodique dans la voix est plus complexe que le traitement de la voix en tant que telle, car la quantité et la nature des informations à encoder n'est pas la même. Un modèle du traitement temporel et cortical en trois étapes de la prosodie émotionnelle a été proposé par Schirmer & Kotz (2006) (figure 4).

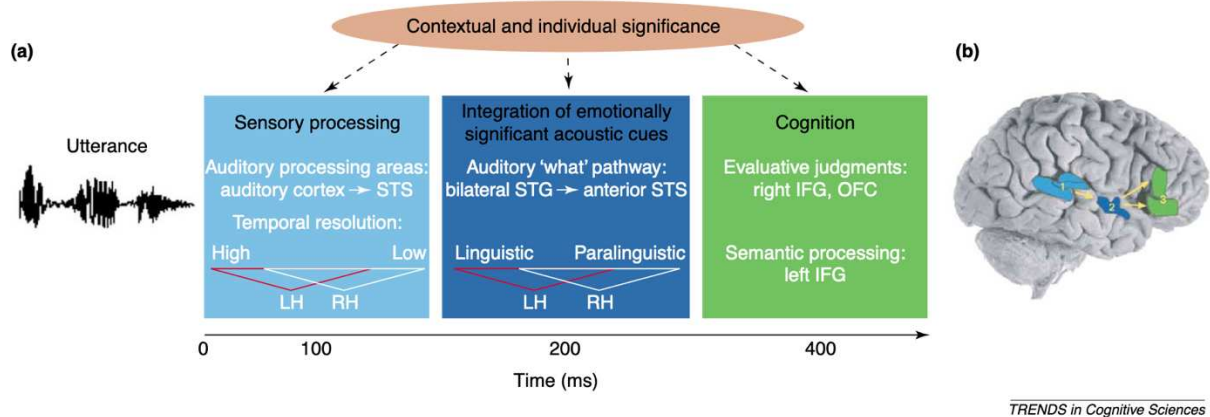
- Étape 1 : traitement des informations acoustiques (fréquence, amplitude, durée...) dans le cortex auditif primaire environ 100 ms après le début du son. Une différence de résolution temporelle entre les 2 hémisphères cérébraux pourrait être à l'origine de la spécialisation de l'hémisphère droit pour le traitement spectral et le gauche pour le traitement temporel.
- Étape 2 : encodage de la dimension émotionnelle dans le sillon temporal supérieur (STS) vers 200ms avec des modulations de la P2<sup>3</sup>.

---

<sup>3</sup> P2 : onde positive du PEA qui culmine à 200ms et qui reflète le traitement attentionnel et l'analyse initiale des stimuli vocaux émotionnels (Schirmer & Kotz, 2006).

- Étape 3 : transmission des informations vers les aires frontales pour un traitement cognitif de haut niveau.

Le modèle de Schirmer et Kotz (2006) se base sur une activation des régions temporales et frontales dans le traitement de la prosodie émotionnelle. Le contenu prosodique dans la voix conduit donc à un traitement différent et de plus haut niveau dans l'étape dite « cognitive » plus frontale et plus tardive.



**Figure 4 : Séquence d'activation cérébrale en réponse à la prosodie émotionnelle**

À partir de Schirmer et Kotz (2006). a) Modèle de travail en trois étapes pour le traitement de la prosodie émotionnelle. Traitement sensoriel (Étape 1) : L'analyse acoustique est médiée par les aires auditives bilatérales. Intégration (Étape 2) : Le traitement le long de la voie auditive du "quoi" intègre les informations acoustiques émotionnellement significatives pour en dériver un "gestalt" émotionnel. Cette voie projette du gyrus temporal supérieur (STG) vers le sillon temporal supérieur (STS) et pourrait être latéralisée dans l'hémisphère droit (RH). Cognition (Étape 3) : Les informations émotionnelles dérivées au niveau du STS sont mises à disposition pour des processus cognitifs de haut niveau. Par exemple, les jugements évaluatifs explicites de la prosodie émotionnelle sont médiés par le gyrus inférieur droit (IFG) et le cortex orbitofrontal (OFC), tandis que l'intégration de la prosodie émotionnelle dans le traitement du langage recrute le gyrus frontal inférieur de l'hémisphère gauche (LH). La signification contextuelle ou individuelle peut faciliter ou améliorer le traitement à chacune des trois étapes. b) Présentation schématique des zones cérébrales impliquées dans le traitement vocal émotionnel dans une vue sagittale droite : cortex auditif primaire, secondaire et tertiaire (bleu clair) s'étendant jusqu'à la partie antérieure du sillon temporal supérieur (bleu foncé), d'où partent des projections vers le gyrus frontal inférieur et le gyrus orbitofrontal (vert). Les flèches (jaunes) indiquent les directions présumées du traitement (les couleurs/nombres correspondent aux étapes de traitement décrites en a).

Par ailleurs, la réponse aux émotions vocales conduit à une plus grande activation du Gyrus Temporal Supérieur médian en comparaison à des intonations de voix neutres (Grandjean et al., 2005, 2005; Wiethoff et al., 2008). De plus, cette activité cérébrale apparaît corrélée aux modulations acoustiques prosodiques (intensité, fréquence de la fondamentale et durée), les

différences d'activation des aires cérébrales du traitement de la voix sont liées aux modulations prosodiques (Wiethoff et al., 2008). L'amygdale est aussi impliquée dans ce traitement (Fecteau et al., 2007; Frühholz et al., 2015; Sander et al., 2005). Une étude sur les lésions chez des patients a montré que des lésions de l'amygdale réduisaient généralement le traitement cortical des voix humaines dans l'hémisphère correspondant à la lésion. Cependant, seule une lésion de l'amygdale gauche modifiait le traitement cortical de l'affect vocal (Frühholz et al., 2015).

Au niveau de la réponse neurophysiologique, les composantes plus tardives seront donc le reflet de cette étape de traitement cognitif tandis que les composantes plus précoces reflèteront le traitement sensoriel et intégratif de la voix et du contenu émotionnel (Bostanov & Kotchoubey, 2004; Paulmann & Kotz, 2008b; Poeppel & Hickok, 2004; Sams et al., 1991; Schirmer et al., 2002). Les études en neurophysiologie ont montré des effets de la prosodie émotionnelle dans le traitement sensoriel de la voix dans l'enfance (Chronaki et al., 2012; Grossmann et al., 2005). Chez l'adulte, les différences entre la réponse à des voix avec un contenu émotionnel et les voix neutres ont été mises en évidence pour la P2 (Fan & Cheng, 2014; Paulmann & Kotz, 2008a; Pinheiro et al., 2017), mais aussi pour des composantes plus tardives, comme la LPC (*Late Positive Component* ; ~ 450-700ms) modulée par le contenu émotionnel de vocalisations (Pell et al., 2015).

### **En résumé**

- La voix est un vecteur de transmission de l'état émotionnel d'un locuteur à travers la prosodie émotionnelle.
- Les émotions transmises dans la modalité auditive se basent sur des représentations mentales semblables entre les individus neurotypiques.
- Le sourire est une émotion particulière qui possède une signature acoustique reconnaissable et identique chez l'adulte neurotypique.
- Une réponse cérébrale spécifique à la voix et aux émotions vocales a été décrite chez l'enfant et l'adulte neurotypique.

### **3. Chapitre 3 : Perception et traitement de la prosodie émotionnelle dans l'autisme**

#### **3.1. Prosodie émotionnelle dans l'autisme**

##### **3.1.1. Atypies de production**

Dès les premières descriptions historiques de l'autisme, des atypies de production de la prosodie ont été mises en évidence (Kanner, 1943). Kanner disait de certains des 11 enfants qu'il a décrits : « Il ne pose jamais de questions sous la forme de questions (avec l'intonation appropriée) », « Il veut [...] toujours prononcer les lettres de la même façon ». Dans ses descriptions initiales il ne rapporte pas ces éléments pour tous les cas qu'il présente (absence pour les cas n°2, 3 et 6). Asperger aussi décrit le volet productif de la prosodie chez ses patients spécifiant une absence de modulation prosodique ou bien des modulations excessives. Ssucharewa en parle aussi dans le volet « Particularités de la voix et du langage » de sa description de l'autisme en évoquant une « voix monotone ». Les atypies de production étaient donc un élément essentiel du diagnostic dès les descriptions princeps de l'autisme.

Plus récemment il a été montré des atypies de production de la prosodie dans différents versants de celle-ci : grammaticale, pragmatique et émotionnelle (McCann & Peppé, 2003). En termes de prosodie émotionnelle il a aussi été mis en évidence des atypies rendant la perception « étrange » (*awkward*) et contribuant aux difficultés de communication sociale dans l'autisme (Grossman et al., 2013; Grossman, 2015). Les modulations de la prosodie chez les personnes autistes ont été expliquées dans la littérature par une plus grande variabilité dans les indices acoustiques lors de la production (Banse & Scherer, 1996; Shriberg et al., 2001). Ces atypies de prosodie émotionnelle conduisent à des difficultés de sociabilisation chez les enfants autistes (Paul, Shriberg, et al., 2005).

Ces atypies prennent une place conséquente dans le diagnostic d'autisme. En effet, dans l'ADOS (Lord et al., 2000), les productions prosodiques sont un des éléments évalués. Dans le DSM-5 les atypies de production de la prosodie émotionnelle font partie des éléments centraux dans l'établissement du diagnostic d'autisme (American Psychiatric Association, 2013).

Néanmoins, les atypies de production de la prosodie dans l'autisme pourraient être la conséquence d'une perception atypique de celle-ci (Kargas et al., 2016; Peppé et al., 2007). En

effet, si la perception des émotions vocales ne se basent pas sur des caractéristiques acoustiques correctes, alors une production adaptée sera plus difficile.

### **3.1.2. Atypies de perception et de reconnaissance des émotions vocales**

Dès les premières descriptions, tout comme la production, la perception et la reconnaissance des émotions vocales ont été identifiées comme étant atypiques dans l'autisme. Kanner disait d'un des patients qu'il a décrit : « Pendant un moment, on a cru qu'il était sourd car il ne montrait aucun changement d'expression lorsqu'on lui parlait ou lorsqu'il était en présence d'autres personnes » ; ou encore pour un autre patient : « Cela ne faisait aucune différence qu'on lui parle de manière amicale ou durement » (Kanner, 1943).

Les difficultés liées à la prosodie réceptive sont fréquemment signalées chez les enfants et les adultes autistes (Peppé et al., 2007). Des difficultés de reconnaissance de la prosodie émotionnelles ont été mises en évidence dans l'autisme chez l'enfant (Yoshimatsu et al., 2016) et chez l'adulte (Hubbard et al., 2017). Il a été proposé à travers une méta-analyse, que les difficultés de reconnaissance des émotions vocales étaient plus marquées chez les adultes que chez les enfants autistes (Leung et al., 2022). Cependant une étude avec des enfants autistes d'environ 9 ans a montré que leurs compétences de reconnaissance des émotions vocales étaient similaires à celles d'enfants neurotypiques de 3-4 ans (Yoshimatsu et al., 2016).

Chez les neurotypiques, la colère et la tristesse sont mieux reconnues que la peur et la joie (Chronaki et al., 2018; Elfenbein & Ambady, 2002; Juslin & Laukka, 2003). Par ailleurs, les difficultés dans la reconnaissance des émotions dans l'autisme apparaissent comme étant dépendantes de la nature de l'émotion. En effet, des altérations spécifiques aux émotions ont été observées chez les adultes autistes dans la reconnaissance de la joie, une émotion positive (Hubbard et al., 2017; J.-E. Wang & Tsao, 2015).

Cependant, d'autres études n'ont pas montré de différence dans la reconnaissance d'émotions vocales entre les autistes et les neurotypiques (Brennand et al., 2011; Brooks & Ploog, 2013; Chevallier et al., 2011; Grossman et al., 2010; Paul, Augustyn, et al., 2005). Il est donc difficile de généraliser ces difficultés qui apparaissent comme étant dépendante du paradigme expérimental et des participants inclus dans les études. En effet, il semble que la

reconnaissance des émotions vocales (Golan et al., 2007) et musicales (Heaton, Davis, et al., 2008; Quintin et al., 2011) soit liée aux compétences verbales.

Malgré les difficultés mises en évidence dans la reconnaissance des émotions vocales, il a été montré dans la littérature que la discrimination de la hauteur des sons (sons complexes, vocaux et purs) était supérieure dans l'autisme (Bonnell et al., 2003, 2010; Eigsti & Fein, 2013; Heaton, Hudry, et al., 2008; Stanutz et al., 2014). Néanmoins, ces bonnes capacités, de discrimination acoustique n'apparaissent pas comme suffisantes pour une bonne reconnaissance de toutes les émotions vocales.

A ce jour, le traitement et la reconnaissance du sourire vocal dans l'autisme n'ont pas été explorés. Pourtant, l'absence de réponse et de production de sourire social est l'un des premiers signes d'alerte chez le jeune enfant (Beall et al., 2008; G. Dawson et al., 1990). Comprendre les indices acoustiques utilisés dans la reconnaissance du sourire vocal dans l'autisme permettrait d'éclairer, dans le contexte de la prosodie émotionnelle, les difficultés de réponse à ce signal social essentiel.

De même, aucune étude n'a cherché à déterminer les représentations internes de la prosodie émotionnelle dans l'autisme. Une seule étude a utilisé la corrélation inverse dans la modalité auditive pour évaluer le ton de la voix et la perception d'une mélodie chez un groupe d'enfants autistes (L. Wang et al., 2022). Dans cette étude, les auteurs ont cherché à identifier les caractéristiques acoustiques associées aux représentations mentales d'une prosodie interrogative mais aussi d'une mélodie (l'équivalent de la comptine *Frère Jacques* en Chine) chez des enfants. Les résultats indiquent que les représentations sont similaires entre les enfants neurotypiques et les enfants autistes, mais une plus grande variabilité des réponses est mise en évidence dans le groupe d'enfants autistes, reflétant une hétérogénéité des représentations mentales. Il n'existe pas pour le moment d'étude cherchant à déterminer les indices acoustiques sur lesquels se basent les représentations mentales des émotions vocales dans l'autisme.

Un traitement et une réponse cérébrale atypique pourrait expliquer l'utilisation d'indices acoustiques atypiques dans les représentations des émotions vocales. Si ces indices sont éloignés de ceux des neurotypiques, alors cela pourrait expliquer, entre autres, les difficultés

de reconnaissance de la prosodie émotionnelle mais aussi les réponses physiologiques et cérébrales atypiques.

### **3.1.1. Traitement cérébral de la voix et de la prosodie émotionnelle dans l'autisme**

Dans l'autisme, peu d'études se sont intéressées au traitement des émotions vocales en comparaison au traitement des visages. En effet, une recherche *Google Scholar* avec les termes « *autism 'vocal emotion' processing* » donne 1490 résultats contre 10 fois plus de résultats pour la recherche « *autism 'facial emotion' processing* » (15100 résultats). Seulement quelques études ont cherché à mesurer les réponses cérébrales en réponse à la voix chez des personnes autistes en IRMf (Imagerie par Résonance Magnétique fonctionnelle). Une première étude, en 2004, avait montré une absence de réponse spécifique à la voix chez 3/5 des adultes autistes testés (Gervais et al., 2004), due à une non activation du sillon temporal supérieur en réponse aux stimulations vocales. Plus récemment d'autres études n'ont pas montré de différence entre des adultes autistes et non-autistes dans l'activation de la TVA (Charpentier et al., 2020; Schelinski et al., 2016). Enfin, une étude a rapporté une connectivité réduite et corrélée négativement avec les traits autistiques entre les aires frontales et temporales en réponse à la voix qui pourrait sous-tendre les difficultés de réponse à la voix dans l'autisme (Hoffmann et al., 2016).

La réponse électrophysiologique à la voix dans l'autisme a été décrite comme étant différente chez l'enfant, avec une FTPV atypique (Bidet-Caulet et al., 2017). Néanmoins, une analyse plus fine des résultats a permis de mettre en évidence que cet effet serait la conséquence d'un traitement atypique des sons de l'environnement et non de la voix (Latinus et al., 2019). Chez les adultes autistes, aucune différence dans le traitement de la voix n'a été constatée. En revanche, chez les enfants, la réponse physiologique varie en fonction de la nature émotionnelle du son (Charpentier, Kovarski, Houy-Durand, et al., 2018). Une autre étude a aussi reporté une absence de différence dans la réponse à la voix chez des adultes autistes (Lepistö et al., 2007). Face aux résultats hétérogènes entre les études, une compréhension fine des indices utilisés lors du traitement de la voix dans l'autisme contribuera à expliquer cette diversité.

Au niveau de la réponse cérébrale aux émotions vocales dans l'autisme, les études ont obtenu des résultats différents selon la complexité des émotions présentées. En effet, une absence de différence dans l'activation des aires fronto-temporales entre un groupe d'adultes

neurotypiques et autistes a été mise en exergue en réponse à des émotions basiques (i.e., colère, tristesse, joie et surprise) (Rosenblau et al., 2016). En revanche, une différence a été observée chez les enfants autistes par rapport au groupe contrôle. Ces enfants présentaient une réduction de l'activation du cortex préfrontal et du gyrus temporal médian en réponse à une prosodie plus complexe, comme l'ironie (A. T. Wang et al., 2007). En plus des atypies dans le traitement des émotions complexes, des modifications de l'activité cérébrale ont également été observées pour les émotions simples. Ainsi, les nouveau-nés à haut risque d'autisme<sup>4</sup> montrent une sensibilité réduite aux vocalisations de tristesse dans le gyrus fusiforme droit et l'hippocampe gauche (Blasi et al., 2015). Plus tard, à l'adolescence, la perception d'émotions vocales provoque des différences d'activation dans les aires frontales, temporales et pariétales (Eigsti et al., 2012). Ces atypies persistent jusqu'à l'âge adulte, les personnes autistes montrant une activation plus forte du gyrus supramarginal par rapport aux neurotypiques, ainsi qu'une désactivation plus faible du réseau par défaut en réponse à des stimuli vocaux émotionnels (Hesling et al., 2010).

Dans le traitement de la prosodie émotionnelle, peu d'études ont mesuré les réponses électrophysiologiques évoquées par des émotions vocales dans l'autisme. Chez l'enfant autiste, la littérature rapporte un allongement de la latence et une diminution de l'amplitude de la réponse (Charpentier, Kovarski, Houy-Durand, et al., 2018; Fan & Cheng, 2014; Korpilahti et al., 2006). Chez l'adulte autiste, il n'a pas été mis en évidence de différence dans les réponses neurophysiologiques en réponse aux émotions vocales (Charpentier, Kovarski, Houy-Durand, et al., 2018). Ces différences entre les adultes et les enfants autistes pourraient être la conséquence de stratégies de compensation mises en place au cours du développement (Hull et al., 2017; Livingston & Happé, 2017).

Le contenu prosodique influe également sur la détection du changement dans l'autisme et dans le développement typique avec une diminution de la latence de la *mismatch negativity* (MMN) pour les déviants émotionnels en comparaison à des sons neutres dans les deux groupes (Charpentier, Kovarski, Roux, et al., 2018). La MMN est une réponse neurophysiologique induite

---

<sup>4</sup> Les enfants à haut-risque d'autisme sont les enfants qui appartiennent à une fratrie avec un-e ainé-e avec un diagnostic d'autisme (Jones et al., 2014). Environ 20% de ces enfants seront eux-mêmes diagnostiqués plus tard (Ozonoff et al., 2011).

par un changement dans une séquence répétitive, elle est considérée comme un indice cérébral de la détection du changement (Näätänen et al., 1978). Les réponses cérébrales impliquées dans le traitement des changements émotionnels spécifiques sont déjà présentes pendant l'enfance, mais elles n'ont pas encore atteint le pattern observé chez les adultes (Charpentier, Kovarski, Roux, et al., 2018). Ces différences de traitement pourraient contribuer à l'amélioration de la perception de la prosodie émotionnelle entre l'enfance et l'âge adulte.

Dans l'autisme, les enfants répondent aussi plus rapidement aux déviants émotionnels mais les amplitudes de MMN sont réduites pour les sons neutres et émotionnels (Charpentier, Kovarski, Houy-Durand, et al., 2018). Chez les adultes autistes, la détection du changement est intacte, quelle que soit la valence du son vocal (Charpentier et al., 2020). Les auteurs suggèrent que l'absence de différence dans le traitement de la voix chez les adultes autistes, malgré les difficultés comportementales rapportées, pourrait s'expliquer par des difficultés présentes dès l'enfance (Charpentier, Kovarski, Houy-Durand, et al., 2018). Pourtant, le traitement des aspects acoustiques du son semble, paradoxalement, préservé, voire amélioré, dans l'autisme. En effet, des études ont montré de meilleures performances dans la discrimination de hauteur chez des adultes autistes (Bonnell et al., 2003, 2010; Samson et al., 2006). Il semblerait donc que l'ajout d'informations sociales à analyser rendent le traitement atypique de la voix et de la prosodie dans l'autisme.

Il n'existe actuellement aucun consensus sur l'intégrité des réponses à la voix et aux émotions vocales chez les personnes autistes, et les recherches sur les processus de détection du changement ont révélé des atypies particulièrement marquées en présence de stimuli émotionnels. Ces différences dans la détection des changements émotionnels pourraient ne pas être dues à des difficultés d'identification de la prosodie émotionnelle elle-même. Elles pourraient ainsi provenir de difficultés à comparer les émotions entendues avec les représentations perceptives établies lors de la phase d'adaptation neurale précédant la détection du changement. La compréhension de ce phénomène dans l'autisme est donc importante et tout particulièrement pour la prosodie émotionnelle. Si des différences existent dans la représentation neurale des émotions vocales chez les personnes autistes (Golan et al., 2007; Uljarevic & Hamilton, 2013), la comparaison entre perception et représentation mentale pourrait entraîner une erreur de prédiction. Cela serait particulièrement marqué dans le cadre

des modèles du cerveau prédictif et bayésien (Gomot & Wicker, 2012; Van de Cruys et al., 2014).

En raison des atypies de détection du changement encore plus marquées dans l'autisme pour des stimuli vocaux émotionnels (Charpentier, Kovarski, Houy-Durand, et al., 2018), l'étude de l'adaptation neurale à la voix et à la prosodie émotionnelle est particulièrement importante pour expliquer ces différences.

### **3.1. Adaptation et encodage de la régularité des émotions vocales.**

Aucune étude n'a, à ce jour, testé l'effet de la prosodie émotionnelle en utilisant un paradigme adapté chez des adultes neurotypiques et autistes pour évaluer l'hypothèse de l'adaptation neurale dans la mise en place de la MMN (Jaaskelainen et al., 2004; Näätänen et al., 1978), et ainsi expliquer les difficultés rencontrées dans l'autisme. Une seule étude a cherché à étudier l'adaptation à des sons de nature électronique et humaine chez des enfants autistes et a montré une réduction de l'adaptation pour les deux natures de sons en comparaison au groupe d'enfants neurotypiques. Toutefois, le protocole de *oddball* utilisé n'était pas adapté à l'étude de l'adaptation neurale à proprement parler (Ruiz-Martínez et al., 2020). Chez l'adulte neurotypique une seule étude s'est intéressée à l'adaptation neurale à des sons vocaux et non-vocaux, cette fois-ci avec un protocole de *roving*, et a montré une adaptation pour les sons vocaux mais pas pour les sons non-vocaux (Ylinen & Huotilainen, 2007). Ces différences selon la nature du son dans l'adaptation neurale ont été attribuées à un effet de familiarité pour les sons vocaux. Cependant, dans cette étude aucune *Repetition Positivity* (i.e., la RP est l'indice préférentiel de mesure de l'adaptation neurale) n'a été mise en évidence à cause du nombre réduits de répétitions comparées, mais aussi à cause de la multiplicité des voyelles testées rendant l'observation de l'adaptation neurale à ces sons difficile. Aucune étude ne s'est pour le moment intéressée à l'adaptation neurale à la voix et aux émotions vocales dans l'autisme.

La suite de ce chapitre est constituée d'une revue de la littérature publiée qui reprend l'état de l'art sur l'habituation comportementale, l'adaptation neurale et les processus sous-jacents dans le développement typique et dans certains troubles du neurodéveloppement.

## **Habituation, Adaptation and Prediction Processes in Neurodevelopmental Disorders: A Comprehensive Review**

*Merchie, A. & Gomot, M.*

*Article publié dans Brain Sciences (2023)*

Dans l'autisme, des difficultés comportementales en réponse à des changements dans l'environnement ont été observées (Crespo et al., 2013). Ces difficultés ont aussi été mises en évidence au niveau neurophysiologique au travers des atypies de détection du changement (Gomot et al., 2002, 2011). Plusieurs processus peuvent être impliqués dans la détection du changement dont l'habituation au contexte. En effet, pour détecter un changement il faut pouvoir le reconnaître au regard d'un environnement stable. De nombreuses études se sont focalisées sur les processus de détection du changement mais peu ont exploré plus précisément le processus d'adaptation qui précède.

Nous nous sommes donc intéressées à la littérature portant sur l'adaptation neurale mais aussi sur l'habituation comportementale dans l'autisme ainsi que dans deux autres troubles du neurodéveloppement dans lesquels des atypies de détection du changement ont aussi été mises en évidence : le TDAH (Trouble du Déficit de l'Attention avec ou sans Hyperactivité) et la schizophrénie.

L'objectif de cette revue de la littérature était de faire le point sur les aspects neurophysiologiques, cognitifs et comportementaux, tant dans le développement typique que dans les troubles du neurodéveloppement. Elle visait également à rassembler les connaissances empiriques sur ces phénomènes au niveau neuronal chez l'homme et l'animal, tout en offrant une synthèse des différents modèles explicatifs proposés.

Review

# Habituation, Adaptation and Prediction Processes in Neurodevelopmental Disorders: A Comprehensive Review

Annabelle Merchie and Marie Gomot \*

UMR 1253 iBrain, Université de Tours, INSERM, 37000 Tours, France; annabelle.merchie@inserm.fr

\* Correspondence: gomot@univ-tours.fr; Tel.: +33-2-4747-8664

**Abstract:** Habituation, the simplest form of learning preserved across species and evolution, is characterized by a response decrease as a stimulus is repeated. This adaptive function has been shown to be altered in some psychiatric and neurodevelopmental disorders such as autism spectrum disorder (ASD), attention-deficit/hyperactivity disorder (ADHD) or schizophrenia. At the brain level, habituation is characterized by a decrease in neural activity as a stimulation is repeated, referred to as neural adaptation. This phenomenon influences the ability to make predictions and to detect change, two processes altered in some neurodevelopmental and psychiatric disorders. In this comprehensive review, the objectives are to characterize habituation, neural adaptation, and prediction throughout typical development and in neurodevelopmental disorders; and to evaluate their implication in symptomatology, specifically in sensitivity to change or need for sameness. A summary of the different approaches to investigate adaptation will be proposed, in which we report the contribution of animal studies as well as electrophysiological studies in humans to understanding of underlying neuronal mechanisms.

**Keywords:** habituation; neural adaptation; development; neurodevelopmental disorders

**Citation:** Merchie, A.; Gomot, M. Habituation, Adaptation and Prediction Processes in Neurodevelopmental Disorders: A Comprehensive Review. *Brain Sci.* **2023**, *13*, 1110. <https://doi.org/10.3390/brainsci13071110>

Academic Editor: Mario Liotti

Received: 13 June 2023

Revised: 12 July 2023

Accepted: 20 July 2023

Published: 21 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Habituation is an essential behavioral process found in all individuals. It allows individuals to adapt to their environment, by being less focused on an irrelevant repeated stimulus, and to detect and respond more quickly to change. This process requires the construction and constant updating of a sensory memory trace following the presentation of a stimulus, through which a repeated stimulus will be considered as regular [1]. A possible processing link to habituation has been described at the neural level, neural adaptation, which results in a decrease in neuronal activity upon repeated presentation of the same stimulus, usually referred to as repetition suppression (RS). Studies on RS were first conducted on animals, revealing this phenomenon at the level of the individual neuron, and then in humans, through function magnetic resonance imagery (fMRI) [2,3], and electroencephalography (EEG) [4,5]. Another phenomenon complementary to RS has been described at the neural level, repetition enhancement (RE), defined as an increase in the neural response with an increase in the number of repetitions, reflecting the anticipation and the expectation of the stimulus appearance [6]. RS and RE, together with prediction error (i.e., difference between the expectancy and the sensory stimuli [7]), are involved in prediction processes. Individuals with neurodevelopmental or psychiatric disorders such as autism spectrum disorder (ASD) or schizophrenia have previously been observed to exhibit atypical behavioral habituation [8,9], and, several studies have also shown atypical adaptation at the neural level [10] and possible atypical prediction processes [11].

The purpose of this comprehensive review is to describe the processes of behavioral habituation, neural adaptation, and prediction and how they have been studied throughout typical development. Adaptation and habituation have been studied extensively, but

we will present and discuss these processes in this review more specifically in light of the more recent theory of predictive coding. Then, we will discuss their maturation and how they are impacted in three neurodevelopmental and psychiatric disorders: autism spectrum disorder (ASD), schizophrenia, and attention-deficit/hyperactivity disorder (ADHD).

## 2. From Behavioral Habituation to Neural Adaptation and Prediction Processes

### 2.1. Behavioral Habituation

Habituation is a behavioral process that has been studied for at least 50 years [12,13] and considered as a simple form of learning [13–15]. It is defined as a decrease in the behavioral response to a repeated stimulus [15]. In animals, habituation has been widely studied in *C. elegans* [16], in rat [17], in *Aplysia* [18] and in many other species (see Rankin et al., 2009 for a review [15]). In humans, habituation has been described in several modalities such as auditory [19,20] and smell [21]. Habituation has been extensively studied in children, but not as many studies have specifically investigated it in adults. A description of the processes during typical development will be given in Section 3.1.

The study of the neural mechanism underlying behavioral habituation are essential to describe the link between habituation and neural adaptation.

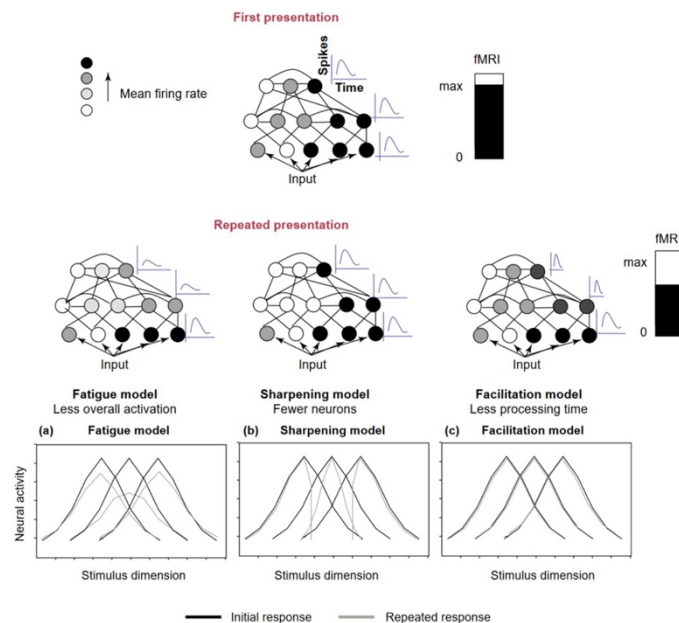
### 2.2. Repetition Suppression (RS)

RS is a cerebral phenomenon reflecting neural adaptation upon presentation of a repeated stimulus. It is defined as a decrease in neural response as the number of repetitions of the same stimulus increases [6,22,23]. Three explanatory models of RS have been proposed in the literature (see Grill-Spector et al. 2006, for an overview [23]).

(1) The fatigue model, according to which each neuron that initially responds to the repeated stimulus sees its response decrease in proportion to its initial involvement. Thus, the stronger the initial response of a neuron, the more its response decreases with repetition. This results in a general decrease in response without modification of its latency or organization (Figure 1a) [24,25].

(2) The sharpening model, in which only the neurons responding to the stimulus and coding for irrelevant features see their response decrease with the repeated presentation of the same stimulus. Therefore, neurons that do not decrease their response at all, or soon after the presentation of a stimulus, will be more responsive to that stimulus (Figure 1b) [22,26].

(3) The facilitation (or accumulation) model where the repetition of a stimulus would lead to faster processing of that stimulus by the neurons involved, resulting in a shorter latency or response time (Figure 1c) [27].



**Figure 1.** Representation of the three models that explain repetition suppression according to the review of Grill-Spector et al. (a) the Fatigue model: with repetition, the overall response will decrease in proportion to the initial response, but there will be no change in the preferred stimulus and tuning width., (b) the Sharpening model: Repeating the stimulus will lead to a reduction in the tuning bandwidth, resulting in a more focused response centered on the preferred stimulus, (c) The Facilitation model: does not offer specific or definitive predictions for changes in tuning curves. (Reprinted from “Repetition and the brain: Neural models of stimulus-specific effects”. Trends in Cognitive Sciences, 10(1), 14–23, Grill-Spector, K., Henson, R., and Martin, A. (2006), with permission from Elsevier) [23].

RS has been studied in both animals, using single-cell recording, and humans, using fMRI, magnetoencephalography (MEG) and EEG.

### 2.2.1. In Animals

Initially, studies revealed RS in animals at the individual neuron level. This phenomenon, called stimulus-specific adaptation (SSA), corresponds to a decrease in the response of a single neuron with an increase in the number of repetitions of the same stimulus (standard) without a decrease in the response to rare stimuli (deviant) [28]. RS was first described in the inferior temporal cortex (IT) of macaque [29–39] and was also found in the rat primary visual cortex [40]. Another study conducted in the IT of macaques also revealed a modulation of the RS by GABA<sub>A</sub>-mediated inhibition [41].

In the auditory modality, the first study observing SSA was conducted in the primary auditory cortex (AC1) and the medial geniculate body (MGB) of cats [42]. A decrease in the response of some neurons during repetition and a larger response to deviant non-impacted by repetition were found only in the AC1. Subsequently, other studies have shown SSA in the AC in other species [43–47]. At the subcortical level, several auditory studies highlighted SSA both in the inferior colliculus (IC) [48–51] and in the medial geniculate body (MGB) [52–55]. Some of these studies were interested in the involvement of

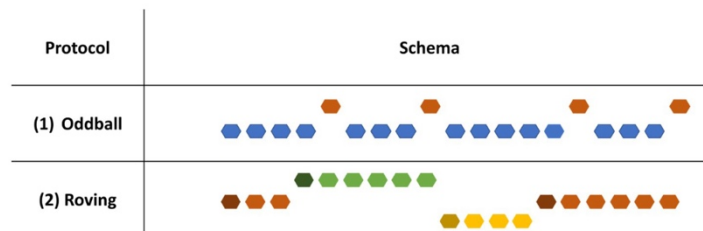
AC in the generation of subcortical SSA and reported remaining SSA even after AC deactivation, although some of the subcortical neurons displayed reduced SSA [48,54], reflecting the involvement of AC in the modulation of subcortical SSA. Some of the studies in rodents highlighted the involvement of GABA<sub>A</sub>-mediated inhibition on the SSA, showing especially a reduced SSA after application of an antagonist of GABA<sub>A</sub> receptors [51,55] or an increased SSA after application of an agonist of GABA<sub>A</sub> receptors [55] suggesting the modulation of SSA by GABA<sub>A</sub>-mediated inhibition.

The SSA is thus a reliable neural phenomenon across species that has been observed, at both the cortical and the subcortical level, in both auditory and visual modalities.

### 2.2.2. In Humans

In humans, many studies have first examined the effects of repetition indirectly through the study of change detection [56,57]. For that, studies were conducted using oddball paradigms (see Figure 2) in which a repeated stimulus, called standard, is occasionally replaced by a new and different stimulus, called deviant. Results from these studies have focused on the response to the deviant stimulus, which reflect violation detection. In fMRI studies, change detection is reflected by an increase in brain activity in response to the deviant compared to the standard stimulus [58,59]. In EEG studies it is reflected by mismatch negativity (MMN), a negative component between 150 and 200 ms obtained by subtracting the response following the repetition of a standard stimulus to the response evoked by a changing deviant stimulus [60]. MMN thus appears to occur when a stimulus is incongruent with the memory representation of the preceding repeated stimuli [61,62], making it an indirect index of neuronal adaptation. However, recent studies have focused directly on the process reflecting the encoding of the repeated stimulus, the RS [57,58,63].

Although oddball paradigms are still used in some EEG studies [57], another type of paradigm more suitable for studying RS has been created, the roving paradigm, in which a stimulus is repeated  $n$  times and then followed by a new stimulus, which is also repeated  $n$  times (Figure 2). This type of paradigm leads to a continuous update of the memory trace that is suppressed at the end of each stimulus train [4,5,64–66]. Using this type of paradigm in the auditory modality, electrophysiological studies have highlighted that neuronal adaptation, in response to repeated sounds, appears to result in a combined decrease in the negative N1 component (a negative deflection of the response around 100 ms) and an increase in the positive P1 and P2 components (positive deflections of the response, respectively, around 100 and 200 ms), reflecting the adaptation of the response following the repeated presentation of a stimulus. By comparing the responses to a small number of repetitions of a new stimulus to the same stimulus after several repetitions, it was possible to isolate an electrophysiological index of auditory regularity encoding, repetition positivity (RP), which corresponds to a positive deflection between 50 and 250 ms, which increases as the number of repetitions increases [4,5,66,67]. Roving paradigms have also shown an “MMN memory trace effect”, i.e., an increase in MMN amplitude with an increase in the number of repetitions [4,5,66,68,69], reinforcing the idea that MMN would be an indirect index of adaptation. Moreover, it has been suggested that RP could be linked to SSA observed in animals [5,42,67].



**Figure 2.** Representation of the different types of paradigms evoked in this review to study neural adaptation (repetition suppression). (1) Classic oddball sequence, in blue: standard, in orange: deviant. (2) Roving sequence: a stimulus is repeated several times followed by another stimulus; the first stimulus of each train (darker in the schema) corresponds to the deviant stimulus.

RP has been observed for sounds of different natures: pure tones [4,56,67,68,70], or more complex acoustic sounds such as vocal sounds [71,72]. However, the complexity and the quantity of information to encode appears to influence the dynamic to reach a stable neural adaptation: to human voice adaptation would require more repetitions for the stabilization of P1 amplitude compared to their equivalent non-vocal sounds [71]. Another study did not draw the same conclusion with no RP for vocal sounds by only compared a few number of repetitions [73]. Using a roving paradigm with vocalizations varying in prosody, an enhanced effect of positive emotional content on the RP compared to angry and neutral vocalizations was observed [72]. Neural adaptation occurs in typical adults for sounds of different natures, but the natures of the sound could influence the pattern to reach a stable adaptation.

In fMRI studies, RS results in a decrease in brain activity with an increase in the number of repetitions of a standard and is referred to as fMRI adaptation [25]. To investigate fMRI adaptation, some studies have used pairs of images, with some of the pairs composed of the same two images and some consisting of two different images [74] and other studies have used classical oddball paradigms as in EEG [75–77]. In 2006, Sayres and Grill-Spector studied the implication of the duration of repetition of the same images [58]. In case of short presentation, the RS was significant, but with a smaller amplitude than for long presentations, suggesting as in other studies the involvement of attention in the adaptation process [24,78,79].

As in EEG, several fMRI studies have been conducted on deviance detection, showing an increase in the cerebral activity in response to the deviant compared to the standard stimulus [75]. Using oddball paradigms, a difference in brain activation in response to a deviant and to a standard stimulus was observed, that could be assimilated to the MMN and P3a (i.e., an index of attention orientation) in EEG studies [75–77,80–82]. In these studies, beside the expected Superior Temporal Gyrus (STG) activation, an involvement of the Inferior Frontal Gyrus (IFG) has repeatedly been reported [75,76,80–82].

More recently, studies using roving protocol have also been done in fMRI to study deviance detection, regularity encoding and prediction. Using a roving protocol with nine pure tones and four train lengths (4, 12, 24 and 36 repetitions) to localize the effect of adaptation, a variation in the intensity of the response to a deviant tone according the number of standards previously presented has been showed, which is consistent with prediction theory, and a dissociation between the regularity encoding and deviance detection in cortical subfields has been demonstrated [83].

### 2.3. Repetition Enhancement (RE)

The RE phenomenon has not been as well studied and described as RS but it is assumed that it corresponds to the expectation of stimulus appearance. It thus rather reflects

prediction, i.e., the ability to anticipate a stimulus on the basis of previous experience [84]. RE is defined as an increase in the neural response with an increase in the number of repetitions [6,85]. Some studies observed RE with repetition, instead of RS, but only in response to degraded stimuli and concluded that RE could be a consequence of lack of access to memory representations through poor stimuli quality, preventing attenuation through the increasing perceptual performance. For example, Turk-Browne et al. [74] found RS when a scene with high visibility was repeated but reported RE when the same low-visibility scene was repeated. Thus, the nature of the stimulus could impact on the presence of either RS or RE. The difference could also reflect the features of the repeated stimulus, with RS related to the form and RE related to the size of the presented object in the same subjects [86].

Several models have proposed explanation for RE [6], among which are the accumulation model [27] and the novel network formation [87].

(1) The accumulation model proposes that the cumulative effect of repeated exposures leads to an enhanced or amplified response to the stimulus compared to its initial presentation, especially in case of qualitatively degraded stimulus (e.g., with low visibility) [27,87].

(2) In the novel network formation model, the presentation of a new stimulus leads to the creation of a new neural network coding for that stimulus. RE would thus reflect the creation of a new representation [6].

Studies intend to determine the difference between RE and RS in the brain regions implicated and in the timing of the setting up, by manipulating the nature of the stimulus or the paradigm. Comparing predictability of conditions, the RP occurs earlier in the more predictable condition, which is consistent with the idea of the novel network formation for RE [5]. In roving paradigms, RS was observed early with a decrease in N100 amplitude in response to pure sounds and then the RE occurs with an increase in P2 amplitude [5].

Neural adaptation has also been studied with magnetoencephalography (MEG), in visual [88] and auditory [67,89] modalities. Trying to localize the source of RS and RE, Recasens et al. demonstrated that the RS was generated in the Superior Temporal Gyrus (STG), the Middle Temporal Gyrus (MTG) and in insular regions whereas the RE sources were located in supratemporal and non-auditory regions (anterior region of the Insula and the Rolandic operculum) [67]. At the temporal level the two phenomena display different dynamic. Indeed, the RS has been identified as an early component (modulation of the early N1m, approximately at 90–150 ms) and the RE as a later component (sustained field, approximately at 230–270 ms) [67], confirming results of Costa-Faidella [5]. Moreover, these two components occur to have different implications. RS would correspond to the phenomenon of adaptation, whereas RE would rather correspond to anticipation of an expected event. Additionally, the dissociation between RS and RE has also been raised in fMRI studies demonstrating activation of different brain areas corresponding to the encoding of a new representation, and of brain regions involved in the retrieval of information [83,90,91]. Finally, the nature of the task and the required engagement of selective attention might influence the RS/RE balance [92]. Considering this modulation, RE would reflect selective attention in case of active oddball for example, whereas RS would be more likely observed in passive paradigms.

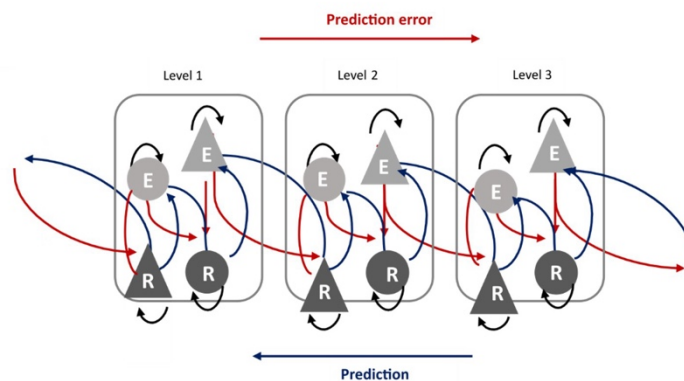
#### 2.4. RS and RE Explained by Predictive Coding

Another model that can explain RS and RE is the predictive coding model [93]. The predictive coding hypothesis is a perceptual inference hypothesis describing the brain as a hierarchically organized cortical system that constantly attempts to anticipate future events. For this, the brain would constantly learn the regularities in the sensory environment to build predictions about future sensory inputs. Comparisons between information provided by sensory inputs (bottom-up) and generated predictions (top-down) of futures sensory inputs are thus performed at each level of the system [93–95]. The predictive coding theory takes up the main concept of the Bayes' theory in which the percept (*posterior*)

is based on belief and knowledge (*prior*) and on sensory input (*likelihood*). According to the strength of the *prior* and of the *likelihood*, the accuracy of the percept will be modified.

According to this hypothesis, perception will be biased by the quality of sensory inputs and predictions. Indeed, the stronger these two parameters are, the stronger will be their influence on perception. In a stable environment, the brain can extract strong regularities, which results in a robust prediction and thus, to a heavy influence of this parameter on perception. However, in an unstable environment, extraction of regularities will become more difficult, resulting in low predictions and thus in a weak influence of this parameter on perception. In parallel, if the sensory input comes from a clear environment, the precision of the incoming information will be high and thus the influence on perception would be stronger. However, if the sensory input comes from a noisy environment, the precision of the information contained will be low and thus the influence of this parameter will be weak [96,97].

Moreover, sometimes sensory input and prediction are incongruent, resulting in the generation of a prediction error that leads to updating of the prediction, depending on the precision of this prediction error. Indeed, through the precision of the prediction error it can be determined if the system needs an update or if the prediction error is merely the result of low-quality sensory inputs or prediction, in which case it can be ignored [98,99]. The stronger the prediction (reflecting a strong regularity) is, the higher the precision of the prediction error will be, inducing prediction updating, which is, in turn, transmitted to lower areas [98,100] (Figure 3).



**Figure 3.** Predictive coding. In light grey: error unit; in dark grey: state unit. Triangle: pyramidal cells. Circle: bouquet cells. Red arrow: prediction error. Blue arrow: prediction. The levels correspond to hierarchical levels of cortical layers (adapted from “Attention, uncertainty, and free energy”, *Frontiers in Human Neuroscience*, 4, 215, Feldman and Friston (2010) [100]).

According to the predictive coding framework, RS, which reflects regularity encoding, would index a decrease in demand that occurs when expected and observed sensory information coincide (lower prediction error) and would thus reflect the increasing precision of the prediction [93]. Conversely, RE would index an increase in the prediction strength when expected and observed sensory information are the same [3,93,99]. MMN, previously described as an electrophysiological index of change detection, is considered to be a marker of prediction error and of updating of the prediction [101–103].

To conclude, the different explanations and models for adaptation and prediction to a repeated stimulus are summarized in Table 1.

**Table 1.** Summary of the processes underlying repetition suppression (RS) and repetition enhancement (RE) according to the different theories/models.

	<b>Repetition Suppression</b>	<b>Repetition Enhancement</b>
	Decrease in neural activity to repeated presentation of the same stimulus → <b>Reflects novelty preference</b>	Increase in the response as the number of repetitions of a stimulus increases. More marked for degraded stimuli, reflecting missing access to memory representation and construction of this representation. → <b>Reflects familiarity preference</b>
<b>Grill-Spector et al.'s Models</b> Grill-Spector et al., 2006 [23]	<b>The fatigue model:</b> decrease in the amplitude of firing of neurons responding to the stimulus, proportional to the initial response <b>The sharpening model:</b> decrease in the number of neurons that respond to the stimulus with repeated presentation	<b>The novel network formation:</b> presentation of a novel stimulus that is learned and its representation is established with the creation of a new neural network coded for this stimulus.
<b>The facilitation (or accumulation) model</b> James and Gauthier, 2006 [27]	Faster processing of the stimulus by the neurons involved, resulting in shorter latency or reaction time	Repeated exposure to stimulus with low visibility generally leads to increased perceptual performance.
<b>Predictive coding</b> Friston, 2005 [93]	Regularity encoding, displaying a decrease in the demand that occurs when expected and observed sensory information coincide. Reflects the increasing precision of prediction, the correct prediction of the upcoming stimulus	Increase in the prediction weight when expected and observed sensory information are the same.

Before discussing the possible implication of these neurophysiological processes in neurodevelopmental and psychiatric disorders, we summarize current knowledge regarding habituation, neural adaptation (RS and RE) and prediction in visual and auditory modalities with different methods (EEG, MEG, fMRI) through typical development.

### 3. Adaptation and Prediction through Typical Development

#### 3.1. Behavioral Habituation and Familiarization

In children, habituation is an essential process for the creation of routine in infancy that improve health and wellbeing in families [104,105]. Several studies have shown that habituation is already present from infants (3 to 6 months old) [106,107], children (8 to 9 years) [108] to adults [15]. Some of these studies used fixation time, also referred as familiarization paradigm, with for example a visual pairs comparison paradigm, consisting of the presentation of pairs of similar and different images [109]. With this type of paradigm, it has been possible to observe preference for novelty, reflected by an increase in fixation time for a novel stimulus relative to the previously repeated stimulus [107,110,111]. However, using this procedure it makes it unclear whether the observed novelty preference indirectly reflects habituation, or not.

Nevertheless, some studies in infants using familiarization paradigms (in which an initial habituation phase composed of a repeated presentation of a stimulus is followed by a test phase with the presentation of a new stimulus [112]) have shown no preference for novelty but a preference for familiarity during the test phase, reflected by an increase in looking toward the familiar stimulus [113]. This result would be dependent on duration of habituation phase. Indeed, if a familiar stimulus has not been well encoded and represented, then it will be preferred during the test phase. The quality of the internal representation of a repeated stimulus, the complexity of the task and the age of the participants,

would therefore influence the number of repetitions needed to reach a complete habituation [74,107,113]. Difference in habituation between adults and children could be the consequence of the effect of familiarity with the sentences used in this study [114]. The complexity of the stimulus to be encoded could also have an impact on the observed response [115]. In keeping with the idea that the RE is in place during the familiarization phase with a stimulation, in children a required number of repetitions was mandatory to eventually model a robust representation of it and produce a decrease in response, neural adaptation [63]. The familiarization phase with a new stimulus may be associated with RE and then the recognition of the familiar stimulus may be related to RS.

### 3.2. Neural Adaptation in Typical Development

In children, although behavioral habituation is present and efficient since the first months of life [106] and a need of routine is observed early in development [104,105], the underlying neural adaptation remains poorly explored. A description of the electrophysiology studies in this population is proposed in the current section.

There is no consensus on neural adaptation in children. Martineau et al. [116] observed RS in response to pure tones in children in an oddball. However, more recently, in a study comparing neural adaptation to electronic and human sounds, a decrease in the P1 amplitude was observed corresponding to RE, but only two repetitions were compared, suggesting that adaptation could take longer to implement [57]. Nevertheless, the presence of RE in children's response is consistent with the idea presented earlier that RE instigates formation of the memory trace and after it is well established, RS is observed. These observations make sense in the predictive coding framework, in which RE could reflect prediction after a first step corresponding to familiarization [117].

To our knowledge, no studies have focused on adaptation in children using fMRI. Studies did not yet focus on the process of neural adaptation in response to different category of sounds. Additionally, no study using a properly designed paradigm to measure cortical adaptation in children has been conducted. The developmental trajectory of the ability of neural adaptation thus remains to be characterized.

The studies on neural adaptation in typically developing children and adults are summarized in Table 2.

**Table 2.** Neural adaptation in typically developing adults and children. SOA: Stimulus Onset Asynchrony.

Study	Cortical/Sub-cortical	Modality	Population	Paradigm	Protocol	Stimuli Duration and SOA (ms)	Results
Baldeweg et al., 2004 [4]	Cortical	Auditory (pure tones)	20 adults	Roving (2, 6, 18 and 38 repetitions)	1 sequence	Stimuli: 25 for standard and 50 for deviant	<ul style="list-style-type: none"> <li>▪ <math>\nearrow</math>MMN with <math>\nearrow</math> number of repetitions</li> <li>▪ Detection of RP</li> </ul>
Haenschel et al., 2005 [66]	Cortical	Auditory (pure tones)	40 adults	Roving (trains of 2, 6 and 36 repetitions)	2 blocks with passive listening 2 blocks with active discrimination	Stimuli: 200 SOA: 500	<ul style="list-style-type: none"> <li>▪ <math>\nearrow</math>MMN with <math>\nearrow</math> number of repetitions</li> <li>▪ Detection of RP (36—2 repetitions)</li> </ul>
Ylinen and Huotilainen, 2007 [73]	Cortical	Auditory (synthesized vowels and vowels-like equivalents)	9 adults	Roving (3–4 repetitions 5–6 repetitions)	1 sequence of familiar stimuli (vowels) 1 sequence of unfamiliar stimuli (vowels-like)	Stimuli: 400 SOA: 700	<ul style="list-style-type: none"> <li>▪ <math>\emptyset</math> RP for both conditions</li> <li>▪ <math>\nearrow</math> N1 amplitude for familiar stimuli only</li> </ul>
Garrido et al., 2009 [118]	Cortical	Auditory (pure tones)	10 adults	Roving	1 sequence	Stimuli: 70 SOA: 570	<ul style="list-style-type: none"> <li>▪ <math>\nearrow</math>MMN with <math>\nearrow</math> number of repetitions</li> </ul>
Costa-Faidella et al., 2011 [56]	Cortical	Auditory (pure tones)	17 adults	Roving (trains of 3, 6 and 12 repetitions)	Predictable condition and unpredictable condition	Stimuli: 50 SOA: 708 (predictable) or 364–1062 (unpredictable)	<ul style="list-style-type: none"> <li>▪ Detection of RP for predictable condition</li> <li>▪ Only later part of RP (&gt;200 ms) observed in unpredictable condition</li> </ul>
Costa-Faidella et al., 2011 [5]	Cortical	Auditory (pure tones)	20 adults	Oddball (2, 6 and 12 repetitions)	1 sequence composed of 2 runs. Run 1: S1 repeated 2, 6 or 12	Stimuli: 40	<ul style="list-style-type: none"> <li>▪ <math>\nearrow</math> P2 amplitude with repetition</li> <li>▪ <math>\nearrow</math> MMN amplitude with repetition</li> <li>▪ <math>\searrow</math> N100 amplitude with repetition</li> </ul>

					times, followed by deviant S2. Run 2: S2 repeated 2, 6 or 12 times, followed by deviant S1		
Cooper et al., 2013 [68]	Cortical	Auditory (pure tones)	24 adults	Roving and oddball (4, 8 or 16 standards)	1 roving sequence 1 oddball sequence	Stimuli: 50 for standard and 100 for deviant	<ul style="list-style-type: none"> <li>▪ RP observed only for roving condition</li> <li>▪ <math>\emptyset</math> increase in MMN amplitude with repetition for both condition</li> </ul>
Recasens et al., 2015 [67]	Cortical	Auditory (pure tones)	13 adults	Roving (3, 12 or 24 repetitions)	2 runs: 198 trains of each length	Stimuli: 50 SOA: 500	<ul style="list-style-type: none"> <li>▪ Early N1m: RS from initial to late repetitions in supratemporal regions and in non-auditory regions as the precuneus</li> <li>▪ Late SF interval: RE from initial to late repetitions in the HG, the STG, the MTG and insular regions</li> </ul>
Gorina-Carreta et al., 2016 [119]	Subcortical	Auditory (consonant-vowel)	30 adults	Sequence of repeated stimuli	Predictable timing condition and unpredictable timing condition (8 blocks of 1001 repetitions per condition)	Stimuli: 170 SOA: 366 (predictable) or 183–549 (unpredictable)	<ul style="list-style-type: none"> <li>▪ <math>\searrow</math> FFR with repetition for both conditions</li> <li>▪ stronger <math>\searrow</math> FFR for predictable condition</li> </ul>
Pinheiro et al., 2017 [72]	Cortical	Auditory (human vocalizations)	23 adults	Oddball	4 blocks of 1050 standards and 150 deviants. Three different stimuli, standard or deviant, depending on the block: neutral, angry and happy	Stimuli 700 SOA: 1200	<ul style="list-style-type: none"> <li>▪ RP amplitude higher for happy vocalizations than angry and neutral and increased with repetitions</li> </ul>

McCleery et al., 2019 [70]	Cortical	Auditory (pure tones)	29 adults	Roving (3, 8 or 33 repetitions)	1 sequence composed of 5 blocks at two different times (2 weeks apart) Stimuli vary in pitch + duration	Stimuli: 50 or 100	<ul style="list-style-type: none"> <li>↗ MMN with repetition</li> <li>↗ amplitude of the last standard of a train (3, 8, 33) with repetition (called RP)</li> </ul>
Fryer et al., 2020 [120]	Cortical	Auditory (pure tones)	241 adults	Oddball	3 blocks consisting of a standard (85%), a deviant in duration (5%), a deviant in frequency (5%), and a deviant in duration and in frequency (5%)	SOA: 500 Stimuli standard in duration: 50 Deviant in duration: 100	<ul style="list-style-type: none"> <li>RP amplitude linearly increased with the repetition of the standard</li> </ul>
Jamal et al., 2020 [121]	Cortical	Auditory (pure tones) and visual (radial checkerboard)	22 children 7.1–12.8 years	Sequences of repeated stimulus	2 sequences (1 visual and 1 auditive) with 300 repetitions of the same stimulus	SOA: 1,116 Stimuli duration: 116	<ul style="list-style-type: none"> <li>Decrease in the ERP amplitude with repetitions for both type of stimuli</li> </ul>
Ruiz-Martínez et al., 2020 [57]	Cortical	Auditory (Electronic and human sounds)	15 children 5–11 years	Oddball	8 blocks (4/sound type) (deviant separated by at least 2 standards; each block begins with 10 standards)	Stimuli: 85 SOA: 685–885	<ul style="list-style-type: none"> <li>↘ P1 response (only two repetitions tested)</li> </ul>
Feuerriegel et al., 2021 [122]	Cortical	Visual (faces)	43 adults	Oddball	34 sequences: 6 faces presented, then oddball face identity (different identity than the base face).	Images were presented at a rate of 6 Hz	<ul style="list-style-type: none"> <li>Surprising compared to neutral: VMN amplitude is more negative in early and late time windows</li> <li>No VMN amplitude difference between expected and neutral conditions in early and late time windows</li> </ul>

---

						Blocks of 6 consecutive sequences were used with two oddball face identities possible, the proportion of both is known
Heurteloup et al., 2022 [71]	Cortical	Auditory (complex tones and human voice)	20 adults (18 to 30 years)	Roving (4, 8 or 16 repetitions)	2 roving sequences: one for vocal sounds and one for complex non-vocal sounds	SOA: 646 Stimuli duration: 300
						<ul style="list-style-type: none"> <li>▪ RP for both categories of sounds</li> <li>▪ Decrease in the ERP amplitude with repetitions for both type of stimuli, but faster for complex sounds compared to vocal sounds</li> </ul>

---

### 3.3. Prediction in Typical Development

Prediction in childhood and during development has not been as widely studied as in adulthood in the framework of the predictive coding theory. To obtain an idea of the prediction processes in this context during typical development, we will focus on the study of the prediction error electrophysiological correlate, i.e., the MMN [101–103]. Automatic detection of deviancy in a repetitive sequence has been studied in newborns [123,124], even in pre terms [125], in toddlers [126] and in children [127,128]. In these studies, a response to change has been observed in response to pitch or phonetic changes [123,124]. Additionally, the brain sources of this mismatch response reflecting the prediction appears to be different between 4 and 6 years old showing that the predictive brain is maturing during this time period [126]. In older children a mismatch response close to that of adults has been observed, with classic age-related differences in amplitudes and/or latencies [126,127]. However, although the characteristics of the response to sounds changes are not yet mature, the organization of the cortical response is already functional at the age of five [126–128]. The presence of prediction error since birth thus demonstrates that predictions are present early and effectively very early in development.

Additionally, as previously presented, RE is observed in children in familiarization and learning phases [63,117], which could reflect prediction and anticipation.

## 4. Adaptation and Prediction Are Altered in Neurodevelopmental and Psychiatric Disorders

What is known about adaptation and prediction in autism spectrum disorder (ASD), schizophrenia and attention-deficit/hyperactivity disorder (ADHD)? In these neurodevelopmental and psychiatric disorders, automatic detection of change has been widely studied through the measure of the MMN. This is true for ASD [129–135], for ADHD [136–139] and for schizophrenia [4,140–143]. In most of these studies, discrepancies in MMN amplitude and/or latency have been demonstrated in the clinical group compared to typically developing, but description of the underlying adaptation process is lacking. From a predictive coding point of view, description and thorough investigation of the prior adaptation process is essential to draw conclusions about difficulties in detection of change. As previously indicated, deficits in deviance detection could be explained by difficulties in underlying neural adaptation; understanding of these two phenomena in neurodevelopmental disorders is therefore important. In these disorders, in which difficulties in adaptation, prediction, and reaction to change have been demonstrated as a part of the symptomatology, a comprehension of the underlying neurophysiology is essential.

### 4.1. Autism Spectrum Disorder (ASD)

ASD is a neurodevelopmental disorder characterized by both social impairments and restrictive, repetitive interests reflecting a need of immutability [144]. One of the main characteristics of the ASD symptomatology is therefore resistance to change [145–147]. Two mechanisms could be involved: an habituation deficit [131] and/or an atypical change detection [133,134]. These adaptative difficulties in autistic people could be a consequence of their difficulties in predicting future events [148]. This idea has been developed in the theories of the Bayesian brain and predictive coding theory of ASD that postulate difficulties in predicting future items and in continually updating internal representations based on what has previously occurred [11,149,150]. According to Pellicano and Burr [11], the prior beliefs of autistic individuals are very different from those of a non-autistic group.

In autistic children a need for repetition and routine is observed, and has a positive influence on children leading to a facilitated adaptation [110,151]. Considering behavioral habituation, a reduction in response to repeated stimuli such as pure tones or vestibular stimulation has been observed in autistic children compared to neurotypical children [9,152]. This reduction of habituation could lead, as previously discussed, to resistance to

change and could be the consequence of atypical neural adaptation. In 1992, Martineau et al. showed a lack of neural adaptation to pure tones through electrophysiological measurements [116]. Other electrophysiological studies have reached similar conclusions with no adaptation or less adaptation in autistic patients than in typically developing children [57,153–155]. However, all these studies draw conclusions on neural adaptation via odd-ball paradigms that are not specifically designed to study this process and by comparing only a restricted number of repetitions [153]. Not much research has used roving paradigms to study the implementation of the encoding of regularity and its associated cue, the RP, to determine whether it is affected in autism. Font-Alaminos et al. [156], using a roving paradigm to study the subcortical adaptation in autism, observed an increase in the amplitude of the frequency-following response (FFR), corresponding to RE. This is consistent with the idea of a neural adaptation deficit in individuals with ASD who do not adapt to a repeated stimulus but remain in a familiarization phase. Same conclusion has been reached by Latinus et al. [157] in an fMRI on cognitive flexibility. The results of EEG studies on neural adaptation are reported in Table 3.

In ASD, a disruption of the Excitatory/Inhibitory (E/I) balance in brain has been hypothesized [158]. In view of the above presented results in animals about the implication of the GABA<sub>A</sub>-mediated inhibition in SSA, the possible imbalance of the E/I in autism [159,160] could partly explain the differences observed in neural adaptation. The level of GABA/Glu also correlates with the ability to make correct prediction in the framework of Bayesian learning, especially in a frontal region in which autistic adults have lower levels of glutamate [161].

Atypicalities reported in adaptation could reflect (or presume) the prediction deficit observed in autism and therefore hypersensitivity to the environment and intolerance to change [148].

Table 3. Neural adaptation in autism. SOA: Stimulus Onset Asynchrony.

Study	Cortical/Sub-cortical	Modality	Population	Paradigm	Protocol	Stimuli Duration and SOA (ms)	Results
Martineau et al., 1992 [116]	Cortical	Auditory (pure tones)	30 children 3–11 years	Oddball	1 sequence of 60 tones	Stimuli: 100	<ul style="list-style-type: none"> <li>No decrease in the AER with the increase in repetitions</li> </ul>
Guiraud et al., 2011 [155]	Cortical	Auditory (Pure tones)	35 infants with high risk for ASD and 21 infants with low risk for ASD 9 months old	Oddball	1 sequence (deviant always followed by two standards)	Stimuli: 100 SOA: 800	<p>High risk: less marked <math>\searrow</math> (less habituation) and less <math>\nearrow</math> response to deviant (reduced discrimination)</p> <p>Low risk:</p> <ul style="list-style-type: none"> <li><math>\searrow</math> in P150 amplitude with standards repetition (habituation)</li> <li><math>\nearrow</math> in response amplitude to deviant (discrimination)</li> </ul> <p>High risk:</p> <ul style="list-style-type: none"> <li>No <math>\searrow</math> in P150 amplitude with standards repetition</li> <li>No <math>\nearrow</math> in response amplitude to deviant</li> </ul>
Gonzalez-Gadea et al., 2015 [154]	Cortical	Auditory (pure tones)	16 children 8–15 years	Oddball	2 blocks of 220 sequences, 3 types of sequence: <ul style="list-style-type: none"> <li>Standard: 5 repetitions of the same tone</li> <li>Expected deviant: repetition of 4 identical tones, the fifth deviant tone is monaural</li> </ul>	Stimuli: 50 SOA: 200 Inter sequence interval: between 700 and 1000	<ul style="list-style-type: none"> <li>No difference in MMN amplitude compared to typically developing participants</li> <li>P3 larger for expected deviant compared to standard</li> <li>No difference in P3 in unexpected deviant compared to typically developing participants</li> <li>Higher P3 amplitude in expected than in unexpected condition</li> </ul>

					Unexpected deviant: repetition of 4 identical tones, the fifth, deviant, tone is interaural	
Kolesnik et al., 2019 [162]	Cortical	Auditory (standard: pure tones Deviants: 1 white noise and 1 pure tone)	116 children with high risk of autism (9.03 +/- 1.1 m—39.05 +/- 3.47 m)	Oddball	1 sequence with standard and deviants (paradigm designed by Guiraud et al., 2011) Measure the standard response after 1, 2 or 3 presentations	Stimuli: 100 Inter-trial interval: 700 <ul style="list-style-type: none"><li>▪ Reduced repetition suppression</li><li>▪ Increased phase locking</li></ul>
Font-Alaminos et al., 2020 [156]	Subcortical	Auditory (pure tones)	17 children (9.1 +/- 1.7 years)	Roving (8, 10 and 12 repetitions)	1 sequence composed of 9 blocks, each block is composed of 30 trains of either 8, 10 and 12 repetitions	Stimuli: 100 SOA: 333 <ul style="list-style-type: none"><li>▪ ↗ FFR amplitude with repetition compared to typically developing participants</li></ul>
Ruiz-Martinez et al., 2020 [57]	Cortical	Auditory (Electronic and human sounds)	16 ASD children 7–10 years	Oddball	8 blocks (4/sound type) (deviant separated by at least 2 standards; each block begins with 10 standard)	Stimuli: 85 SOA: 685–885 <ul style="list-style-type: none"><li>▪ Reduced <math>\sphericalangle</math> P1 response (reduced habituation)</li><li>▪ Reduced MMN amplitude (reduced discrimination)</li></ul>

Jamal et al., 2020 [121]	Cortical	Auditory (pure tones) and visual (radial checkerboard)	13 children 7.4–12.8 years	Sequences of repeated stimulus	2 sequences (1 visual and 1 auditory) with 300 repetitions of the same stimulus	SOA: 1,116 Stimuli duration: 116	<ul style="list-style-type: none"> <li>▪ Reduced adaptation for both type of stimulus</li> <li>▪ increase in the amplitude of the response</li> </ul>
Cary et al., 2023 [153]	Cortical	Auditory (pure tones)	13 children 12.81 ± 2.63 years	Oddball	1 sequence of 1000 trials (80% standard)	SOA: 600 Stimuli duration: 360	<ul style="list-style-type: none"> <li>▪ Reduced P1 adaptation (between the first and the second standard)</li> <li>▪ No difference between MMN amplitude with typically developing participants</li> </ul>

In the predictive coding framework in ASD, the difficulties observed in response to deviance could reflect the lack of power of *priors* [11] or impaired encoding of sensory information [149,163,164]. Destabilization of the balance between the weight of sensory inputs and the weight of priors would lead to a high rates of prediction error and thus to impairment in deviance detection [165]. In autism there is a reduced ability to predict events and this could lead to deficits in understanding surroundings, which could explain difficulties in coping and may explain the emphasis on sameness, sensory hypersensitivities, difficulties in interacting with dynamic objects as well as difficulties in social cognition [166]. These different theories and assumptions were presented in a recent literature review on prediction in autistic people [167]. However, they do not cover all the major symptoms of autism, such as hyposensitivity, which would benefit further exploration and refined conceptualization. Randeniya et al. suggested that the processes of sensory learning and adaptation are important to interpret prediction error atypicalities in autism [168] and should also be further investigated.

Neural adaptation deficits, manifest in EEG and in fMRI, appear to be correlated with behavioral symptoms in ASD. High social impairment and sensory processing difficulties are indeed associated with reduced brain adaptation [121,169]. Differences related to the nature of the stimulus to be encoded have also been observed, with reduced fMRI adaptation for faces, but not for objects, in autistic adults, associated with challenges in social communication [121]. In response to sounds, fMRI adaptation in autistic adults also appears reduced in comparison to a group of typically developing adults [170].

In conclusion, in ASD the process of habituation/adaptation appeared to be impaired, and this may partially explain hypersensitivity to change, and difficulties observed in detecting change. Further studies of these phenomena in autism, with more ecological stimuli, social for example, could make it possible to determine the possible implication of the context, and of the nature of the repeated stimuli in resistance to change. A few studies have compared adaptation to social and non-social information in autism and have revealed differences according to the nature of the stimuli, with less (or an absence of) adaptation for social stimuli in comparison to non-social stimuli [10,171]. In other words, reduced sensory adaptation in autism could contribute to hypersensitivities but also to difficulties in social communication and adaptation to the environment.

#### 4.2. Schizophrenia

Schizophrenia is a psychiatric disorder characterized by several symptoms such as disturbances in thought, perception and behavior, according to the DSM-5 [144]. Habituation also seems to be impacted in this disorder. It has been proposed that in schizophrenia there is an atypical interpretation of incoming input due to inappropriate creation and use of stored regularities [172–174]. We propose here a review of the current state of knowledge about habituation and adaptation in this disorder.

Regarding behavioral habituation in schizophrenia, impairment has been observed, with less marked habituation to repeated events (i.e., sounds, images) compared to typically developing [8,175,176]. The habituation deficit in schizophrenia may be related to the memory impairment observed in this disorder. fMRI studies have shown that the lack of habituation was associated with a lack of activation of the hippocampus, a brain region involved in memory, in response to repeated stimuli [177,178].

There is no consensus on neural adaptation in schizophrenia based on EEG studies. Atypical deviance detection has repeatedly been observed in schizophrenia patients compared to neurotypical, with reduced MMN [4,70,71,179–181]. Studies using the roving paradigm to explore regularity encoding and neural adaptation by measuring the RP have reached divergent conclusions. In some studies, a reduced RP was observed in schizophrenic patients compared to typically developing participants [4,121], while another study reported no difference in the RP between patients and neurotypical [70]. In an odd-ball study, RS on several electrophysiological components (N100, P50 and P2) was observed [180]. As presented above, neural adaptation can be established in patients with

schizophrenia, but not as efficiently as in controls, which could partially explain the deficit in habituation observed in schizophrenia. A summary of these different electrophysiological studies and their conclusions are reported in Table 4.

**Table 4.** Neural adaptation in schizophrenia. SOA: Stimulus Onset Asynchrony.

Study	Cortical/Sub-cortical	Modality	Population	Paradigm	Protocol	Stimuli Duration and SOA (ms)	Results
Baldeweg et al., 2004 [4]	Cortical	Auditory (pure tones)	28 adults	Roving (2, 6, 18 and 38 repetitions)	1 sequence	Stimuli: 25 for standard and 50 for deviant	<ul style="list-style-type: none"> <li>▪ Reduced MMN compared to typically developing participants                             <ul style="list-style-type: none"> <li>▪ Ø increase RP with repetition</li> <li>▪ Reduced RP compared to controls</li> </ul> </li> </ul>
Rentzsch et al., 2015 [182]	Cortical	Auditory (pure tones and click sound)	25 adults	Oddball and paired click paradigm	Oddball: standard and deviant stimuli differed in frequency Click: pair of identical stimuli repeated	Oddball: Pure tone duration: 80 ms Inter-stimulus interval: between 350 and 650 ms Click sound: SOA: 500 and inter-trial interval of 3400	<ul style="list-style-type: none"> <li>▪ Reduced MMN compared to typically developing participants                             <ul style="list-style-type: none"> <li>▪ No effect on RS for P50</li> <li>▪ N100 and P200 RS reduced</li> </ul> </li> </ul>
Coffman et al., 2017 [180]	Cortical	Auditory (pure tones)	26 adults	Oddball	<ul style="list-style-type: none"> <li>▪ 2 tasks:                                     <ul style="list-style-type: none"> <li>▪ RS task: 5 repeated similar tones</li> <li>▪ MMN task: 5 repeated tones and pitch or duration mismatch</li> </ul> </li> </ul>	Stimuli: 50 SOA: 330 Inter-group: 750 Stimuli duration mismatch: 100	<ul style="list-style-type: none"> <li>▪ RS in P50, N100 and P200 responses</li> <li>▪ Reduced MMN compared to control</li> </ul>
McCleery et al., 2019 [70]	Cortical	Auditory (pure tones)	43 adults	Roving (3, 8 or 33 repetitions)	1 sequence composed of 5 blocks at two different times (2 weeks apart). Pitch + duration stimuli variation	Stimuli: 50 or 100	<ul style="list-style-type: none"> <li>▪ ↗ MMN with repetition but reduced MMN compared to controls</li> <li>▪ ↗ amplitude of the last standard of a train (3, 8, 33) with repetition (called RP) as in typically developing participants</li> </ul>

Fryer et al., 2020 [120]	Cortical	Auditory (pure tones)	54 adults with Psychosis Risk syndrome—Conversion (PRS-C)	Oddball	3 blocks consisting of a standard (85%), a deviant in duration (5%), a deviant in frequency (5%), and a deviant in duration and in frequency (5%)	SOA: 50 Stimuli standard in duration: 50 Deviant in duration: 100	<ul style="list-style-type: none"> <li>RP amplitude smaller than typically developing participants</li> <li>Effect of the position in the train especially in position 3, 8–10 and 11+ in the train</li> </ul>
Koshiyama et al., 2020 [181]	Cortical	Auditory (pure tones)	25 adults	Oddball and many-standards paradigm	<p>Oddball: 2 sequences</p> <ul style="list-style-type: none"> <li>Duration: 2 tones with different durations</li> <li>Frequency: 2 tones with different frequencies</li> </ul> <p>Many-standards paradigm: 2 sequences</p> <ul style="list-style-type: none"> <li>Duration: 10 tones with different durations</li> <li>Frequencies: 10 tones with different frequencies</li> </ul>	<p>Stimulus standard: 50</p> <p>Oddball duration sequence → deviants 100</p> <p>Many-standards paradigm: between 10 and 225</p>	<ul style="list-style-type: none"> <li>Reduced MMN compared to typically developing participants</li> <li>No difference in adaptation between patients and typically developing participants</li> <li>Reduced deviant detection component of the MMN in patients</li> </ul>
Mazer et al., 2021 [183]	Cortical	Auditory (bird song, voice and pure tone)	26 adults	Roving	7 repetitions for trains for bird song and voice—target tone appears between train	<p>Complex sounds (bird songs and voice): 200</p> <p>Inter-stimulus interval: 1000</p> <p>Pure tone: 70</p>	<ul style="list-style-type: none"> <li>No decrease in P2 amplitude for bird songs</li> <li>Decrease in P2 amplitude for human voice</li> <li>No difference in N1 habituation between patients and typically developing participants</li> </ul>

Moreover, as in ASD, it has been argued that the predictive coding theory accounts for the underlying mechanisms involved in this altered MMN [142] since there is a reduced effect of the repetition on the amplitude of the MMN reflecting an impairment in prediction [184].

It remains complicated to draw conclusions regarding neural adaptation in schizophrenia. Indeed, there is still no consensus on fMRI adaptation in these patients. Some studies have found similar fMRI adaptation for repeated faces between typically developing participants and schizophrenic patients in specialized brain areas such as the fusiform face area (FFA) and occipital face area (OFA) [185,186]. However, in Williams et al.'s study [186], no RS was found in less specialized brain regions, with no decrease in the response amplitude with repetition in the primary visual cortex and in the hippocampus in comparison to neurotypical. Authors conclude that adaptation appears to be preserved in ultra-specialized brain areas in schizophrenia, but not in the more general areas responsible for encoding basic features of visual stimuli [186]. Furthermore, another study found smaller fMRI adaptation in patients with schizophrenia compared to typically developing participants in response to pictures of objects in the lateral occipital cortex (LOC), the brain area responsive to the category of visual stimuli [187]. This study represents a complementary finding to what was previously presented; adaptation at the brain level appears to be possible in specialized areas but appears to be impaired in more generic brain areas responsible for visual processing, reflecting alteration in the whole brain response but not in ultra-specialized areas in schizophrenia.

To summarize, in schizophrenia the habituation difficulties observed in behavior are not necessarily related to basic deficits in neural adaptation as observed in EEG and/or in fMRI. New theories about predictive coding in schizophrenia may help to draw conclusions about the mechanisms involved in these difficulties.

#### 4.3. Attention-Deficit/Hyperactivity Disorder (ADHD)

According to the DSM-5, ADHD is a neurodevelopmental disorder characterized by a persistent pattern of inattention and/or hyperactivity-impulsivity [144]. The deficits in attention could lead to adaptation and habituation impairment, because of difficulties in regularity encoding due to early inattention.

At the behavioral level, habituation has been shown to be impaired in ADHD in visual tasks [188,189]. However, the nature of the measure has an impact on the habituation ability, with some skin conductance studies revealing enhanced and accelerated habituation to sounds in startle reflex paradigms compared to typically developing adults [190,191]. This quicker habituation would induce a less efficient reinforcement of the memory trace and consequently a less sustainable habituation [192].

In ADHD there is no study focusing on neural regularity encoding. The possible adaptation and habituation deficit is mostly studied by observing changes in attention, through the P3, an index of attention orientation. P3 has been shown to be altered in ADHD, with no difference in amplitude between expected and unexpected conditions for example [154]. This attentional switch deficit could lead to an adaptation deficit. Indeed, in case of impairment in the regularity encoding and in adaptation (less expectation) to standard stimuli, the detection of change and novelty involved in attention switching become more difficult. Predictive coding theory could provide one explanation: in ADHD the observed defect in P3 could be related to too much involvement of sensory inputs at the expense of predictions [154,193].

Other studies have found no difference in MMN amplitude between neurotypical and individuals with ADHD [138,154], indicating that even if attentional switching is atypical in ADHD, deviance detection does not seem to be impaired. However, a meta-analysis on MMN in ADHD patients revealed that in these patients the amplitude of this index is reduced compared to typically developing children [194]. As in ASD and schizophrenia, change detection appears to be impaired. The joint action of change detection

and attentional switching deficits could therefore be indicative of a neural adaptation deficit, but to our knowledge no study has yet addressed this cascade of processes in ADHD.

### 5. Future Directions

Future study of neural adaptation throughout typical development and in neurodevelopmental disorders would be interesting, especially in response to more complex stimuli than pure tones or basic shapes. Indeed, the study of the combination of social cues on adaptation, especially in ASD, could allow for better understanding of the effects that have already been observed regarding change detection, and could provide a link with the symptomatic dyad, sameness and aloneness, initially described by Kanner [195] and Soukhareva [196]. Additionally, the study of neural adaptation in neurodevelopmental disorders will help answer the research question of whether there is a problem in change detection itself or a problem in subadjacent adaptation that is behind the MMN atypicalities. The response to this question would have a major impact on targeted behavioral intervention in these populations.

### 6. Conclusions

In summary, habituation and adaptation are important phenomena in the encoding of a steady context and in the detection of changes in the environment. These mechanisms are studied at different levels, but recently a more adapted electrophysiological cue has been measured: repetition positivity, which reflects regularity encoding and therefore neural adaptation in the framework of the roving paradigm. In the neurodevelopmental disorders discussed in this review, behavioral habituation seems to be impaired, but specifically designed studies of neural adaptation are not yet widespread. In future studies, it would be relevant to investigate these adaptation mechanisms using more complex stimuli that may contain social cues, for example in order to study their involvement in other components of the symptomatology of neurodevelopmental disorders, such as socio-communicative difficulties.

**Author Contributions:** A.M. and M.G. were involved in preparing writing and reviewing the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the INSERM (National Institute for Health and Medical Research) and the ANR (ANR-19-CE37-0022-01). The funding sources were not involved in the study design, data collection and analysis, or the writing of the report.

**Institutional review board statement:** Not applicable.

**Informed consent statement:** Not applicable.

**Data availability statement:** Please contact the corresponding author for data and experimental stimuli requests.

**Acknowledgments:** We would like to thank Laurie Tuller for her help with the revision of the manuscript.

**Conflicts of Interest:** All authors declare that they have no conflict of interest.

### References

1. Winkler, I.; Denham, S.L.; Nelken, I. Modeling the Auditory Scene: Predictive Regularity Representations and Perceptual Objects. *Trends Cogn. Sci.* **2009**, *13*, 532–540. <https://doi.org/10.1016/j.tics.2009.09.003>.
2. Andrews, T.J.; Ewbank, M.P. Distinct Representations for Facial Identity and Changeable Aspects of Faces in the Human Temporal Lobe. *Neuroimage* **2004**, *23*, 905–913. <https://doi.org/10.1016/j.neuroimage.2004.07.060>.
3. Summerfield, C.; Trittschuh, E.H.; Monti, J.M.; Mesulam, M.-M.; Egner, T. Neural Repetition Suppression Reflects Fulfilled Perceptual Expectations. *Nat. Neurosci.* **2008**, *11*, 1004–1006. <https://doi.org/10.1038/nn.2163>.
4. Baldeweg, T.; Klugman, A.; Gruzelier, J.; Hirsch, S.R. Mismatch Negativity Potentials and Cognitive Impairment in Schizophrenia. *Schizophr. Res.* **2004**, *69*, 203–217. <https://doi.org/10.1016/j.schres.2003.09.009>.

5. Costa-Faidella, J.; Grimm, S.; Slabu, L.; Díaz-Santaella, F.; Escera, C. Multiple Time Scales of Adaptation in the Auditory System as Revealed by Human Evoked Potentials: Adaptation in the Human Auditory System. *Psychophysiology* **2011**, *48*, 774–783. <https://doi.org/10.1111/j.1469-8986.2010.01144.x>.
6. Segaert, K.; Weber, K.; de Lange, F.P.; Petersson, K.M.; Hagoort, P. The Suppression of Repetition Enhancement: A Review of fMRI Studies. *Neuropsychologia* **2013**, *51*, 59–66. <https://doi.org/10.1016/j.neuropsychologia.2012.11.006>.
7. Parras, G.G.; Nieto-Diego, J.; Carbajal, G.V.; Valdés-Baizabal, C.; Escera, C.; Malmierca, M.S. Neurons along the Auditory Pathway Exhibit a Hierarchical Organization of Prediction Error. *Nat. Commun.* **2017**, *8*, 2148. <https://doi.org/10.1038/s41467-017-02038-6>.
8. Avery, S.N.; Armstrong, K.; Blackford, J.U.; Woodward, N.D.; Cohen, N.; Heckers, S. Impaired Relational Memory in the Early Stage of Psychosis. *Schizophr. Res.* **2019**, *212*, 113–120. <https://doi.org/10.1016/j.schres.2019.07.060>.
9. Ramaswami, M. Network Plasticity in Adaptive Filtering and Behavioral Habituation. *Neuron* **2014**, *82*, 1216–1229. <https://doi.org/10.1016/j.neuron.2014.04.035>.
10. Ewbank, M.P.; Pell, P.J.; Powell, T.E.; von dem Hagen, E.A.H.; Baron-Cohen, S.; Calder, A.J. Repetition Suppression and Memory for Faces Is Reduced in Adults with Autism Spectrum Conditions. *Cereb. Cortex* **2017**, *27*, 92–103. <https://doi.org/10.1093/cercor/bhw373>.
11. Pellicano, E.; Burr, D. When the World Becomes “Too Real”: A Bayesian Explanation of Autistic Perception. *Trends Cogn. Sci.* **2012**, *16*, 504–510. <https://doi.org/10.1016/j.tics.2012.08.009>.
12. Groves, P.M.; Thompson, R.F. Habituation: A Dual-Process Theory. *Psychol. Rev.* **1970**, *77*, 419–450. <https://doi.org/10.1037/h0029810>.
13. Thompson, R.F.; Spencer, W.A. Habituation: A Model Phenomenon for the Study of Neuronal Substrates of Behavior. *Psychol. Rev.* **1966**, *73*, 16–43. <https://doi.org/10.1037/h0022681>.
14. Hepper, P.G.; Leader, L.R. Fetal Habituation. *Fet. Matern. Med. Rev.* **1996**, *8*, 109–123. <https://doi.org/10.1017/S0965539500001534>.
15. Rankin, C.H.; Abrams, T.; Barry, R.J.; Bhatnagar, S.; Clayton, D.F.; Colombo, J.; Coppola, G.; Geyer, M.A.; Glanzman, D.L.; Marsland, S.; et al. Habituation Revisited: An Updated and Revised Description of the Behavioral Characteristics of Habituation. *Neurobiol. Learn. Mem.* **2009**, *92*, 135–138. <https://doi.org/10.1016/j.nlm.2008.09.012>.
16. Giles, A.C.; Rankin, C.H. Behavioral and Genetic Characterization of Habituation Using *Caenorhabditis Elegans*. *Neurobiol. Learn. Mem.* **2009**, *92*, 139–146. <https://doi.org/10.1016/j.nlm.2008.08.004>.
17. Bronstein, P.M.; Neiman, H.; Wolkoff, F.D.; Levine, M.J. The Development of Habituation in the Rat. *Anim. Learn. Behav.* **1974**, *2*, 92–96. <https://doi.org/10.3758/BF03199129>.
18. Pinsker, H.; Kupfermann, I.; Castellucci, V.; Kandel, E. Habituation and Dishabituation of the Gill-Withdrawal Reflex in Aplysia. *Science* **1970**, *167*, 1740–1742. <https://doi.org/10.1126/science.167.3926.1740>.
19. Bell, R.; Röer, J.P.; Dentale, S.; Buchner, A. Habituation of the Irrelevant Sound Effect: Evidence for an Attentional Theory of Short-Term Memory Disruption. *J. Exp. Psychol. Learn. Mem. Cogn.* **2012**, *38*, 1542–1557. <https://doi.org/10.1037/a0028459>.
20. Ornitz, E.M.; Guthrie, D. Long-Term Habituation and Sensitization of the Acoustic Startle Response in the Normal Adult Human. *Psychophysiology* **1989**, *26*, 166–173. <https://doi.org/10.1111/j.1469-8986.1989.tb03149.x>.
21. Pellegrino, R.; Sinding, C.; de Wijk, R.A.; Hummel, T. Habituation and Adaptation to Odors in Humans. *Physiol. Behav.* **2017**, *177*, 13–19. <https://doi.org/10.1016/j.physbeh.2017.04.006>.
22. Desimone, R. Neural Mechanisms for Visual Memory and Their Role in Attention. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 13494–9. <https://doi.org/10.1073/pnas.93.24.13494>.
23. Grill-Spector, K.; Henson, R.; Martin, A. Repetition and the Brain: Neural Models of Stimulus-Specific Effects. *Trends Cogn. Sci.* **2006**, *10*, 14–23. <https://doi.org/10.1016/j.tics.2005.11.006>.
24. Avidan, G.; Hasson, U.; Hendler, T.; Zohary, E.; Malach, R. Analysis of the Neuronal Selectivity Underlying Low fMRI Signals. *Curr. Biol.* **2002**, *12*, 964–972. [https://doi.org/10.1016/s0960-9822\(02\)00872-2](https://doi.org/10.1016/s0960-9822(02)00872-2).
25. Grill-Spector, K.; Malach, R. fMR-Adaptation: A Tool for Studying the Functional Properties of Human Cortical Neurons. *Acta Psychol.* **2001**, *107*, 293–321. [https://doi.org/10.1016/S0001-6918\(01\)00019-1](https://doi.org/10.1016/S0001-6918(01)00019-1).
26. Wiggs, C.L.; Martin, A. Properties and Mechanisms of Perceptual Priming. *Curr. Opin. Neurobiol.* **1998**, *8*, 227–233. [https://doi.org/10.1016/S0959-4388\(98\)80144-X](https://doi.org/10.1016/S0959-4388(98)80144-X).
27. James, T.W.; Gauthier, I. Repetition-Induced Changes in BOLD Response Reflect Accumulation of Neural Activity. *Hum. Brain Mapp.* **2006**, *27*, 37–46. <https://doi.org/10.1002/hbm.20165>.
28. Ulanovsky, N.; Las, L.; Farkas, D.; Nelken, I. Multiple time scales of adaptation in auditory cortex neurons. *J. Neurosci.* **2004**, *24*, 10440–10453. <https://doi.org/10.1523/JNEUROSCI.1905-04.2004>.
29. Baylis, G.C.; Rolls, E.T. Responses of Neurons in the Inferior Temporal Cortex in Short Term and Serial Recognition Memory Tasks. *Exp. Brain Res.* **1987**, *65*, 614–622. <https://doi.org/10.1007/BF00235984>.
30. De Baene, W.; Vogels, R. Effects of Adaptation on the Stimulus Selectivity of Macaque Inferior Temporal Spiking Activity and Local Field Potentials. *Cereb. Cortex* **2010**, *20*, 2145–2165. <https://doi.org/10.1093/cercor/bhp277>.
31. Dragoi, V.; Sharma, J.; Sur, M. Adaptation-Induced Plasticity of Orientation Tuning in Adult Visual Cortex. *Neuron* **2000**, *28*, 287–298. [https://doi.org/10.1016/s0896-6273\(00\)00103-3](https://doi.org/10.1016/s0896-6273(00)00103-3).
32. Kaliukhovich, D.A.; Vogels, R. Stimulus Repetition Probability Does Not Affect Repetition Suppression in Macaque Inferior Temporal Cortex. *Cereb. Cortex* **2011**, *21*, 1547–1558. <https://doi.org/10.1093/cercor/bhq207>.

33. Kaliukhovich, D.A.; Vogels, R. Stimulus Repetition Affects Both Strength and Synchrony of Macaque Inferior Temporal Cortical Activity. *J. Neurophysiol.* **2012**, *107*, 3509–3527. <https://doi.org/10.1152/jn.00059.2012>.
34. Kaliukhovich, D.A.; Vogels, R. Neurons in Macaque Inferior Temporal Cortex Show No Surprise Response to Deviants in Visual Oddball Sequences. *J. Neurosci.* **2014**, *34*, 12801–12815. <https://doi.org/10.1523/JNEUROSCI.2154-14.2014>.
35. Liu, Y.; Murray, S.O.; Jagadeesh, B. Time Course and Stimulus Dependence of Repetition-Induced Response Suppression in Inferotemporal Cortex. *J. Neurophysiol.* **2009**, *101*, 418–436. <https://doi.org/10.1152/jn.90960.2008>.
36. Miller, E.K.; Li, L.; Desimone, R. A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* **1991**, *254*, 1377–1379. <https://doi.org/10.1126/science.1962197>.
37. Müller, J.R.; Metha, A.B.; Krauskopf, J.; Lennie, P. Rapid Adaptation in Visual Cortex to the Structure of Images. *Science* **1999**, *285*, 1405–1408. <https://doi.org/10.1126/science.285.5432.1405>.
38. Ringo, J.L. Stimulus Specific Adaptation in Inferior Temporal and Medial Temporal Cortex of the Monkey. *Behav. Brain Res.* **1996**, *76*, 191–197. [https://doi.org/10.1016/0166-4328\(95\)00197-2](https://doi.org/10.1016/0166-4328(95)00197-2).
39. Sawamura, H.; Orban, G.A.; Vogels, R. Selectivity of Neuronal Adaptation Does Not Match Response Selectivity: A Single-Cell Study of the fMRI Adaptation Paradigm. *Neuron* **2006**, *49*, 307–318. <https://doi.org/10.1016/j.neuron.2005.11.028>.
40. Vinken, K.; Vogels, R. Adaptation Can Explain Evidence for Encoding of Probabilistic Information in Macaque Inferior Temporal Cortex. *Curr. Biol.* **2017**, *27*, R1210–R1212. <https://doi.org/10.1016/j.cub.2017.09.018>.
41. Kuravi, P.; Vogels, R. GABAergic and Cholinergic Modulation of Repetition Suppression in Inferior Temporal Cortex. *Sci. Rep.* **2018**, *8*, 13160. <https://doi.org/10.1038/s41598-018-31515-1>.
42. Ulanovsky, N.; Las, L.; Nelken, I. Processing of Low-Probability Sounds by Cortical Neurons. *Nat. Neurosci.* **2003**, *6*, 391–398. <https://doi.org/10.1038/nn1032>.
43. Fishman, Y.I.; Steinschneider, M. Searching for the Mismatch Negativity in Primary Auditory Cortex of the Awake Monkey: Deviance Detection or Stimulus Specific Adaptation? *J. Neurosci.* **2012**, *32*, 15747–15758. <https://doi.org/10.1523/JNEUROSCI.2835-12.2012>.
44. Nelken, I.; Yaron, A.; Polterovich, A.; Hershenhoren, I. Stimulus-Specific Adaptation beyond Pure Tones. *Adv. Exp. Med. Biol.* **2013**, *787*, 411–418. [https://doi.org/10.1007/978-1-4614-1590-9\\_45](https://doi.org/10.1007/978-1-4614-1590-9_45).
45. Taaseh, N.; Yaron, A.; Nelken, I. Stimulus-Specific Adaptation and Deviance Detection in the Rat Auditory Cortex. *PLoS ONE* **2011**, *6*, e23369. <https://doi.org/10.1371/journal.pone.0023369>.
46. von der Behrens, W.; Bäuerle, P.; Kössl, M.; Gaese, B.H. Correlating Stimulus-Specific Adaptation of Cortical Neurons and Local Field Potentials in the Awake Rat. *J. Neurosci.* **2009**, *29*, 13837–13849. <https://doi.org/10.1523/JNEUROSCI.3475-09.2009>.
47. Yaron, A.; Hershenhoren, I.; Nelken, I. Sensitivity to Complex Statistical Regularities in Rat Auditory Cortex. *Neuron* **2012**, *76*, 603–615. <https://doi.org/10.1016/j.neuron.2012.08.025>.
48. Anderson, L.A.; Malmierca, M.S. The Effect of Auditory Cortex Deactivation on Stimulus-Specific Adaptation in the Inferior Colliculus of the Rat. *Eur. J. Neurosci.* **2013**, *37*, 52–62. <https://doi.org/10.1111/ejn.12018>.
49. Ayala, Y.A.; Malmierca, M.S. Stimulus-Specific Adaptation and Deviance Detection in the Inferior Colliculus. *Front. Neural Circuits* **2012**, *6*, 89. <https://doi.org/10.3389/fncir.2012.00089>.
50. Pérez-González, D.; Malmierca, M.S.; Covey, E. Novelty Detector Neurons in the Mammalian Auditory Midbrain. *Eur. J. Neurosci.* **2005**, *22*, 2879–2885. <https://doi.org/10.1111/j.1460-9568.2005.04472.x>.
51. Pérez-González, D.; Malmierca, M.S. Variability of the Time Course of Stimulus-Specific Adaptation in the Inferior Colliculus. *Front. Neural Circuits* **2012**, *6*, 107. <https://doi.org/10.3389/fncir.2012.00107>.
52. Anderson, L.A.; Christianson, G.B.; Linden, J.F. Stimulus-Specific Adaptation Occurs in the Auditory Thalamus. *J. Neurosci.* **2009**, *29*, 7359–7363. <https://doi.org/10.1523/JNEUROSCI.0793-09.2009>.
53. Antunes, F.M.; Nelken, I.; Covey, E.; Malmierca, M.S. Stimulus-Specific Adaptation in the Auditory Thalamus of the Anesthetized Rat. *PLoS ONE* **2010**, *5*, e14071. <https://doi.org/10.1371/journal.pone.0014071>.
54. Antunes, F.M.; Malmierca, M.S. Effect of Auditory Cortex Deactivation on Stimulus-Specific Adaptation in the Medial Geniculate Body. *J. Neurosci.* **2011**, *31*, 17306–17316. <https://doi.org/10.1523/JNEUROSCI.1915-11.2011>.
55. Duque, D.; Malmierca, M.S.; Caspary, D.M. Modulation of Stimulus-Specific Adaptation by GABA(A) Receptor Activation or Blockade in the Medial Geniculate Body of the Anaesthetized Rat. *J. Physiol.* **2014**, *592*, 729–743. <https://doi.org/10.1113/jphysiol.2013.261941>.
56. Costa-Faidella, J.; Baldeweg, T.; Grimm, S.; Escera, C. Interactions between “What” and “When” in the Auditory System: Temporal Predictability Enhances Repetition Suppression. *J. Neurosci.* **2011**, *31*, 18590–18597. <https://doi.org/10.1523/JNEUROSCI.2599-11.2011>.
57. Ruiz-Martínez, F.J.; Rodríguez-Martínez, E.I.; Wilson, C.E.; Yau, S.; Saldaña, D.; Gómez, C.M. Impaired P1 Habituation and Mismatch Negativity in Children with Autism Spectrum Disorder. *J. Autism Dev. Disord.* **2020**, *50*, 603–616. <https://doi.org/10.1007/s10803-019-04299-0>.
58. Sayres, R.; Grill-Spector, K. Object-Selective Cortex Exhibits Performance-Independent Repetition Suppression. *J. Neurophysiol.* **2006**, *95*, 995–1007. <https://doi.org/10.1152/jn.00500.2005>.
59. Snyder, K.A.; Keil, A. Repetition Suppression of Induced Gamma Activity Predicts Enhanced Orienting toward a Novel Stimulus in 6-Month-Old Infants. *J. Cogn. Neurosci.* **2008**, *20*, 2137–2152. <https://doi.org/10.1162/jocn.2008.20149>.
60. Näätänen, R.; Gaillard, A.W.; Mäntysalo, S. Early Selective-Attention Effect on Evoked Potential Reinterpreted. *Acta Psychol.* **1978**, *42*, 313–329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9).

61. Näätänen, R.; Paavilainen, P.; Rinne, T.; Alho, K. The Mismatch Negativity (MMN) in Basic Research of Central Auditory Processing: A Review. *Clin. Neurophysiol.* **2007**, *118*, 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>.
62. Winkler, I.; Schröger, E.; Cowan, N. The Role of Large-Scale Memory Organization in the Mismatch Negativity Event-Related Brain Potential. *J. Cogn. Neurosci.* **2001**, *13*, 59–71. <https://doi.org/10.1162/089892901564171>.
63. Grossmann, T.; Gliga, T.; Johnson, M.H.; Mareschal, D. The Neural Basis of Perceptual Category Learning in Human Infants. *J. Cogn. Neurosci.* **2009**, *21*, 2276–2286. <https://doi.org/10.1162/jocn.2009.21188>.
64. Baldeweg, T.; Williams, J.D.; Gruzelier, J.H. Differential Changes in Frontal and Sub-Temporal Components of Mismatch Negativity. *Int. J. Psychophysiol.* **1999**, *33*, 143–148. [https://doi.org/10.1016/s0167-8760\(99\)00026-4](https://doi.org/10.1016/s0167-8760(99)00026-4).
65. Cowan, N.; Winkler, I.; Teder, W.; Naatanen, R. Memory Prerequisites of Mismatch Negativity in the Auditory Event-Related Potential (ERP). *J. Exp. Psychol. Learn. Mem. Cogn.* **1993**, *13*, 909–921. <https://doi.org/10.1037//0278-7393.19.4.909>.
66. Haenschel, C.; Vernon, D.J.; Dwivedi, P.; Gruzelier, J.H.; Baldeweg, T. Event-Related Brain Potential Correlates of Human Auditory Sensory Memory-Trace Formation. *J. Neurosci.* **2005**, *25*, 10494–10501. <https://doi.org/10.1523/JNEUROSCI.1227-05.2005>.
67. Recasens, M.; Leung, S.; Grimm, S.; Nowak, R.; Escera, C. Repetition Suppression and Repetition Enhancement Underlie Auditory Memory-Trace Formation in the Human Brain: An MEG Study. *NeuroImage* **2015**, *108*, 75–86. <https://doi.org/10.1016/j.neuroimage.2014.12.031>.
68. Cooper, R.J.; Atkinson, R.J.; Clark, R.A.; Michie, P.T. Event-Related Potentials Reveal Modelling of Auditory Repetition in the Brain. *Int. J. Psychophysiol.* **2013**, *88*, 74–81. <https://doi.org/10.1016/j.ijpsycho.2013.02.003>.
69. Garrido, M.I.; Kilner, J.M.; Stephan, K.E.; Friston, K.J. The Mismatch Negativity: A Review of Underlying Mechanisms. *Clin. Neurophysiol.* **2009**, *11*, 453–463. <https://doi.org/10.1016/j.clinph.2008.11.029>.
70. McCleery, A.; Mathalon, D.H.; Wynn, J.K.; Roach, B.J.; Helleman, G.S.; Marder, S.R.; Green, M.F. Parsing Components of Auditory Predictive Coding in Schizophrenia Using a Roving Standard Mismatch Negativity Paradigm. *Psychol. Med.* **2019**, *49*, 1195–1206. <https://doi.org/10.1017/S0033291718004087>.
71. Heurteloup, C.; Merchie, A.; Roux, S.; Bonnet-Brilhault, F.; Escera, C.; Gomot, M. Neural Repetition Suppression to Vocal and Non-Vocal Sounds. *Cortex* **2022**, *148*, 1–13. <https://doi.org/10.1016/j.cortex.2021.11.020>.
72. Pinheiro, A.P.; Barros, C.; Vasconcelos, M.; Obermeier, C.; Kotz, S.A. Is Laughter a Better Vocal Change Detector than a Growl? *Cortex* **2017**, *92*, 233–248. <https://doi.org/10.1016/j.cortex.2017.03.018>.
73. Ylinen, S.; Huotilainen, M. Is There a Direct Neural Correlate for Memory-Trace Formation in Audition? *NeuroReport* **2007**, *18*, 1281–1284. <https://doi.org/10.1097/WNR.0b013e32826fb38a>.
74. Turk-Browne, N.; Yi, D.-J.; Leber, A.; Chun, M. Visual Quality Determines the Direction of Neural Repetition Effects. *Cereb. Cortex* **2007**, *17*, 425–433. <https://doi.org/10.1093/cercor/bhj159>.
75. Doeller, C.F.; Opitz, B.; Mecklinger, A.; Krick, C.; Reith, W.; Schröger, E. Prefrontal Cortex Involvement in Preattentive Auditory Deviance Detection: Neuroimaging and Electrophysiological Evidence. *NeuroImage* **2003**, *20*, 1270–1282. [https://doi.org/10.1016/S1053-8119\(03\)00389-6](https://doi.org/10.1016/S1053-8119(03)00389-6).
76. Opitz, B.; Rinne, T.; Mecklinger, A.; von Cramon, D.Y.; Schröger, E. Differential Contribution of Frontal and Temporal Cortices to Auditory Change Detection: fMRI and ERP Results. *NeuroImage* **2002**, *15*, 167–174. <https://doi.org/10.1006/nimg.2001.0970>.
77. Opitz, B.; Schröger, E.; Cramon, D.Y.V. Sensory and Cognitive Mechanisms for Preattentive Change Detection in Auditory Cortex. *Eur. J. Neurosci.* **2005**, *21*, 531–535. <https://doi.org/10.1111/j.1460-9568.2005.03839.x>.
78. Eger, E.; Henson, R.N.A.; Driver, J.; Dolan, R.J. BOLD Repetition Decreases in Object-Responsive Ventral Visual Areas Depend on Spatial Attention. *J. Neurophysiol.* **2004**, *92*, 1241–1247. <https://doi.org/10.1152/jn.00206.2004>.
79. Murray, S.O.; Wojciulik, E. Attention Increases Neural Selectivity in the Human Lateral Occipital Complex. *Nat. Neurosci.* **2003**, *7*, 70–74. <https://doi.org/10.1038/nn1161>.
80. Gomot, M.; Bernard, F.A.; Davis, M.H.; Belmonte, M.K.; Ashwin, E.T.; Baron-Cohen, S. Change Detection in Children with Autism: An Auditory Event-Related fMRI Study. *NeuroImage* **2006**, *29*, 475–484. <https://doi.org/10.1016/j.neuroimage.2005.07.027>.
81. Sabri, M.; Kareken, D.A.; Dzemidzic, M.; Lowe, M.J.; Melara, R.D. Neural Correlates of Auditory Sensory Memory and Automatic Change Detection. *NeuroImage* **2004**, *21*, 69–74. <https://doi.org/10.1016/j.neuroimage.2003.08.033>.
82. Schall, U.; Johnston, P.; Todd, J.; Ward, P.B.; Michie, P.T. Functional Neuroanatomy of Auditory Mismatch Processing: An Event-Related fMRI Study of Duration-Deviant Oddballs. *NeuroImage* **2003**, *20*, 729–736. [https://doi.org/10.1016/S1053-8119\(03\)00398-7](https://doi.org/10.1016/S1053-8119(03)00398-7).
83. Cacciaglia, R.; Costa-Faidella, J.; Zarnowiec, K.; Grimm, S.; Escera, C. Auditory Predictions Shape the Neural Responses to Stimulus Repetition and Sensory Change. *NeuroImage* **2019**, *186*, 200–210. <https://doi.org/10.1016/j.neuroimage.2018.11.007>.
84. Sokolov, E.N.; Sokolov, Y.N.; Sokolov, E.N.; Waydenfeld, S.W.; Worters, R.; Clarke, A.D.B. *Perception and the Conditioned Reflex*; Pergamon Press: Oxford, UK, 1963.
85. Vogels, R. Sources of Adaptation of Inferior Temporal Cortical Responses. *Cortex* **2016**, *80*, 185–195. <https://doi.org/10.1016/j.cortex.2015.08.024>.
86. Peel, H.J.; Chouinard, P.A. fMRI Form Adaptation and Size Repetition Enhancement in Different Subdivisions of the Lateral Occipital Complex. *Cortex* **2022**, *154*, 135–148. <https://doi.org/10.1016/j.cortex.2022.04.020>.
87. Henson, R.; Shallice, T.; Dolan, R. Neuroimaging Evidence for Dissociable Forms of Repetition Priming. *Science* **2000**, *287*, 1269–1272. <https://doi.org/10.1126/science.287.5456.1269>.
88. Ishai, A.; Bickle, P.C.; Ungerleider, L.G. Temporal Dynamics of Face Repetition Suppression. *Brain Res. Bull.* **2006**, *70*, 289–295. <https://doi.org/10.1016/j.brainresbull.2006.06.002>.

89. Todorovic, A.; de Lange, F.P. Repetition Suppression and Expectation Suppression Are Dissociable in Time in Early Auditory Evoked Fields. *J. Neurosci.* **2012**, *32*, 13389–13395. <https://doi.org/10.1523/JNEUROSCI.2227-12.2012>.
90. van Turennout, M. Modulation of Neural Activity during Object Naming: Effects of Time and Practice. *Cereb. Cortex* **2003**, *13*, 381–391. <https://doi.org/10.1093/cercor/13.4.381>.
91. van Turennout, M.; Ellmore, T.; Martin, A. Long-Lasting Cortical Plasticity in the Object Naming System. *Nat. Neurosci.* **2000**, *3*, 1329–1334. <https://doi.org/10.1038/81873>.
92. Lange, K. The Reduced N1 to Self-Generated Tones: An Effect of Temporal Predictability?: N1 Reduction and Temporal Predictability. *Psychophysiology* **2011**, *48*, 1088–1095. <https://doi.org/10.1111/j.1469-8986.2010.01174.x>.
93. Friston, K. A Theory of Cortical Responses. *Phil. Trans. R. Soc.* **2005**, *360*, 815–836. <https://doi.org/10.1098/rstb.2005.1622>.
94. Grotheer, M.; Kovács, G. Can Predictive Coding Explain Repetition Suppression? *Cortex* **2016**, *80*, 113–124. <https://doi.org/10.1016/j.cortex.2015.11.027>.
95. Wacongne, C.; Labyt, E.; van Wassenhove, V.; Bekinschtein, T.; Naccache, L.; Dehaene, S. Evidence for a Hierarchy of Predictions and Prediction Errors in Human Cortex. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20754–20759. <https://doi.org/10.1073/pnas.1117807108>.
96. Tassinari, H.; Hudson, T.E.; Landy, M.S. Combining Priors and Noisy Visual Cues in a Rapid Pointing Task. *J. Neurosci.* **2006**, *26*, 10154–10163. <https://doi.org/10.1523/JNEUROSCI.2779-06.2006>.
97. Tong, J.; Ngo, V.; Goldreich, D. Tactile Length Contraction as Bayesian Inference. *J. Neurophysiol.* **2016**, *116*, 369–379. <https://doi.org/10.1152/jn.00029.2016>.
98. Clark, A. Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science. *Behav. Brain Sci.* **2013**, *36*, 181–204. <https://doi.org/10.1017/S0140525X12000477>.
99. Friston, K. The Free-Energy Principle: A Unified Brain Theory? *Nat. Rev. Neurosci.* **2010**, *11*, 127–138. <https://doi.org/10.1038/nrn2787>.
100. Feldman, H.; Friston, K.J. Attention, Uncertainty, and Free-Energy. *Front. Hum. Neurosci.* **2010**, *4*, 215. <https://doi.org/10.3389/fnhum.2010.00215>.
101. Bendixen, A.; SanMiguel, I.; Schröger, E. Early Electrophysiological Indicators for Predictive Processing in Audition: A Review. *Int. J. Psychophysiol.* **2012**, *83*, 120–131. <https://doi.org/10.1016/j.ijpsycho.2011.08.003>.
102. Winkler, I. Interpreting the Mismatch Negativity. *J. Psychophysiol.* **2007**, *21*, 147–163. <https://doi.org/10.1027/0269-8803.21.34.147>.
103. Winkler, I.; Czigler, I. Mismatch Negativity: Deviance Detection or the Maintenance of the “Standard”. *Neuroreport* **1998**, *9*, 3809–3813. <https://doi.org/10.1097/00001756-199812010-00008>.
104. Mindell, J.A.; Williamson, A.A. Benefits of a Bedtime Routine in Young Children: Sleep, Development, and Beyond. *Sleep Med. Rev.* **2018**, *40*, 93–108. <https://doi.org/10.1016/j.smrv.2017.10.007>.
105. Rodger, S.; Umaibalan, V. The Routines and Rituals of Families of Typically Developing Children Compared with Families of Children with Autism Spectrum Disorder: An Exploratory Study. *Br. J. Occup. Ther.* **2011**, *74*, 20–26. <https://doi.org/10.4276/030802211X12947686093567>.
106. Nakano, T.; Watanabe, H.; Homae, F.; Taga, G. Prefrontal Cortical Involvement in Young Infants’ Analysis of Novelty. *Cereb. Cortex* **2009**, *19*, 455–463. <https://doi.org/10.1093/cercor/bhn096>.
107. Turk-Browne, N.B.; Scholl, B.J.; Chun, M.M. Babies and Brains: Habituation in Infant Cognition and Functional Neuroimaging. *Front. Hum. Neurosci.* **2008**, *2*, 16. <https://doi.org/10.3389/neuro.09.016.2008>.
108. Kulig, J.W.; Tighe, T.J. Habituation in Children within a Behavior Suppression Paradigm. *J. Exp. Child Psychol.* **1981**, *32*, 425–442. [https://doi.org/10.1016/0022-0965\(81\)90106-5](https://doi.org/10.1016/0022-0965(81)90106-5).
109. Nguyen, L.Y.; Spehar, B. Visual Adaptation to Natural Scene Statistics and Visual Preference. *Vis. Res.* **2021**, *180*, 87–95. <https://doi.org/10.1016/j.visres.2020.11.011>.
110. Fantz, R.L. visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science* **1964**, *146*, 668–670. <https://doi.org/10.1126/science.146.3644.668>.
111. Nordt, M.; Hoehl, S.; Weigelt, S. The Use of Repetition Suppression Paradigms in Developmental Cognitive Neuroscience. *Cortex* **2016**, *80*, 61–75. <https://doi.org/10.1016/j.cortex.2016.04.002>.
112. Fagan, J.F., 3rd. Infants’ Recognition Memory for Faces. *J. Exp. Child Psychol.* **1972**, *14*, 453–476. [https://doi.org/10.1016/0022-0965\(72\)90065-3](https://doi.org/10.1016/0022-0965(72)90065-3).
113. Hunter, M.A.; Ames, E.W. A Multifactor Model of Infant Preferences for Novel and Familiar Stimuli. In *Advances in Infancy Research*; Ablex Publishing: Westport, CT, USA, 1988; Volume 5, pp. 69–95, ISBN 0-89391-378-2.
114. Dehaene-Lambertz, G.; Dehaene, S.; Anton, J.-L.; Campagne, A.; Ciuciu, P.; Dehaene, G.P.; Denghien, I.; Jobert, A.; LeBihan, D.; Sigman, M.; et al. Functional Segregation of Cortical Language Areas by Sentence Repetition. *Hum. Brain Mapp.* **2006**, *27*, 360–371. <https://doi.org/10.1002/hbm.20250>.
115. Bouchon, C.; Nazzi, T.; Gervain, J. Hemispheric Asymmetries in Repetition Enhancement and Suppression Effects in the Newborn Brain. *PLoS ONE* **2015**, *10*, e0140160. <https://doi.org/10.1371/journal.pone.0140160>.
116. Martineau, J.; Roux, S.; Garreau, B.; Adrien, J.L.; Lelord, G. Unimodal and Crossmodal Reactivity in Autism: Presence of Auditory Evoked Responses and Effect of the Repetition of Auditory Stimuli. *Biol. Psychiatry* **1992**, *31*, 1190–1203. [https://doi.org/10.1016/0006-3223\(92\)90338-Z](https://doi.org/10.1016/0006-3223(92)90338-Z).

117. von Koss Torkildsen, J.; Hansen, H.F.; Svangstu, J.M.; Smith, L.; Simonsen, H.G.; Moen, I.; Lindgren, M. Brain Dynamics of Word Familiarization in 20-Month-Olds: Effects of Productive Vocabulary Size. *Brain Lang.* **2009**, *108*, 73–88. <https://doi.org/10.1016/j.bandl.2008.09.005>.
118. Garrido, M.I.; Kilner, J.M.; Kiebel, S.J.; Stephan, K.E.; Baldeweg, T.; Friston, K.J. Repetition Suppression and Plasticity in the Human Brain. *Neuroimage* **2009**, *48*, 269–279. <https://doi.org/10.1016/j.neuroimage.2009.06.034>.
119. Gorina-Careta, N.; Zarnowiec, K.; Costa-Faidella, J.; Escera, C. Timing Predictability Enhances Regularity Encoding in the Human Subcortical Auditory Pathway. *Sci. Rep.* **2016**, *6*, 37405. <https://doi.org/10.1038/srep37405>.
120. Fryer, S.L.; Roach, B.J.; Hamilton, H.K.; Bachman, P.; Belger, A.; Carrión, R.E.; Duncan, E.; Johannesen, J.; Light, G.A.; Niznikiewicz, M.; et al. Deficits in Auditory Predictive Coding in Individuals with the Psychosis Risk Syndrome: Prediction of Conversion to Psychosis. *J. Abnorm. Psychol.* **2020**, *129*, 599–611. <https://doi.org/10.1037/abn0000513>.
121. Jamal, W.; Cardinaux, A.; Haskins, A.J.; Kjelgaard, M.; Sinha, P. Reduced Sensory Habituation in Autism and Its Correlation with Behavioral Measures. *J. Autism Dev. Disord.* **2020**, *51*, 3153–3164. <https://doi.org/10.1007/s10803-020-04780-1>.
122. Feuerriegel, D.; Yook, J.; Quek, G.L.; Hogendoorn, H.; Bode, S. Visual Mismatch Responses Index Surprise Signalling but Not Expectation Suppression. *Cortex* **2021**, *134*, 16–29. <https://doi.org/10.1016/j.cortex.2020.10.006>.
123. Alho, K.; Sainio, K.; Sajaniemi, N.; Reinikainen, K.; Näätänen, R. Event-Related Brain Potential of Human Newborns to Pitch Change of an Acoustic Stimulus. *Electroencephalogr. Clin. Neurophysiol.* **1990**, *77*, 151–155. [https://doi.org/10.1016/0168-5597\(90\)90031-8](https://doi.org/10.1016/0168-5597(90)90031-8).
124. Dehaene-Lambertz, G.; Dehaene, S. Speed and Cerebral Correlates of Syllable Discrimination in Infants. *Nature* **1994**, *370*, 292–295. <https://doi.org/10.1038/370292a0>.
125. Cheour, M.; Alho, K.; Ceponiené, R.; Reinikainen, K.; Sainio, K.; Pohjavuori, M.; Aaltonen, O.; Näätänen, R. Maturation of Mismatch Negativity in Infants. *Int. J. Psychophysiol.* **1998**, *29*, 217–226. [https://doi.org/10.1016/s0167-8760\(98\)00017-8](https://doi.org/10.1016/s0167-8760(98)00017-8).
126. Rapaport, H.; Seymour, R.A.; Benikos, N.; He, W.; Pellicano, E.; Brock, J.; Sowman, P.F. Investigating Predictive Coding in Younger and Older Children Using MEG and a Multi-Feature Auditory Oddball Paradigm. *Cereb. Cortex* **2023**, *33*, 7489–7499. <https://doi.org/10.1093/cercor/bhad054>.
127. Gomot, M.; Giard, M.-H.; Roux, S.; Barthélémy, C.; Bruneau, N. Maturation of Frontal and Temporal Components of Mismatch Negativity (MMN) in Children. *NeuroReport* **2000**, *11*, 3109–3112. <https://doi.org/10.1097/00001756-200009280-00014>.
128. Cheour, M.; Leppänen, P.H.; Kraus, N. Mismatch Negativity (MMN) as a Tool for Investigating Auditory Discrimination and Sensory Memory in Infants and Children. *Clin. Neurophysiol.* **2000**, *111*, 4–16. [https://doi.org/10.1016/S1388-2457\(99\)00191-1](https://doi.org/10.1016/S1388-2457(99)00191-1).
129. Charpentier, J.; Kovarski, K.; Houy-Durand, E.; Malvy, J.; Saby, A.; Bonnet-Brilhault, F.; Latinus, M.; Gomot, M. Emotional Prosodic Change Detection in Autism Spectrum Disorder: An Electrophysiological Investigation in Children and Adults. *J. Neurodev. Disord.* **2018**, *10*, 28. <https://doi.org/10.1186/s11689-018-9246-9>.
130. Charpentier, J.; Kovarski, K.; Roux, S.; Houy-Durand, E.; Saby, A.; Bonnet-Brilhault, F.; Latinus, M.; Gomot, M. Brain Mechanisms Involved in Angry Prosody Change Detection in School-Age Children and Adults, Revealed by Electrophysiology. *Cogn. Affect. Behav. Neurosci.* **2018**, *18*, 748–763. <https://doi.org/10.3758/s13415-018-0602-8>.
131. Chen, T.-C.; Hsieh, M.H.; Lin, Y.-T.; Chan, P.-Y.S.; Cheng, C.-H. Mismatch Negativity to Different Deviant Changes in Autism Spectrum Disorders: A Meta-Analysis. *Clin. Neurophysiol.* **2020**, *131*, 766–777. <https://doi.org/10.1016/j.clinph.2019.10.031>.
132. Cléry, H.; Bonnet-Brilhault, F.; Lenoir, P.; Barthelemy, C.; Bruneau, N.; Gomot, M. Atypical Visual Change Processing in Children with Autism: An Electrophysiological Study. *Psychophysiology* **2013**, *50*, 240–252. <https://doi.org/10.1111/psyp.12006>.
133. Gomot, M.; Giard, M.-H.; Adrien, J.-L.; Barthelemy, C.; Bruneau, N. Hypersensitivity to Acoustic Change in Children with Autism: Electrophysiological Evidence of Left Frontal Cortex Dysfunctioning. *Psychophysiology* **2002**, *39*, 577–584. <https://doi.org/10.1017/S0048577202394058>.
134. Gomot, M.; Blanc, R.; Cléry, H.; Roux, S.; Barthelemy, C.; Bruneau, N. Candidate Electrophysiological Endophenotypes of Hyper-Responsivity to Change in Autism. *J. Autism Dev. Disord.* **2011**, *41*, 705–714. <https://doi.org/10.1007/s10803-010-1091-y>.
135. Kovarski, K.; Charpentier, J.; Roux, S.; Batty, M.; Houy-Durand, E.; Gomot, M. Emotional Visual Mismatch Negativity: A Joint Investigation of Social and Non-Social Dimensions in Adults with Autism. *Transl. Psychiatry* **2021**, *11*, 10. <https://doi.org/10.1038/s41398-020-01133-5>.
136. Franken, I.H.A.; Nijs, I.; Van Strien, J.W. Impulsivity Affects Mismatch Negativity (MMN) Measures of Preattentive Auditory Processing. *Biol. Psychol.* **2005**, *70*, 161–167. <https://doi.org/10.1016/j.biopsycho.2005.01.007>.
137. Rydkjær, J.; Jepsen, J.R.M.; Pagsberg, A.K.; Fagerlund, B.; Glenthøj, B.Y.; Oranje, B. Mismatch Negativity and P3a Amplitude in Young Adolescents with First-Episode Psychosis: A Comparison with ADHD. *Psychol. Med.* **2017**, *47*, 377–388. <https://doi.org/10.1017/S0033291716002518>.
138. Yang, M.-T.; Hsu, C.-H.; Yeh, P.-W.; Lee, W.-T.; Liang, J.-S.; Fu, W.-M.; Lee, C.-Y. Attention Deficits Revealed by Passive Auditory Change Detection for Pure Tones and Lexical Tones in ADHD Children. *Front. Hum. Neurosci.* **2015**, *9*, 470. <https://doi.org/10.3389/fnhum.2015.00470>.
139. Chen, C.; Sung, J.-Y.; Cheng, Y. Neural Dynamics of Emotional Salience Processing in Response to Voices during the Stages of Sleep. *Front. Behav. Neurosci.* **2016**, *10*, 117. <https://doi.org/10.3389/fnbeh.2016.00117>.
140. Baldeweg, T.; Hirsch, S.R. Mismatch Negativity Indexes Illness-Specific Impairments of Cortical Plasticity in Schizophrenia: A Comparison with Bipolar Disorder and Alzheimer's Disease. *Int. J. Psychophysiol.* **2015**, *95*, 145–155. <https://doi.org/10.1016/j.ijpsycho.2014.03.008>.

141. Erickson, M.A.; Ruffle, A.; Gold, J.M. A Meta-Analysis of Mismatch Negativity in Schizophrenia: From Clinical Risk to Disease Specificity and Progression. *Biol. Psychiatry* **2016**, *79*, 980–987. <https://doi.org/10.1016/j.biopsych.2015.08.025>.
142. Kirihaara, K.; Tada, M.; Koshiyama, D.; Fujioka, M.; Usui, K.; Araki, T.; Kasai, K. A Predictive Coding Perspective on Mismatch Negativity Impairment in Schizophrenia. *Front. Psychiatry* **2020**, *11*, 660. <https://doi.org/10.3389/fpsy.2020.00660>.
143. Nääätänen, R.; Kähkönen, S. Central Auditory Dysfunction in Schizophrenia as Revealed by the Mismatch Negativity (MMN) and Its Magnetic Equivalent MMNm: A Review. *Int. J. Neuropsychopharm.* **2009**, *12*, 125–135. <https://doi.org/10.1017/S1461145708009322>.
144. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Association: Washington, DC, USA, 2013; ISBN 0-89042-555-8.
145. Baranek, G.T.; Barnett, C.R.; Adams, E.M.; Wolcott, N.A.; Watson, L.R.; Crais, E.R. Object Play in Infants with Autism: Methodological Issues in Retrospective Video Analysis. *Am. J. Occup. Ther.* **2005**, *59*, 20–30. <https://doi.org/10.5014/ajot.59.1.20>.
146. Liss, M.; Saulnier, C.; Fein, D.; Kinsbourne, M. Sensory and Attention Abnormalities in Autistic Spectrum Disorders. *Autism* **2006**, *10*, 155–172. <https://doi.org/10.1177/1362361306062021>.
147. Rogers, S.J.; Hepburn, S.; Wehner, E. Parent Reports of Sensory Symptoms in Toddlers with Autism and Those with Other Developmental Disorders. *J. Autism Dev. Disord.* **2003**, *33*, 631–642. <https://doi.org/10.1023/b:jadd.000006000.38991.a7>.
148. Gomot, M.; Wicker, B. A Challenging, Unpredictable World for People with Autism Spectrum Disorder. *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol.* **2012**, *83*, 240–247. <https://doi.org/10.1016/j.ijpsycho.2011.09.017>.
149. Brock, J. Alternative Bayesian Accounts of Autistic Perception: Comment on Pellicano and Burr. *Trends Cogn. Sci.* **2012**, *16*, 573–574. <https://doi.org/10.1016/j.tics.2012.10.005>.
150. Karaminis, T.; Cicchini, G.M.; Neil, L.; Cappagli, G.; Aagten-Murphy, D.; Burr, D.; Pellicano, E. Central Tendency Effects in Time Interval Reproduction in Autism. *Sci. Rep.* **2016**, *6*, 28570. <https://doi.org/10.1038/srep28570>.
151. Crespo, C.; Santos, S.; Canavaro, M.C.; Kielikowski, M.; Pryor, J.; Féres-Carneiro, T. Family Routines and Rituals in the Context of Chronic Conditions: A Review. *Int. J. Psychol.* **2013**, *48*, 729–746. <https://doi.org/10.1080/00207594.2013.806811>.
152. Rogers, S.J.; Ozonoff, S. Annotation: What Do We Know about Sensory Dysfunction in Autism? A Critical Review of the Empirical Evidence. *J. Child Psychol. Psychiatry* **2005**, *46*, 1255–1268. <https://doi.org/10.1111/j.1469-7610.2005.01431.x>.
153. Cary, E.; Pacheco, D.; Kaplan-Kahn, E.; McKernan, E.; Matsuba, E.; Prieve, B.; Russo, N. Brain Signatures of Early and Late Neural Measures of Auditory Habituation and Discrimination in Autism and Their Relationship to Autistic Traits and Sensory Overresponsivity. *J. Autism Dev. Disord.* **2023**. <https://doi.org/10.1007/s10803-022-05866-8>.
154. Gonzalez-Gadea, M.L.; Chennu, S.; Bekinschtein, T.A.; Rattazzi, A.; Beraudi, A.; Tripicchio, P.; Moyano, B.; Soffita, Y.; Steinberg, L.; Adolphi, F.; et al. Predictive Coding in Autism Spectrum Disorder and Attention Deficit Hyperactivity Disorder. *J. Neurophysiol.* **2015**, *114*, 2625–2636. <https://doi.org/10.1152/jn.00543.2015>.
155. Guiraud, J.A.; Kushnerenko, E.; Tomalski, P.; Davies, K.; Ribeiro, H.; Johnson, M.H. Differential Habituation to Repeated Sounds in Infants at High Risk for Autism. *NeuroReport* **2011**, *22*, 845–849. <https://doi.org/10.1097/WNR.0b013e32834c0bec>.
156. Font-Alaminos, M.; Cornella, M.; Costa-Faidella, J.; Hervás, A.; Leung, S.; Rueda, L.; Escera, C. Increased Subcortical Neural Responses to Repeating Auditory Stimulation in Children with Autism Spectrum Disorder. *Biol. Psychol.* **2020**, *149*, 107807. <https://doi.org/10.1016/j.biopsycho.2019.107807>.
157. Latinus, M.; Cléry, H.; Andersson, F.; Bonnet-Brilhault, F.; Fonlupt, P.; Gomot, M. Inflexibility in Autism Spectrum Disorder: Need for Certainty and Atypical Emotion Processing Share the Blame. *Brain Cogn.* **2019**, *136*, 103599. <https://doi.org/10.1016/j.bandc.2019.103599>.
158. Rubenstein, J.L.R.; Merzenich, M.M. Model of Autism: Increased Ratio of Excitation/Inhibition in Key Neural Systems. *Genes Brain Behav.* **2003**, *2*, 255–267. <https://doi.org/10.1034/j.1601-183x.2003.00037.x>.
159. Briend, F.; Barantin, L.; Cléry, H.; Cottier, J.-P.; Bonnet-Brilhault, F.; Houy-Durand, E.; Gomot, M. Glutamate Levels of the Right and Left Anterior Cingulate Cortex in Autistic Adults. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry* **2023**, *126*, 110801. <https://doi.org/10.1016/j.pnpbp.2023.110801>.
160. Galineau, L.; Arlicot, N.; Dupont, A.-C.; Briend, F.; Houy-Durand, E.; Tauber, C.; Gomot, M.; Gissot, V.; Barantin, L.; Lefevre, A.; et al. Glutamatergic Synapse in Autism: A Complex Story for a Complex Disorder. *Mol. Psychiatry* **2023**, *28*, 801–809. <https://doi.org/10.1038/s41380-022-01860-9>.
161. Sapey-Triomphe, L.-A.; Temmerman, J.; Puts, N.A.J.; Wagemans, J. Prediction Learning in Adults with Autism and Its Molecular Correlates. *Mol. Autism* **2021**, *12*, 64. <https://doi.org/10.1186/s13229-021-00470-6>.
162. Kolesnik, A.; Begum Ali, J.; Gliga, T.; Guiraud, J.; Charman, T.; Johnson, M.H.; Jones, E.J.H. Increased Cortical Reactivity to Repeated Tones at 8 Months in Infants with Later ASD. *Transl. Psychiatry* **2019**, *9*, 46. <https://doi.org/10.1038/s41398-019-0393-x>.
163. Friston, K.J.; Lawson, R.; Frith, C.D. On Hyperpriors and Hypopriors: Comment on Pellicano and Burr. *Trends Cogn. Sci.* **2013**, *17*, 1. <https://doi.org/10.1016/j.tics.2012.11.003>.
164. Lawson, R.P.; Rees, G.; Friston, K.J. An Aberrant Precision Account of Autism. *Front. Hum. Neurosci.* **2014**, *8*, 302. <https://doi.org/10.3389/fnhum.2014.00302>.
165. Van de Cruys, S.; Evers, K.; Van der Hallen, R.; Van Eylen, L.; Boets, B.; de-Wit, L.; Wagemans, J. Precise Minds in Uncertain Worlds: Predictive Coding in Autism. *Psychol. Rev.* **2014**, *121*, 649. <https://doi.org/10.1037/a0037665>.
166. Sinha, P.; Kjelgaard, M.M.; Gandhi, T.K.; Tsourides, K.; Cardinaux, A.L.; Pantazis, D.; Diamond, S.P.; Held, R.M. Autism as a Disorder of Prediction. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15220–15225. <https://doi.org/10.1073/pnas.1416797111>.

167. Cannon, J.; O'Brien, A.M.; Bungert, L.; Sinha, P. Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. *Autism Res.* **2021**, *14*, 604–630. <https://doi.org/10.1002/aur.2482>.
168. Randeniya, R.; Mattingley, J.B.; Garrido, M.I. Increased Context Adjustment Is Associated with Auditory Sensitivities but Not with Autistic Traits. *Autism Res.* **2022**, *15*, 1457–1468. <https://doi.org/10.1002/aur.2759>.
169. Kleinhans, N.M.; Johnson, L.C.; Richards, T.; Mahurin, R.; Greenson, J.; Dawson, G.; Aylward, E. Reduced Neural Habituation in the Amygdala and Social Impairments in Autism Spectrum Disorders. *Am. J. Psychiatry* **2009**, *166*, 467–475. <https://doi.org/10.1176/appi.ajp.2008.07101681>.
170. Millin, R.; Kolodny, T.; Flevaris, A.V.; Kale, A.M.; Schallmo, M.-P.; Gerds, J.; Bernier, R.A.; Murray, S. Reduced Auditory Cortical Adaptation in Autism Spectrum Disorder. *eLife* **2018**, *7*, e36493. <https://doi.org/10.7554/eLife.36493>.
171. Lawson, R.P.; Aylward, J.; Roiser, J.P.; Rees, G. Adaptation of Social and Non-Social Cues to Direction in Adults with Autism Spectrum Disorder and Neurotypical Adults with Autistic Traits. *Dev. Cogn. Neurosci.* **2018**, *29*, 108–116. <https://doi.org/10.1016/j.dcn.2017.05.001>.
172. Hemsley, D.R. The Development of a Cognitive Model of Schizophrenia: Placing It in Context. *Neurosci. Biobehav. Rev.* **2005**, *29*, 977–988. <https://doi.org/10.1016/j.neubiorev.2004.12.008>.
173. Sterzer, P.; Adams, R.A.; Fletcher, P.; Frith, C.; Lawrie, S.M.; Muckli, L.; Petrovic, P.; Uhlhaas, P.; Voss, M.; Corlett, P.R. The Predictive Coding Account of Psychosis. *Biol. Psychiatry* **2018**, *84*, 634–643. <https://doi.org/10.1016/j.biopsych.2018.05.015>.
174. Thakkar, K.N.; Silverstein, S.M.; Brascamp, J.W. A Review of Visual Aftereffects in Schizophrenia. *Neurosci. Biobehav. Rev.* **2019**, *101*, 68–77. <https://doi.org/10.1016/j.neubiorev.2019.03.021>.
175. Bolino, F.; Di Michele, V.; Di Cicco, L.; Manna, V.; Daneluzzo, E.; Casacchia, M. Sensorimotor Gating and Habituation Evoked by Electro-Cutaneous Stimulation in Schizophrenia. *Biol. Psychiatry* **1994**, *36*, 670–679. [https://doi.org/10.1016/0006-3223\(94\)91176-2](https://doi.org/10.1016/0006-3223(94)91176-2).
176. Freedman, R. Inhibitory Gating of an Evoked Response to Repeated Auditory Stimuli in Schizophrenic and Normal Subjects: Human Recordings, Computer Simulation, and an Animal Model. *Arch. Gen. Psychiatry* **1996**, *53*, 1114. <https://doi.org/10.1001/archpsyc.1996.01830120052009>.
177. Avery, S.N.; McHugo, M.; Armstrong, K.; Blackford, J.U.; Vandekar, S.; Woodward, N.D.; Heckers, S. Habituation during Encoding: A New Approach to the Evaluation of Memory Deficits in Schizophrenia. *Schizophr. Res.* **2020**, *223*, 179–185. <https://doi.org/10.1016/j.schres.2020.07.007>.
178. Holt, D.J.; Weiss, A.P.; Rauch, S.L.; Wright, C.I.; Zalesak, M.; Goff, D.C.; Ditman, T.; Welsh, R.C.; Heckers, S. Sustained Activation of the Hippocampus in Response to Fearful Faces in Schizophrenia. *Biol. Psychiatry* **2005**, *57*, 1011–1019. <https://doi.org/10.1016/j.biopsych.2005.01.033>.
179. Bodatsch, M.; Brockhaus-Dumke, A.; Klosterkötter, J.; Ruhrmann, S. Forecasting Psychosis by Event-Related Potentials—Systematic Review and Specific Meta-Analysis. *Biol. Psychiatry* **2015**, *77*, 951–958. <https://doi.org/10.1016/j.biopsych.2014.09.025>.
180. Coffman, B.A.; Haigh, S.M.; Murphy, T.K.; Salisbury, D.F. Impairment in Mismatch Negativity but Not Repetition Suppression in Schizophrenia. *Brain Topogr.* **2017**, *30*, 521–530. <https://doi.org/10.1007/s10548-017-0571-1>.
181. Koshiyama, D.; Kirihara, K.; Tada, M.; Nagai, T.; Fujioka, M.; Usui, K.; Araki, T.; Kasai, K. Reduced Auditory Mismatch Negativity Reflects Impaired Deviance Detection in Schizophrenia. *Schizophr. Bull.* **2020**, *46*, 937–946. <https://doi.org/10.1093/schbul/sbaa006>.
182. Rentzsch, J.; Shen, C.; Jockers-Scherübl, M.C.; Gallinat, J.; Neuhaus, A.H. Auditory Mismatch Negativity and Repetition Suppression Deficits in Schizophrenia Explained by Irregular Computation of Prediction Error. *PLoS ONE* **2015**, *10*, e0126775. <https://doi.org/10.1371/journal.pone.0126775>.
183. Mazer, P.; Macedo, I.; Paiva, T.O.; Ferreira-Santos, F.; Pasion, R.; Barbosa, F.; Almeida, P.; Silveira, C.; Cunha-Reis, C.; Marques-Teixeira, J. Abnormal Habituation of the Auditory Event-Related Potential P2 Component in Patients With Schizophrenia. *Front. Psychiatry* **2021**, *12*. <https://doi.org/10.3389/fpsy.2021.630406>.
184. McCleery, A.; Wynn, J.K.; Green, M.F. Hallucinations, Neuroplasticity, and Prediction Errors in Schizophrenia. *Scand. J. Psychol.* **2018**, *59*, 41–48. <https://doi.org/10.1111/sjop.12413>.
185. Kovács, G.; Grotheer, M.; Münke, L.; Kéri, S.; Nenadić, I. Significant Repetition Probability Effects in Schizophrenia. *Psychiatry Res. Neuroimaging* **2019**, *290*, 22–29. <https://doi.org/10.1016/j.psychres.2019.05.006>.
186. Williams, L.E.; Blackford, J.U.; Luksik, A.; Gauthier, I.; Heckers, S. Reduced Habituation in Patients with Schizophrenia. *Schizophr. Res.* **2013**, *151*, 124–132. <https://doi.org/10.1016/j.schres.2013.10.017>.
187. Lee, J.; Reavis, E.A.; Engel, S.A.; Altshuler, L.L.; Cohen, M.S.; Glahn, D.C.; Nuechterlein, K.H.; Wynn, J.K.; Green, M.F. fMRI Evidence of Aberrant Neural Adaptation for Objects in Schizophrenia and Bipolar Disorder. *Hum. Brain Mapp.* **2019**, *40*, 1608–1617. <https://doi.org/10.1002/hbm.24472>.
188. Jansiewicz, E.M.; Newschaffer, C.J.; Denckla, M.B.; Mostofsky, S.H. Impaired Habituation in Children with Attention Deficit Hyperactivity Disorder. *Cogn. Behav. Neurol.* **2004**, *17*, 1–8. <https://doi.org/10.1097/00146965-200403000-00001>.
189. Massa, J.; O'Desky, I.H. Impaired Visual Habituation in Adults With ADHD. *J. Atten. Disord.* **2011**, *16*, 553–561. <https://doi.org/10.1177/1087054711423621>.
190. Conzelmann, A.; Pauli, P.; Mucha, R.F.; Jacob, C.P.; Gerdes, A.B.M.; Romanos, J.; Bähne, C.G.; Heine, M.; Boreatti-Hümmer, A.; Alpers, G.W.; et al. Early Attentional Deficits in an Attention-to-Prepulse Paradigm in ADHD Adults. *J. Abnorm. Psychol.* **2010**, *119*, 594. <https://doi.org/10.1037/a0019859>.

191. Herpertz, S.C.; Mueller, B.; Wenning, B.; Qunaibi, M.; Lichtenfeld, C.; Herpertz-Dahlmann, B. Autonomic Responses in Boys with Externalizing Disorders. *J. Neural Transm* **2003**, *110*, 1181–1195. <https://doi.org/10.1007/s00702-003-0026-6>.
192. Lloyd, D.R.; Medina, D.J.; Hawk, L.W.; Fosco, W.D.; Richards, J.B. Habituation of Reinforcer Effectiveness. *Front. Integr. Neurosci.* **2014**, *7*, 107. <https://doi.org/10.3389/fnint.2013.00107>.
193. Friston, K. The Free-Energy Principle: A Rough Guide to the Brain? *Trends Cogn. Sci.* **2009**, *13*, 293–301. <https://doi.org/10.1016/j.tics.2009.04.005>.
194. Cheng, C.-H.; Chan, P.-Y.S.; Hsieh, Y.-W.; Chen, K.-F. A Meta-Analysis of Mismatch Negativity in Children with Attention Deficit-Hyperactivity Disorders. *Neurosci. Lett.* **2016**, *612*, 132–137. <https://doi.org/10.1016/j.neulet.2015.11.033>.
195. Kanner, L. Autistic Disturbances of Affective Contact. *Nerv. Child* **1943**, *2*, 217–250.
196. Andronikof, A.; Fontan, P. Grounia Efimovna Soukhareva : La Première Description Du Syndrome Dit d'Asperger. *Neuropsychiatr. L'enfance L'adolescence* **2016**, *64*, 58–70. <https://doi.org/10.1016/j.neurenf.2015.07.007>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

## **Habituation, Adaptation and Prediction Processes in Neurodevelopmental Disorders: A Comprehensive Review**

***Merchie, A. & Gomot, M.***

*Article publié dans Brain Sciences (2023)*

Dans cette revue, des difficultés dans l'habituation comportementale et dans l'adaptation neurale ont été mises en lumière dans les troubles du neurodéveloppement. Ces difficultés ont été expliquées par certains auteurs comme étant la conséquence des déficits de prédiction retrouvées dans le contexte du cerveau Bayésien (ou codage prédictif) résultant dans une mise à jour constante des représentations internes et donc à une adaptation plus difficile.

Au niveau comportemental, un besoin de routine et de répétitions est observé chez les enfants autistes. Toutefois, une réduction de l'habituation comportementale a été mise en lumière tout comme une réduction de l'adaptation neurale. Toutes les études sur l'adaptation neurale dans l'autisme ont utilisé des paradigmes qui ne sont pas adaptés rendant difficile de tirer des conclusions. Une seule étude a cherché à mesurer ce phénomène avec un protocole adapté chez des enfants autistes et à montrer une absence d'adaptation neurale au niveau sous-cortical. Des difficultés d'adaptation dans le TDAH et la schizophrénie ont aussi été mises en évidence au niveau comportemental et neural, avec des paradigmes de Roving notamment dans la schizophrénie.

Pour conclure, dans l'autisme, des difficultés d'adaptation et d'habituation pourraient expliquer en partie l'hypersensibilité au changement observée au niveau comportemental. La compréhension plus précise de ce phénomène dans un contexte plus écologique ou social pourrait permettre non seulement d'apporter des connaissances sur les comportements répétitifs et restreints, mais aussi sur les difficultés dans la communication et les interactions sociales.

Cette revue a permis de mettre en évidence un manque d'études spécifiquement adaptées pour mesurer l'adaptation neurale dans les troubles du neurodéveloppement et surtout en réponse à des stimuli plus écologiques.

### **En résumé**

- Dans l'autisme des atypies de production et de perception des émotions vocales rendent la communication avec les autres parfois difficile.
- L'identification des indices acoustiques liés à la prosodie émotionnelle et utilisés par les personnes autistes permettrait de mieux appréhender les difficultés associées.
- Le traitement de la voix et des émotions vocales apparaît typique ou atypique selon la littérature dans l'autisme, surtout chez les adultes, reflétant de potentielles stratégies de compensation avec l'âge et l'expérience.
- Une adaptation neurale atypique dans l'autisme, pourrait expliquer les atypies de détection du changement observées et d'autant plus pour des sons vocaux et/ou émotionnels.

## 4. Chapitre 4 : Réponse aux émotions vocales et contagion émotionnelle

Comme décrit jusqu'à présent dans cette introduction, dans l'autisme il est observé des atypies de réponses aux émotions vocales tant au niveau neurophysiologique qu'au niveau comportemental où des difficultés ont été observées avec une moins bonne reconnaissance. Cependant, les différences dans l'expression des émotions pourraient aussi expliquer les atypies de réponse automatique aux émotions dans l'autisme. Dans ce dernier chapitre introductif, la réponse automatique aux émotions vocales dans le développement typique et dans l'autisme est décrite, avec une focalisation sur la contagion émotionnelle et une description des théories et modèles historiques de la réponse comportementale aux émotions.

### 4.1. Contagion émotionnelle chez le sujet neurotypique

Le terme de contagion émotionnelle est apparu au XIX<sup>ème</sup> siècle dans la littérature avec l'idée de *contagion of feeling* (Baldwin, 1897), puis ce terme est devenu un concept en psychologie dans les années 1980 (Reber, 1985).

Une des fonctions essentielles des émotions dans la communication est la coordination des interactions sociales (Anderson et al., 2003; Frijda & Mesquita, 1994) et leur partage entre les individus (Rimé, 2009). Quand les émotions se transmettent d'un individu à un autre, il s'agit de contagion émotionnelle (Hatfield et al., 1993, 2014). Celle-ci a été définie par Hatfield, Cacioppo et Rapson comme « une tendance à imiter automatiquement et en synchronie les expressions, les vocalisations, les postures et les mouvements d'autrui et par conséquent à converger émotionnellement » (Hatfield et al., 1993). Cette transmission d'état émotionnel peut se faire grâce à plusieurs indices de natures différentes : visuels, auditifs, olfactifs, posturaux ou tactiles, (Haag, 2019; Hatfield et al., 1993) entraînant une forme de mimétisme (*mimicry*). La contagion émotionnelle active conjointement l'état émotionnel, à travers la cartographie automatique des émotions (Damasio & Blanc, 1995), et l'expression de l'émotion chez l'interlocuteur. Cette compétence a longtemps été considérée comme faisant référence à l'empathie, « faculté intuitive de se mettre à la place d'autrui, de percevoir ce qu'il ressent » (M. H. Davis, 1983). Cependant aucun consensus n'existe aujourd'hui quant à l'interaction entre l'empathie et la contagion émotionnelle.

#### **4.1.1. Résonance motrice émotionnelle**

Un des éléments majeurs de la contagion émotionnelle est le mimétisme facial (*facial mimicry*) (Hatfield et al., 2014). Le mimétisme facial, défini comme « le partage de l'expression émotionnelle en réponse à l'expression émotionnelle de quelqu'un d'autre » (U. Hess & Fischer, 2014), se manifeste uniquement par le partage des expressions faciales. Ce mimétisme nécessite un contact visuel entre les individus lors de l'échange (Mauersberger et al., 2022; Niedenthal et al., 2010, 2016). Par conséquent, la résonance motrice induite par une émotion purement auditive, sans contact visuel direct, n'est pas considérée comme du mimétisme facial.

Néanmoins, dans les interactions impliquant la modalité visuelle, le mimétisme facial contribue au partage et à la contagion émotionnelle (Olszanowski et al., 2020; Olszanowski & Wróbel, 2022; Prochazkova & Kret, 2017). La résonance motrice est étroitement liée à la contagion émotionnelle, définie comme « le résultat de tout processus interactionnel dans lequel les individus en viennent à ressentir l'émotion de l'autre » (Hess et al., 2014).

L'encadré n°2 reprend la définition utilisée pour chacun de ces concepts dans le cadre de cette thèse. Plusieurs théories ont essayé d'expliquer les bases neurales et cérébrales de cette contagion émotionnelle. Ces différentes théories sont brièvement décrites dans la suite de ce chapitre.

### Définitions des concepts

Dans la littérature les termes mimétisme (*mimicry*), résonance et contagion sont utilisés de manière interchangeable pour caractériser la réponse automatique à un stimulus émotionnel. Une définition de ces concepts dans le cadre de cette thèse est donc proposée dans cet encadré :

- **Résonance** : réactivité musculaire faciale (**résonance motrice**) ou **autonome** en réponse à une émotion perçue quelle que soit la modalité d'expression (visuelle ou auditive).
- **Mimétisme** : cas particulier de la résonance dans la modalité visuelle, il peut s'exprimer à travers l'activité autonome (**mimétisme pupillaire**) ou musculaire (**mimétisme facial**).
- **Contagion**: « tendance à imiter automatiquement et en synchronie les expressions, vocalisations, postures et mouvements d'autrui ; et par conséquent à converger émotionnellement » (Hatfield et al., 1993).

#### *Encadré 2 : définition des concepts de contagion, résonance et mimétisme émotionnel*

##### 4.1.1.1. Modèles historiques des émotions

Deux modèles périphériques s'opposent concernant le ressenti et les réponses aux émotions : la théorie de James-Lange (James, 1884) et la théorie de Cannon-Bard (Cannon, 1927). Dans la théorie périphérique de James-Lange, les émotions seraient la conséquence de manifestations périphériques (je suis triste parce que je pleure). La théorie centrale, quant à elle, postule qu'il y a d'abord une intégration sous-corticale du stimulus émotionnel, suivie par la prise de conscience de l'émotion et enfin par les manifestations somatiques (je pleure parce que je suis triste). Par la suite, Schachter et Singer ont proposé une autre théorie, celle des « émotions à deux facteurs » (Schachter & Singer, 1962), qui suggère que l'activation physiologique est suivie par la recherche d'indices permettant son étiquetage et l'identification de l'émotion ressentie. Dans un premier temps, l'organisme est activé par le Système Nerveux Autonome (SNA) de manière aspécifique. Cette activation par le SNA est traitée comme un signal d'alerte par le

corps. Dans un second temps, ce signal est étiqueté cognitivement et la réponse émotionnelle adéquate est identifiée. Enfin, une réaction comportementale et subjective en découle.

Plus récemment, dans les années 1990-2000, LeDoux et Damasio ont chacun proposé un nouveau modèle expliquant le ressenti et la contagion émotionnelle. LeDoux dans son ouvrage *Le cerveau des émotions* (2005), se base en partie sur la théorie de James-Lange en proposant un système à deux voies parallèles qui permettent la réponse émotionnelle. Une première voie dite « basse » qui conduit à une réponse rapide en activant des réseaux de neurones thalamiques spécifiques avant un relais amygdalien. En parallèle, la voie « haute » se compose d'inférences nerveuses entre le thalamus et le cortex sensoriel qui vont ensuite projeter sur l'amygdale. La réponse sera alors déclenchée dans un premier temps par la voie basse, plus rapide, puis par la voie haute, toutes deux conduisant à une activation de l'amygdale. Ensuite, des afférences corporelles provoquent une rétroaction au niveau cérébral, activant les zones cérébrales correspondantes impliquées dans l'action. Contrairement à James, pour LeDoux ce n'est pas seulement la rétroaction du corps sur le cerveau qui est importante, c'est surtout la spécialisation des zones activées. Damasio, un autre neurobiologiste, propose lui aussi sa théorie sur le ressenti et la réponse émotionnelle avec des boucles neuro-somatiques et des marqueurs somatiques (Damasio & Blanc, 1995). Il propose en plus de ces boucles, un renforcement du ressenti émotionnel par l'activation des images et des représentations mentales de l'émotion qui seront mises à jour en cas d'activation.

### 4.1.1.2. Les neurones miroirs

Les neurones miroirs ont été décrit pour la première fois au niveau du cortex pré-moteur chez le macaque (Rizzolatti et al., 1996), puis chez l'humain (Keysers & Gazzola, 2010). Chez l'humain, on ne parle pas de neurones mais plutôt de « système miroir » car il n'a pas pu être mesurée d'activité précise de ces neurones. Ce système regroupe un ensemble de structures cérébrales (le gyrus frontal inférieur, le cortex prémoteur dorsal et ventral, et les lobes pariétaux supérieurs et inférieurs) impliquées dans les mécanismes inconscients de repérage, de compréhension et d'imitation du comportement et des états affectifs d'autrui (Molnar-Szakacs & Overy, 2006; Rizzolatti & Craighero, 2004).

Le rôle du système miroir dans la perception des émotions a été confirmé par l'étude des substrats neuronaux du rire et du sourire (Krolak-Salmon et al., 2006). L'aire motrice

supplémentaire gauche est responsable à la fois des processus moteurs et perceptifs dans la reconnaissance des visages, et elle contribue également à la genèse de l'émotion (Krolak-Salmon et al., 2006). La résonance neuronale intersubjective se manifeste non seulement dans les zones motrices, mais également dans les structures sous-corticales associées au ressenti émotionnel (Anders et al., 2011). Ainsi, le système miroir facilite la compréhension des émotions d'autrui. Ces neurones seraient impliqués dans la résonance motrice (Bonaiuto & Arbib, 2010; Oztop et al., 2006; Rizzolatti & Craighero, 2004) et dans la contagion émotionnelle (Gallese, 2005; Iacoboni, 2009; Likowski et al., 2012). Il a été proposé que ce système était une des bases de l'empathie dans la perception des émotions éprouvées par autrui (Gallese & Goldman, 1998; Preston & Waal, 2002).

### 4.1.1.3. Cognition incarnée

La théorie de la cognition incarnée suggère que la simulation des émotions d'autrui est un moyen d'accéder à ses états affectifs (Niedenthal et al., 2010). Par ailleurs, l'hypothèse de la rétroaction faciale propose que la simple production d'une expression faciale peut modifier nos propres émotions (Strack et al., 1988; Wagenmakers et al., 2016). Ainsi, le mimétisme facial serait un moyen de partager les émotions d'autrui. Des études ont tenté de comprendre les implications causales du mimétisme facial en bloquant les réactions faciales. Par exemple, Oberman, Winkielman et Ramachandran (2007) ont montré que mordre un stylo ou mâcher du chewing-gum pouvait altérer la reconnaissance des expressions émotionnelles. De plus, Neal et Chartrand (2011) ont manipulé la capacité des participants à imiter les expressions faciales via des injections de Botox (réduisant les contractions des muscles faciaux et donc le *feedback* musculaire) ou des produits de comblement dermique (n'affectant pas le *feedback*). Les résultats, conformes aux théories incarnées, indiquent que la modulation des signaux de *feedback* facial peut diminuer ou améliorer la capacité à interpréter les émotions faciales d'autrui.

Dans la modalité auditive, une étude récente a étudié les effets d'une inhibition des émotions faciales dans la reconnaissance d'émotions vocales. Une absence d'effet a été montrée, avec des performances similaires dans la reconnaissance de la prosodie émotionnelle, que les expressions faciales soient bloquées, ou non (Vilaverde et al., 2024). Les effets de la rétroaction

faciale sur le jugement correct des émotions peuvent donc être à nuancer, notamment en fonction de la modalité.

Pour résumer, la contagion émotionnelle dans le cadre de la théorie de la cognition incarnée peut se définir comme suit : la rétroaction faciale influence la perception des émotions en suivant trois étapes (Hatfield et al., 2014) : i) une imitation subtile et inconsciente de l'expression faciale par l'observateur ; ii) cette contraction musculaire génère un signal de rétroaction musculaire afférent du visage ; iii) l'utilisation de cette rétroaction pour simuler et contribuer à inférer par la suite la signification de l'expression perçue.

### 4.1.1.4. Modèle neurocognitif de la contagion émotionnelle (NMEC)

En 2017, Prochazkova et Kret ont proposé un Modèle Neurocognitif de la Contagion Émotionnelle (NMEC) qui s'appuie sur le mimétisme automatique : l'activation neuronale conjointe et la rétroaction faciale pour expliquer la contagion émotionnelle et l'empathie (Prochazkova & Kret, 2017). Ce modèle s'intéresse à la manière dont le mimétisme automatique des expressions faciales et les réponses autonomes contribuent à la contagion émotionnelle et au développement de l'empathie. Ce modèle associe les différentes approches présentées plus tôt avec l'implication des neurones miroirs et la théorie de la rétroaction faciale des émotions.

Le mimétisme, qu'il soit moteur ou autonome, crée une base sur laquelle la contagion émotionnelle peut se développer. En percevant et en imitant les réponses émotionnelles d'un partenaire, un individu ressent des émotions similaires, facilitant ainsi une compréhension empathique plus précise. Le modèle NMEC intègre ces mécanismes pour expliquer comment le mimétisme, à différents niveaux (moteur et autonome), mène à la contagion émotionnelle (figure 5). Plusieurs étapes vont ainsi conduire à la compréhension empathique : i) l'observateur perçoit l'émotion de l'émetteur ce qui va ii) activer les mêmes représentations neuronales chez l'observateur grâce au système miroir conduisant à iii) un mimétisme automatique moteur et autonome. Une fois ces effecteurs activés iv) le mimétisme conduit à un feedback physiologique et émotionnel, induisant des émotions similaires chez l'observateur permettant v) la compréhension de l'état mental de l'émetteur, la compréhension empathique.

Dans ce modèle, des variations de la fréquence cardiaque, de la réponse électrodermale ou encore des modulations du diamètre pupillaire sont proposées comme effecteurs du système nerveux autonome reflétant la contagion émotionnelle.

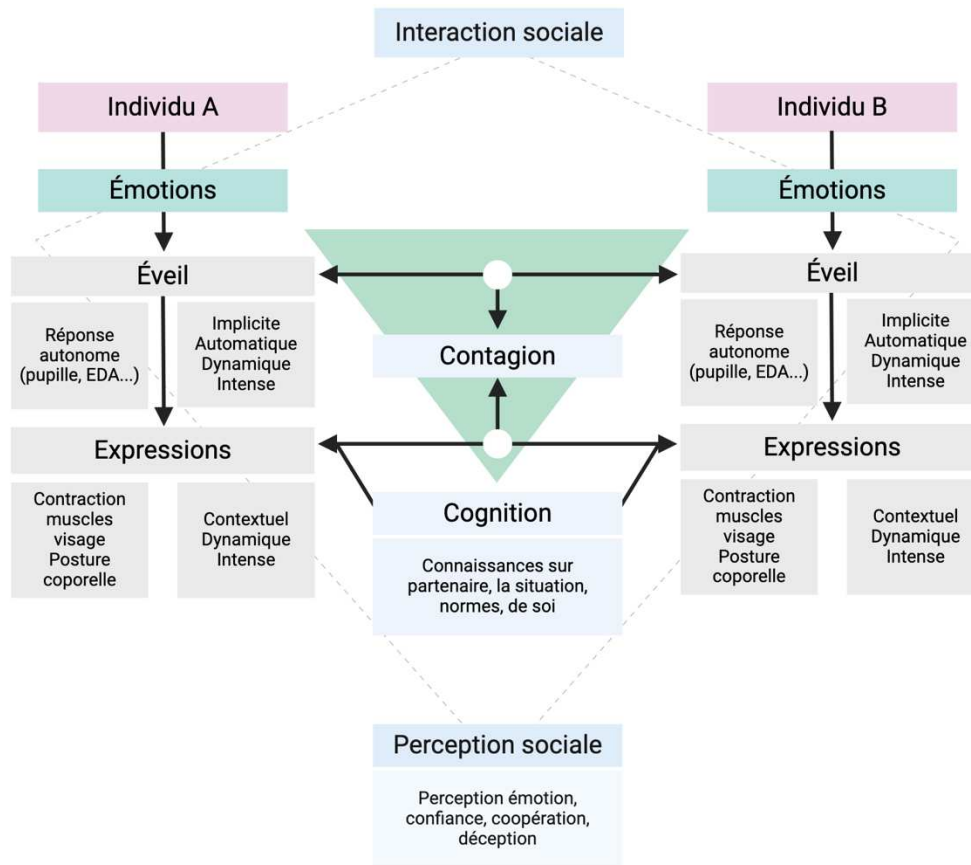


Figure 5 : Modèle de la contagion émotionnelle adapté de Prochazkova & Kret (2017).

#### 4.1.2. Contagion autonome et pupillométrie

Plusieurs indices du Système Nerveux Autonome (SNA) reflètent la contagion émotionnelle : l'activité électrodermale (Lang et al., 1993), le rythme cardiaque (Lang et al., 1993) ou encore la pupille (E. H. Hess & Polt, 1964). La pupille reflète plusieurs fonctions principales : l'adaptation à la luminosité environnante par le réflexe photomoteur, l'accommodation de la vision et enfin l'augmentation de l'éveil ou en fonction de la complexité de l'information à encoder (Mathôt, 2018). Les modulations du diamètre de la pupille reflètent l'activité du LC-NE (noyau du *Locus-Coeuruleus-NorEpinéphrine*) (Aston-Jones & Cohen, 2005). Une description de la physiologie de la dilatation et de la constriction pupillaire est faite dans la partie méthodologie de la thèse (2.3.3.2. Pupillométrie).

Dans la modalité visuelle, le mimétisme pupillaire, ou contagion pupillaire, c'est-à-dire l'adaptation automatique du diamètre de la pupille en fonction de celle de l'interlocuteur, a été décrite dans la littérature chez les primates (Kret et al., 2014), l'adulte et l'enfant humain (Fawcett et al., 2016, 2017; Harrison et al., 2006, 2009; Kret et al., 2014, 2015; Kret & De Dreu, 2017). Chez l'Homme, dans les premières descriptions du phénomène, un lien entre cette contagion pupillaire et l'attractivité a été montré, les participants jugeaient plus attractifs les personnes avec une pupille plus dilatée (E. H. Hess, 1965). Ce résultat ne s'est pas confirmé dans d'autres études (Demos et al., 2008; Mathôt & Naber, 2018). Dans le NMEC, Prochazkova et Kret proposent une composante autonome à la contagion émotionnelle qui passerait entre autre par ce mimétisme pupillaire (Prochazkova & Kret, 2017). Ce modèle est ainsi cohérent avec l'idée que la contagion pupillaire et la sensibilité aux signaux pupillaires d'autres personnes seraient corrélées positivement avec l'empathie (Harrison et al., 2007).

La réponse pupillaire n'est pas seulement sensible aux modulations du diamètre de la pupille d'autrui, mais c'est aussi un indice de la réactivité émotionnelle. Il a été montré des modulations du diamètre pupillaire en réponse à des images émotionnelles (Aguillon-Hernandez et al., 2020; M. M. Bradley et al., 2008; Geangu et al., 2011) mais aussi à des émotions faciales (Aguillon-Hernandez et al., 2020; Burley & Daughters, 2020). La valence de l'émotion exerce aussi une influence sur la dilatation pupillaire. En effet, pour certains auteurs les émotions négatives induisent une plus large dilatation que les émotions positives (Aguillon-Hernandez et al., 2020). Pour d'autres auteurs, l'effet inverse a été observé avec une plus large dilatation en réponse à des émotions positives que négatives (Burley & Daughters, 2020). En revanche, la pupille est surtout sensible à l'authenticité des stimuli présentés et réagit davantage quand ils sont sociaux (Aguillon-Hernandez et al., 2020; Cosme et al., 2021).

### **4.1.3. La contagion émotionnelle auditive**

Comme il a été montré dans ce chapitre, la contagion émotionnelle est un phénomène qui a été largement étudié et théorisé dans la littérature. Un nombre important de ces études s'est focalisé sur la réponse et la réaction à des stimulations visuelles. Dans le cadre de ce travail de thèse, nous nous focaliserons principalement sur la prosodie émotionnelle. Ce paragraphe présentera donc les quelques études, portant sur la contagion émotionnelle dans la modalité auditive, disponibles dans la littérature.

Une résonance motrice congruente à des émotions entendues a été mise en évidence au travers de la mesure de l'activité des muscles de la face pour des vocalisations (Larsen et al., 2003; Verona et al., 2004), de la musique (Bullack et al., 2018) et en réponse à un contenu prosodique émotionnel (Hietanen et al., 1998).

Il existe une forme de réaction émotionnelle à la musique, avec l'expression physique et physiologique de l'émotion perçue lors de l'écoute musicale (Justin, 2013; Justin & Västfjäll, 2008). Cette contagion émotionnelle, observée lors de l'écoute musicale, a également été mise en évidence au niveau cérébral, avec une synchronisation entre plusieurs individus écoutant de la musique simultanément lors d'un même concert (Chabin et al., 2022). Cependant, cette contagion liée à la musique pourrait dépendre davantage de la contagion émotionnelle entre les personnes environnantes que du contenu émotionnel de la musique elle-même (Holochwost & Izard, 2008).

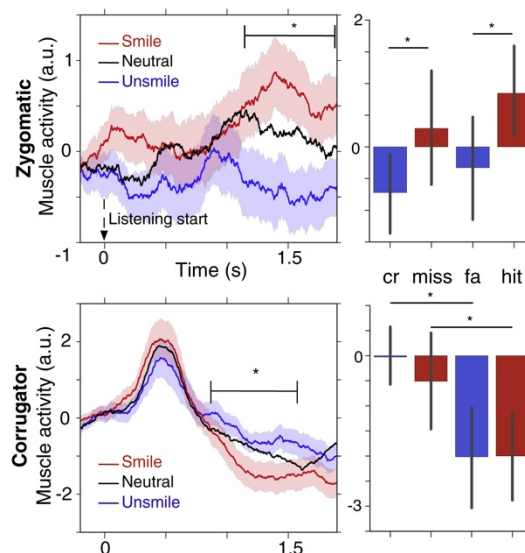
La pupille est aussi sensible à des stimuli non-visuels (Bufo et al., 2022; Zekveld et al., 2018). En effet, il a été montré que la pupille réagissait à des sons de différentes natures (Cherng et al., 2020; Legris et al., 2022; Zekveld et al., 2018). Tout comme dans la modalité visuelle, les sons émotionnels induisent une dilatation pupillaire plus large que les neutres (Babiker et al., 2015; Jin et al., 2015; Jürgens et al., 2018; Partala & Surakka, 2003). Un effet de la valence est alors constaté : les émotions négatives entraînent une dilatation pupillaire plus marquée (Babiker et al., 2015; Jürgens et al., 2018). Cependant, ces quelques études ont mesuré les réponses à des émotions auditives très marquées, telles que des cris de bébé ou des sons de foule en délire (vocalisations, base de stimuli IADS M. Bradley & Lang, 2007), plutôt que sur des sons avec des modulations prosodiques subtiles (Kuchinke et al. 2015 ; Olivia & Anikin, 2018 ; Cosme et al., 2021).

### **4.1.4. Cas particulier du sourire**

Le sourire est l'émotion classiquement utilisée dans l'étude de la résonance motrice et de la contagion émotionnelle. En effet, la contagion au sourire est présente dans plusieurs espèces (Burrows et al., 2006) et chez l'homme dès le plus jeune âge (Simpson et al., 2014). Le sourire étant un indice social important, la réponse automatique au sourire de quelqu'un conduit à une synchronisation des sentiments (sensation de récompense) (Shore & Heerey, 2011) et à la réciprocité sociale (U. Hess & Bourgeois, 2010). Le sourire dans la modalité visuelle conduit à

un mimétisme facial (Dimberg & Thunberg, 1998) qui dans certaines études a été combiné avec un report de l'état émotionnel du participant pour estimer si une synchronisation des états était observée ; on parle alors de contagion émotionnelle (Hatfield et al., 1992, 1993; Soussignan, 2002).

Enfin, comme présenté dans le chapitre 2, le sourire peut être transmis par la seule voix. Il a été démontré que le sourire vocal, pouvait provoquer une résonance motrice congruente à l'écoute. Cela se manifeste par une contraction du muscle zygomatique à l'écoute de phrases prononcées de manière souriantes, même lorsqu'elles ne sont pas correctement identifiées comme telles, et par une contraction du corrugateur pour des phrases jugées non souriantes. Ces effets ont été observés chez des adultes neurotypiques et des aveugles congénitaux (Arias, Bellmann, et al., 2021; Arias et al., 2018) (figure 6).



**Figure 6 : Résonance motrice au sourire vocal à partir de Arias et al. (2018)**

*Cr : correct rejection ; fa : false alarm ; miss et hit sont des indices issus de la théorie de la détection du signal pour classifier les réponses.*

## 4.2. Contagion émotionnelle dans l'autisme

### 4.2.1. Résonance motrice et contagion émotionnelle

Dans l'autisme, des atypies de production des émotions faciales spontanées en réponse à des émotions faciales (mimétisme facial) (McIntosh et al., 2006) ou volontaires (Brewer et al., 2016) ont été avancées. Ces atypies peuvent être liées aux difficultés observées dans les processus d'imitation, définis comme la "reproduction d'une mimique, d'un geste, d'une

posture, ou d'une situation vécue" (Meltzoff & Moore, 1977) et qui ont été largement documentées dans l'autisme (Nadel & Potier, 2002).

#### 4.2.1.1. Expressions faciales

Un mimétisme facial absent aux émotions visuelles a été mis en évidence chez des adultes autistes (McIntosh et al., 2006). Ces difficultés de réaction automatique aux émotions ont été suggérées comme étant associées à l'absence de contagion face à d'autres phénomènes comme le bâillement dans l'autisme (Helt et al., 2010; Senju et al., 2007). En effet, dans l'autisme on retrouve une absence de contagion au bâillement dans la modalité visuelle (Helt et al., 2010) mais aussi auditive (Giganti & Esposito ZIELLO, 2009). Cependant tous les comportements sensibles à la contagion ne semblent pas altérés dans le TSA, certains apparaissent même comme amplifiés comme la contagion aux démangeaisons (Helt, de Marchena, et al., 2021; Helt, Sorensen, et al., 2021).

Le sourire est un stimulus social particulier dans l'autisme. En effet, un des premiers signes cliniques d'alerte dans les premières années de vie est l'absence de réponse au sourire social chez les enfants (Beall et al., 2008; G. Dawson et al., 1990), contrairement aux bébés neurotypiques qui répondent au sourire dès 3 mois de vie (Dondi et al., 2010). Le sourire est un exemple d'émotion faciale qui conduit à une résonance motrice (U. Hess et al., 2014; Mauersberger & Hess, 2019; Olszanowski et al., 2020) qui n'est pas retrouvée dans l'autisme (Beall et al., 2008; McIntosh et al., 2006; Oberman et al., 2009).

L'absence de résonance motrice aux émotions faciales dans l'autisme apparaît aussi comme étant dépendante du contexte. Certaines études n'ont pas mis en évidence de différences dans les réponses motrices faciales entre les groupes de participants autistes et neurotypiques (Magnée et al., 2007; Press et al., 2010). En effet, aucune distinction n'a été observée entre les individus autistes et neurotypiques concernant l'imitation volontaire des expressions émotionnelles. Si la mise en place de ce mimétisme peut s'avérer plus lente dans l'autisme, elle reste la même que chez des individus neurotypiques (Oberman et al., 2009).

Il a été proposé que ces difficultés d'imitation et de résonance dans l'autisme seraient la conséquence d'atypies d'activation des systèmes miroirs (Dapretto et al., 2006; Hamilton, 2008; Iacoboni & Dapretto, 2006). Néanmoins, cette théorie est débattue (Hamilton, 2013) et

aucun consensus n'existe quand à une défaut d'activation du système miroir dans l'autisme (Chan & Han, 2020).

#### 4.2.1.2. Contagion pupillaire

Une contagion pupillaire aux émotions a été montré dans l'autisme dans la modalité visuelle malgré un temps passé sur la région des yeux moins important (Galazka et al., 2019). Une autre étude rapporte même une contagion pupillaire plus importante chez des enfants avec un haut-risque d'autisme que chez des enfants neurotypiques (Wagner et al., 2016). Aguillon-Hernandez et ses collègues (2020) ont mis en évidence que la contagion pupillaire en réponse aux émotions visuelles est possible chez les enfants neurotypiques, néanmoins elle dépend de la nature écologique des stimuli présentés. En effet, une dilatation pupillaire en réponse à des visages émotionnels a été observée chez ces enfants uniquement avec des stimuli dynamiques, sans différence notable pour les images fixes. Au contraire, chez les enfants autistes, aucune différence n'a été mise en évidence en réponse à des visages neutres et émotionnels, qu'ils soient statiques ou dynamiques (Aguillon-Hernandez et al., 2020).

Quelques études se sont intéressées à la réponse pupillaire aux émotions auditives dans l'autisme. Une étude a montré qu'une réponse plus importante pouvait être observée en réponse à la prosodie émotionnelle chez des adultes autistes, toutefois cette réponse pouvait être dépendante de la nature de la tâche (Kuchinke et al., 2011). En effet, aucune modulation de la réponse pupillaire n'a été observée chez les adultes autistes pour une phrase prononcée avec une prosodie positive, par rapport à une phrase neutre, dans une condition où les participants n'avaient pas à inférer l'état émotionnel du locuteur (condition passive). En revanche, dans la condition active, une augmentation du diamètre de la pupille a été observée pour des sons avec une prosodie positive et négative chez des adultes autistes et neurotypiques en comparaison à une prosodie neutre (Kuchinke et al., 2011). Comme pour la résonance motrice (McIntosh et al., 2006), il apparaît donc que la nature de la tâche proposée a une influence majeure sur l'observation ou non de ces phénomènes.

La littérature a révélé des atypies dans la contagion émotionnelle et la résonance motrice visuelle chez les personnes autistes, toutefois les études explorant ces phénomènes en réponse aux émotions dans d'autres modalités restent rares. L'exploration combinée des aspects de résonance motrice et autonome en réponse à des émotions vocales subtiles

#### Chapitre 4 : Réponse aux émotions vocales et contagion émotionnelle

permettrait de mieux caractériser le phénomène de contagion émotionnelle dans l'autisme. Comprendre ces phénomènes et les étapes potentiellement atypiques dans l'autisme pourrait améliorer la prise en charge des personnes autistes, en particulier pour la reconnaissance et la réponse aux émotions vocales.

### En résumé

- Plusieurs théories ont cherché à expliquer la réponse automatique aux émotions et la contagion émotionnelle.
- La contagion émotionnelle a été décrite chez le sujet neurotypique en réponse aux émotions faciales et vocales, mais jamais à l'aide de la mesure combinée de plusieurs effecteurs autonomes et moteurs.
- Dans l'autisme, une absence de résonance motrice et autonome a été décrite mais aucune étude ne s'est intéressée spécifiquement à ces phénomènes, simultanément, en réponse aux émotions vocales.

## 5. Hypothèses et objectifs

La revue de la littérature a permis de mettre en évidence que dans l'autisme les différentes étapes conduisant à une résonance émotionnelle pourraient être altérées ou atypiques. Bien que certaines études aient examiné ces différentes étapes chez des adultes ou des enfants autistes, aucune n'a cherché à les étudier de manière intégrative chez un même individu afin de déterminer la ou lesquelles sont les plus atypiques dans l'autisme en réponse à un stimulus en particulier, le sourire vocal.

Ce travail de thèse s'intéresse aux différentes étapes chez des adultes autistes en comparaison à des adultes neurotypiques en se focalisant sur les émotions vocales. En effet, ces processus ont été plus largement décrits dans la modalité visuelle. Cependant, la voix est un vecteur important dans la communication à travers le contenu linguistique, mais aussi à travers la prosodie et en particulier la prosodie émotionnelle. L'objectif principal de ce travail est de caractériser le traitement et la réponse aux émotions vocales dans l'autisme au cours des trois étapes de la boucle *Perception – Représentation – Action*. Cet objectif principal est sous-tendu par plusieurs questions :

- 1) **Quelles sont les caractéristiques de l'adaptation neurale à des sons avec un contenu social et prosodique émotionnel dans l'autisme ?**
- 2) **Quelles sont les représentations perceptives des émotions vocales dans l'autisme ?**
- 3) **Quelles sont les caractéristiques de la contagion émotionnelle (résonance motrice et autonome) dans l'autisme ? Est-ce que cette contagion est modulée par la représentation perceptive ?**

Dans le cadre de ce travail de thèse, trois grandes hypothèses pouvaient être posées pour étudier le traitement de la prosodie émotionnelle dans l'autisme.

- **Hypothèse 1** : l'adaptation neurale à des sons pourrait être atypique dans l'autisme. Ces atypies pourraient être plus marquées pour des sons vocaux et émotionnels, en lien avec les difficultés importantes du traitement de la voix et des informations sociales et émotionnelles.

## Hypothèses et objectifs

- **Hypothèse 2** : les représentations perceptives des émotions vocales sont différentes dans l'autisme, avec l'utilisation de paramètres acoustiques différents de ceux des adultes neurotypiques. Si ces représentations sont différentes la dernière étape de la boucle en réponse à ces émotions vocales serait également altérée.
- **Hypothèse 3** : les représentations perceptives se basent sur les mêmes indices acoustiques dans l'autisme et dans le développement typique. Les particularités de réponse aux émotions vocales seraient alors spécifiques aux atypies de résonance motrice (et autonome).

Les 3 étapes de la boucle ont été étudiées à l'aide de méthodes d'exploration différentes (EEG, psychophysique, EMG facial et pupillométrie) afin de tester nos hypothèses dans le cadre du traitement des émotions vocales dans un groupe d'adultes autistes, ainsi que dans un groupe d'adultes neurotypiques. Par la suite, les résultats d'adultes neurotypiques et autistes seront comparés afin de mieux caractériser les potentielles difficultés.

Ce travail a ainsi mené à la réalisation de cinq études différentes ayant donné lieu à cinq articles présentés dans les parties suivantes, qui reprennent les étapes de la boucle *Perception* (**études 1 et 2**) – *Représentation* (**étude 3**) – *Action* (**études 4 et 5**) (figure 7).

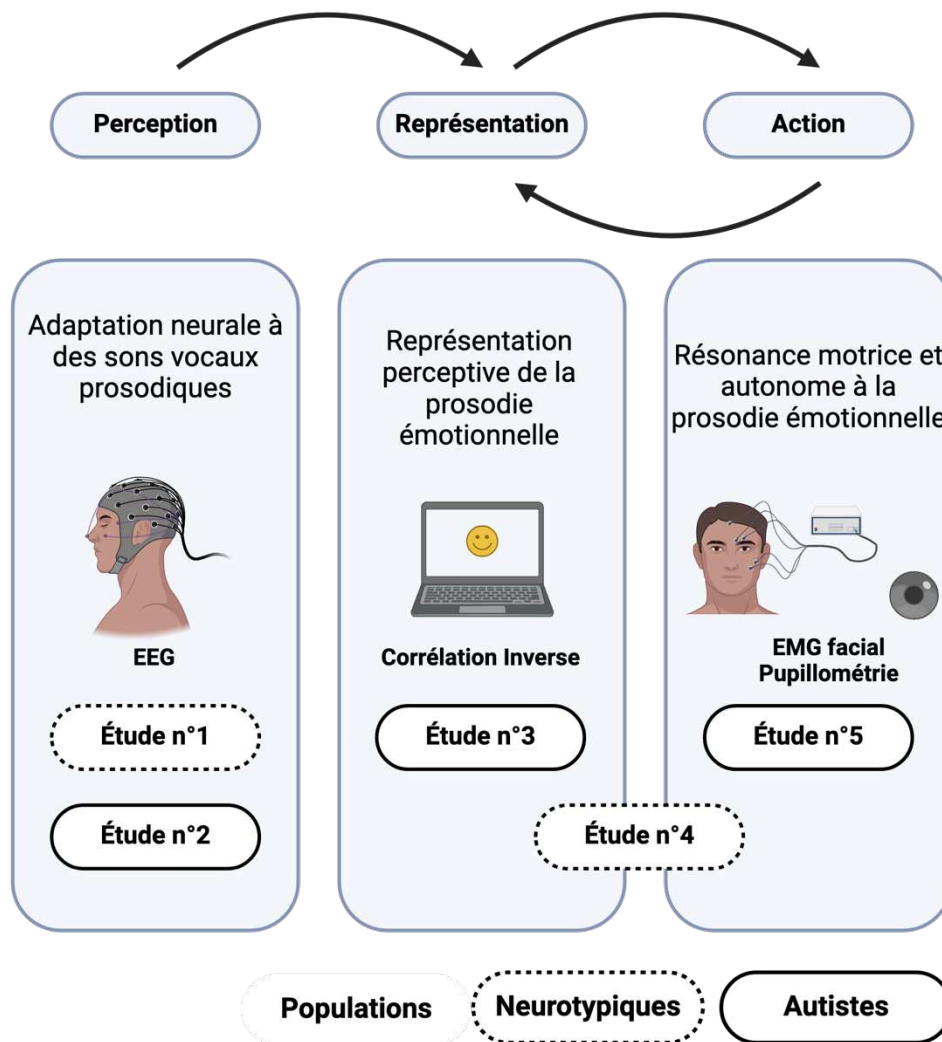


Figure 7 : Représentation des différentes études présentées dans la thèse

**Étude n°1 :** *Neural repetition suppression to vocal and non-vocal sounds*, publié dans la revue *Cortex*, Heurteloup<sup>\*</sup>, **Merchie<sup>\*</sup>** et al., 2022 (<sup>\*</sup>co-premières autrices)

- Objectif : déterminer si la nature sociale d'un son a un impact sur les processus d'adaptation neurale chez des adultes neurotypiques.

**Étude n°2 :** *Unraveling neural adaptation to vocal and non-vocal sounds in autism: effect of emotional prosody*, **Merchie** et al. (2025), publié dans *Clinical Neurophysiology*

- Objectif : étudier l'adaptation neurale dans l'autisme à des sons de différentes natures afin d'évaluer l'effet du contenu social sur ce processus.

**Étude n°3** *Unveiling Mental Representations of Vocal Smile and Internal Noise in Autism: a Reverse Correlation Study*, **Merchie** et al., en révision dans *Research in Autism Spectrum Disorders*

## Hypothèses et objectifs

- Objectif : déterminer si la représentation perceptive du sourire vocal se base sur les mêmes indices acoustiques dans l'autisme que dans la population neurotypique à l'aide d'une tâche de corrélation inverse.

**Étude n°4** : *Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance*, **Merchie** et al. (2024), publié dans *Scientific reports*

- Objectif : étudier la contagion émotionnelle à travers la combinaison de mesures de résonance motrice et pupillaire dans un groupe d'adultes neurotypiques en réponse à des sons émotionnels.

**Étude n°5** : *Autism : When my face is not listening to your smile*, **Merchie** et al., soumis dans *Nature Human Behavior*.

- Objectif : étudier la résonance motrice implicite et explicite au sourire vocal dans l'autisme en lien avec les représentations perceptives

## **Méthodologie**

---



## **1. Populations**

### ***1.1.1. Cadre éthique***

Pour pouvoir réaliser les différentes études de ce projet de thèse des autorisations éthiques concernant la recherche impliquant la personne humaine étaient nécessaires. Ces recherches ont été réalisées sous la promotion du CHRU de Tours dans le cadre d'une autorisation par le Comité de Protection des Personnes Est I (PROSCEA 2017/23 ; ID RCB: 2017-A00756-47). Les expérimentations menées dans cette thèse respectent donc la réglementation en vigueur sur les recherches impliquant la personne humaine. Les participants (et/ou leur tuteur légal en cas de tutelle ou de curatelle) ont tous reçu une information et ont signé un formulaire de consentement éclairé quant à leur participation à cette étude, conformément à la déclaration d'Helsinki. Les participants autistes (s'ils n'étaient pas sous tutelle ou curatelle) et neurotypiques recevaient chacun une compensation financière de 20€, pour leur participation à l'étude.

### ***1.1.2. Populations recrutées***

Les sujets inclus dans le cadre de cette thèse étaient des personnes majeures. Tous les sujets présentaient une audition normale et une vision normale ou corrigée, et ne prenaient aucun traitement médicamenteux pouvant modifier l'électrogénèse au moment de leur participation. Les critères d'exclusion comprenaient : une absence de maladie infectieuse ou métabolique sévère, pas d'épilepsie ni d'antécédents familiaux ou personnels de maladies du système nerveux central ou d'antécédents psychiatriques. Pour les participants neurotypiques, une absence de délai dans l'acquisition du langage et de la marche était également requise. Pour les patients, un diagnostic de Trouble du Spectre de l'Autisme (TSA) selon les critères du DSM-5 (American Psychiatric Association, 2013) était nécessaire.

Le recrutement des adultes autistes a été réalisé en majorité par l'intermédiaire du Centre Ressources Autisme Centre Val de Loire à Tours en collaboration avec les équipes cliniques, en particulier la Dr. Emmanuelle Houy-Durand. Les adultes neurotypiques ont été recrutés à l'aide de listes de diffusion étudiante ou bien par des connaissances afin de recruter une population assez diversifiée.

## Matériel et méthodes

Un récapitulatif des échantillons recrutés pour chacun des protocoles et pour chaque groupe est présenté dans le tableau 2. Les participants inclus pouvaient participer indépendamment à chacun des protocoles, ainsi les participants ne sont pas tous les mêmes pour chacune des études. Par ailleurs, en fonction du niveau intellectuel des adultes autistes nous pouvions leur proposer de participer à l'un ou l'autre des protocoles qui nécessitait d'être actif ou non.

	Neurotypiques	Autistes
EEG	23	21
RevCor	29	30
EMG/Pupillométrie	28	20

**Tableau 2 : Récapitulatif des participants inclus dans les protocoles définitifs présentés dans ce travail de thèse**

Les résultats de l'étude seront communiqués par courrier et/ou par email à l'ensemble des participants à l'issue de ce travail de thèse.

L'annexe 1 récapitule le nombre total de participants vus dans le cadre de cette thèse et du projet SEPIA (*Sensory and Emotional Processing In Autism spectrum disorders*) en fonction des protocoles et des versions qui ne sont pas toutes reprises dans le corps du manuscrit. Les résultats qui ne sont pas présentés dans la partie résultats de la thèse, le sont brièvement dans les annexes 2 et 3.

## 2. Matériel et méthodes

Dans le but de caractériser la boucle *Perception – Représentation – Action* dans l'autisme en réponse à des émotions vocales, plusieurs méthodes d'exploration ont été utilisées : l'électroencéphalographie (EEG), la corrélation inverse (RevCor), l'électromyographie faciale (fEMG) et la pupillométrie. Dans la partie suivante de ce manuscrit une description concise de chacune de ces méthodes et des paradigmes expérimentaux mis en place est proposée.

### 2.1. Électroencéphalographie

#### 2.1.1. Principe de l'EEG

L'électroencéphalogramme (EEG) a été développé par Hans Berger (1929). À partir de cette technique le principe des potentiels évoqués a ensuite été proposé par P.A. Davis (1939) et G.D. Dawson (1954).

## Matériel et méthodes

Le cerveau humain présente une activité biochimique et électrique spontanée. Cette activité intracérébrale produit des fluctuations de potentiels entre différents points du scalp, enregistrées sous forme d'électroencéphalogramme. Les potentiels enregistrés proviennent de l'activité électrique des cellules nerveuses sous-jacentes. Lorsqu'un neurone est excité, des canaux ioniques s'ouvrent dans sa membrane. En raison des différences de composition ionique entre l'intérieur et l'extérieur des cellules, cette ouverture entraîne des mouvements ioniques entre les milieux intracellulaire et extracellulaire. À l'échelle cellulaire, ces mouvements génèrent un potentiel d'action qui se propage du corps cellulaire à l'extrémité de l'axone, formant ainsi un dipôle électrique. Ce dipôle crée un petit potentiel électrique dans les milieux environnants. Si un nombre suffisant de cellules sont excitées et orientées parallèlement, comme les cellules pyramidales, leurs potentiels s'additionnent et deviennent détectables au niveau du scalp. Le signal EEG correspond donc à la sommation des potentiels post-synaptiques des cellules pyramidales des couches 3 à 5 du cortex cérébral.

L'EEG permet de mesurer l'activité cérébrale spontanée et globale, mais il permet aussi de mesurer la réponse à des stimuli spécifiques. La réponse est alors qualifiée de potentiel évoqué (PE). Ces PE permettent d'étudier le décours temporel de la réponse cérébrale spécifique à une stimulation. Les PE étudiés dans ce travail sont qualifiés de corticaux, car générés au niveau cortical et mesurés à des latences « longues » en comparaison aux PE du tronc cérébral qui ont une latence courte ( $<10\text{ms}$ ). Les PE ont une amplitude moindre (environ  $10\mu\text{V}$ ) en comparaison à l'activité globale de l'EEG (environ  $100\mu\text{V}$ ). Afin de visualiser ces réponses la méthode du moyennage est donc utilisée. Pour ce faire, une fenêtre temporelle de quelques centaines de millisecondes est isolée autour de chaque répétition du stimulus d'intérêt. Toutes les réponses obtenues dans cette fenêtre sont moyennées. La répétition du stimulus permet d'augmenter le rapport signal sur bruit. En effet, la puissance du potentiel évoqué augmente car elle survient de manière similaire à chaque répétition, tandis que le bruit généré par l'activité cérébrale de fond s'annule, étant considéré comme "aléatoire" et donc non synchronisée avec la stimulation. Les PE sont ensuite caractérisés selon leur polarité (Négative ou Positive), leur latence d'apparition (après la présentation du stimulus) et le site où ils culminent, on parlera alors de différentes composantes (par exemple N100, N170, P300...).

### **2.1.2. Séquence de stimulation**

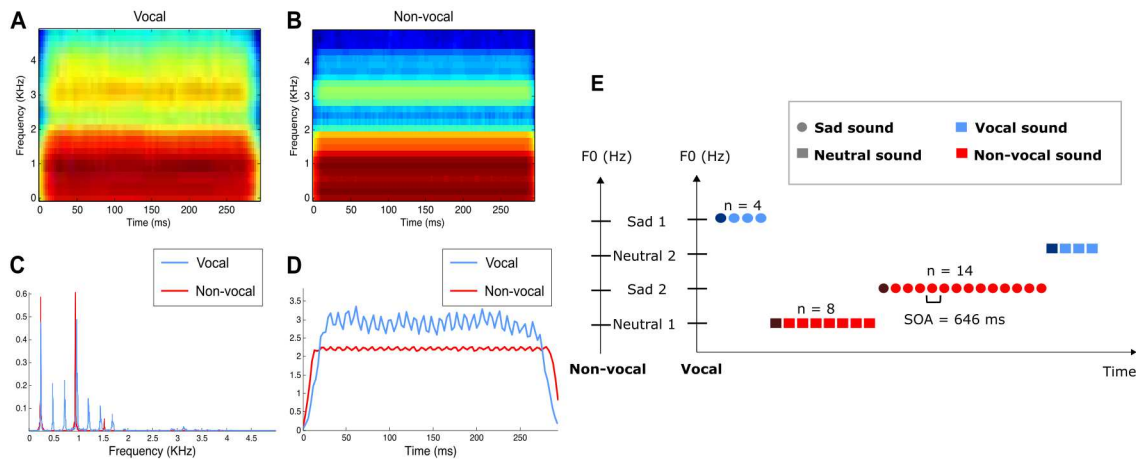
Afin de mesurer et d'étudier l'adaptation neurale à différents types de son, un paradigme de *roving* est utilisé. Ce type de paradigme permet de mesurer la détection du changement mais aussi l'adaptation. Dans ce paradigme la séquence de sons est organisée en trains de répétitions de longueurs variables (Cowan et al., 1993). Le premier son de chaque train correspondra alors au stimulus déviant alors qu'à partir de la deuxième répétition on parlera de standard. Afin d'étudier le décours temporel de la mise en place de l'adaptation neurale des trains de 3 longueurs différentes sont utilisés 4, 8 ou 14 répétitions, dans l'étude n°2 par exemple.

La durée du protocole était de 1h pour le premier (**étude n°1**) et de 25 minutes pour le second (**étude n°2**). Afin de choisir les sons et la séquence la plus adaptée pour étudier l'adaptation à des sons de natures différentes une étude préliminaire a été réalisée. Cette étude préliminaire est présentée en Annexe 2.

Les sons utilisés dans ces paradigmes étaient des sons vocaux et non-vocaux pour les 2 études. Dans l'étude n°1, les sons vocaux étaient des phonèmes /a/ prononcés par 8 locutrices différentes avec une prosodie neutre. La prosodie a été testée et validée dans une étude précédente (Charpentier, Kovarski, Houy-Durand, et al., 2018). Les sons non-vocaux ont ensuite été synthétisés à partir des caractéristiques acoustiques des sons vocaux. Pour cela, la fondamentale et les quatre premiers formants des sons vocaux ont été extraits sur Praat® (Boersma, 2002). Ensuite, les sons complexes non-vocaux ont été synthétisés en utilisant ces fréquences et en adaptant aux caractéristiques spectrales des voix (Adobe Audition®). Tous les sons ont été normalisés en énergie et une atténuation de l'intensité en début et en fin de son a été appliquée.

Pour l'étude n°2, la même méthode a été utilisée pour la création de sons complexes non-vocaux : deux catégories prosodiques de voix ont été choisies, neutre et triste, prononcées par 2 locutrices différentes. La prosodie de ces sons a aussi été validée dans une précédente étude comme étant bien reconnue (Charpentier, Kovarski, Houy-Durand, et al., 2018). Le choix des sons et de la séquence de stimulation a été fait à partir de l'étude préliminaire (Annexe 2). La figure 8 reprend la séquence de Roving utilisée dans l'étude n°2 ainsi qu'un exemple de sons vocaux et non-vocaux utilisés dans cette étude.

## Matériel et méthodes



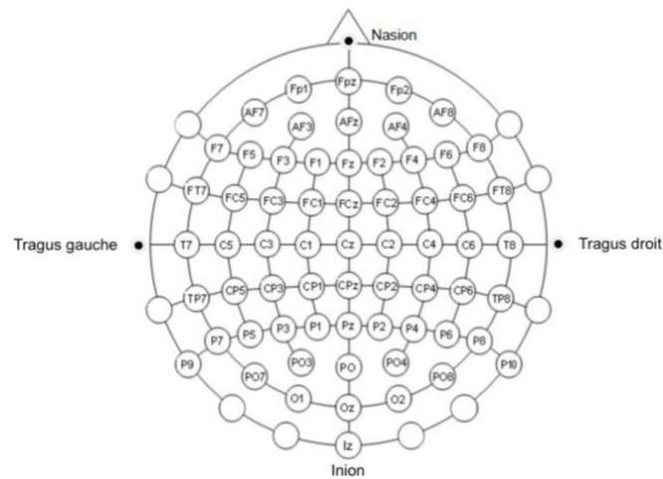
**Figure 8: Caractéristiques des sons et de la séquence utilisés pour l'étude en EEG de l'adaptation neurale**

A et B : spectrogramme d'un son vocal (A) et de son équivalent non-vocal (B). C et D : énergie de ces mêmes sons, selon les fréquences (C) et le temps (D). E : séquence de stimulation de l'étude n°2.

Pendant l'enregistrement les participants étaient confortablement installés dans un fauteuil dans une salle isolée acoustiquement et devaient rester les plus immobile possible pendant que les sons leur étaient présentés et qu'ils visionnaient un film muet (*Les Temps Modernes*, de Charlie Chaplin, 1936). Le visionnage du film permettait de maintenir l'attention des participants sur des informations visuelles afin qu'ils ne se concentrent pas sur les sons présentés (Näätänen, 2000).

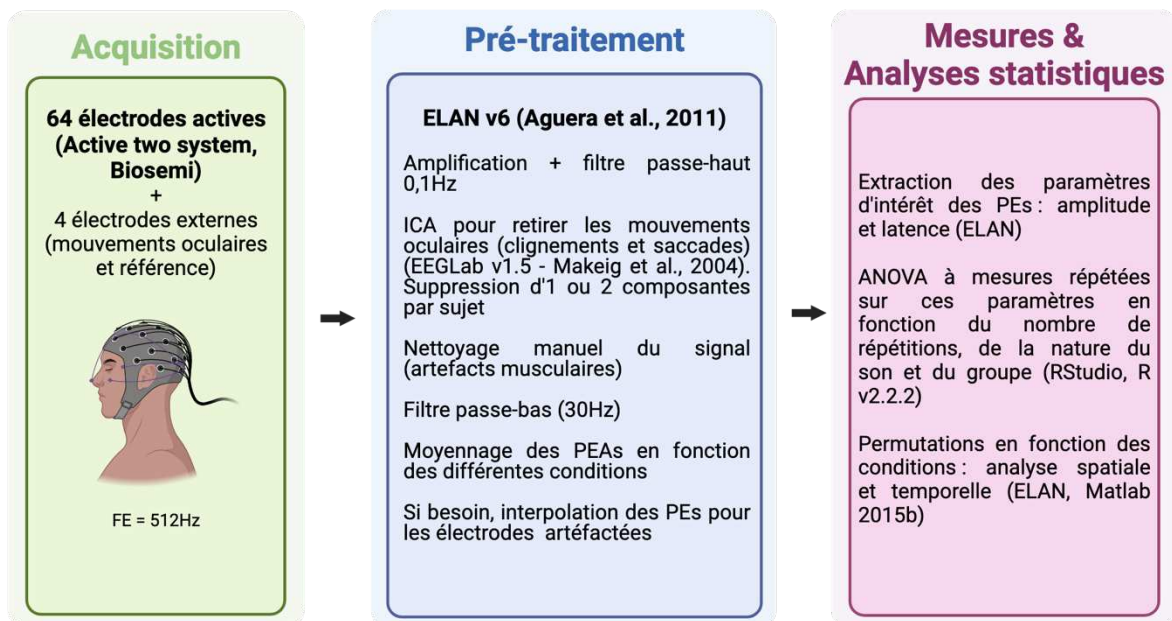
### 2.1.3. Acquisition et traitement du signal

Les enregistrements EEG ont été effectués avec le système à 64 électrodes actives de Biosémi (Active two systems, Biosemi®, The Netherlands) avec une fréquence d'échantillonnage de 512Hz et une bande passante de 0 à 104Hz. Le matériel utilisé se composait d'un bonnet en tissu ajusté au périmètre crânien du participant dans lequel étaient insérées 64 électrodes avec du gel conducteur. Afin de bien positionner le casque sur la tête du participant, l'électrode Cz servait de repère, elle devait être centrée selon l'axe inion-nasion et l'axe tragus gauche – tragus droit, selon le système 10-20 (Oostenveld & Praamstra, 2001) (figure 9). Trois électrodes externes étaient ajoutées dans le montage afin d'enregistrer les mouvements oculaires verticaux (clignement) et horizontaux (saccades). Enfin, une dernière électrode externe était placée sur le bout du nez servant de référence pour le montage. L'enregistrement EEG était couplé avec le système Présentation (NeuroBehavioral Systems®) qui permettait la délivrance des stimuli auditifs depuis un second ordinateur.



**Figure 9 : Positionnement des électrodes d'EEG dans le système 10-20**

Avant d'extraire les potentiels évoqués le signal EEG doit être prétraité. Les différentes étapes de ce pré-traitement sont présentées dans la figure 10.



**Figure 10 : Synthèse de la méthode d'acquisition et de pré-traitement des données EEG**  
 FE : fréquence d'échantillonnage ; PE : Potentiel Évoqué ; PEA : Potentiel Évoqué Auditif ; ICA : Independent Component Analysis

## 2.2. Corrélation inverse

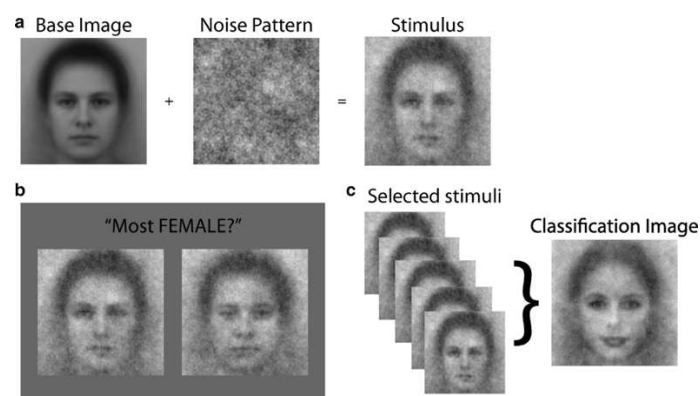
### 2.2.1. Principe de la corrélation inverse

La corrélation inverse (*Reverse Correlation*, RevCor) est une méthode de psychophysique permettant d'estimer les paramètres physiques (acoustiques ou visuels) associés à la

## Matériel et méthodes

représentation mentale d'une catégorie de stimuli (Brinkman et al., 2017). L'objectif de la corrélation inverse est d'identifier comment des variations spécifiques dans les stimuli influencent les réponses perceptives des participants. Le terme de modèle inverse est utilisé car plutôt que de partir des stimuli pour prédire les réponses, ce sont les réponses qui vont permettre d'inférer les aspects des stimuli qui ont permis de les classifier. On parle de corrélation inverse car contrairement à l'estimation d'une corrélation entre les paramètres et les réponses des participants, cette méthode mesure la corrélation entre une variable de réponse fixe et des stimuli aléatoires afin de modéliser les paramètres qui permettent d'établir le choix des participants (Todorov et al., 2011).

Cette méthode a été largement utilisée en visuel dans l'estimation de paramètres associés au genre (Dotsch et al., 2011; Mangini & Biederman, 2004; Nestor & Tarr, 2008), à l'identité (Mangini & Biederman, 2004; A. I. Young et al., 2014) ou encore à l'appartenance à un groupe (Brinkman et al., 2017; Brown-Iannuzzi et al., 2017; Dotsch et al., 2008; Imhoff et al., 2011, 2013). Cette technique consiste à présenter un grand nombre de stimuli à des participants. Ces stimuli sont tous créés à partir d'un même stimulus de base sur lequel du bruit a été ajouté. Les participants doivent ainsi comparer 2 stimuli pour chaque essai dans un choix forcé (2-AFC : *two Alternative Forced Choice*). Ainsi, après une multitude d'essais, une estimation du ou des paramètres utilisés pour classifier le stimulus peut être faite. Le résultat est appelé « image de classification » pour la modalité visuelle (figure 11).



**Figure 11 : Exemple de paradigme de corrélation inverse en visuel**  
« Quelle image représente le plus une femme ? » (Brinkman et al., 2019)

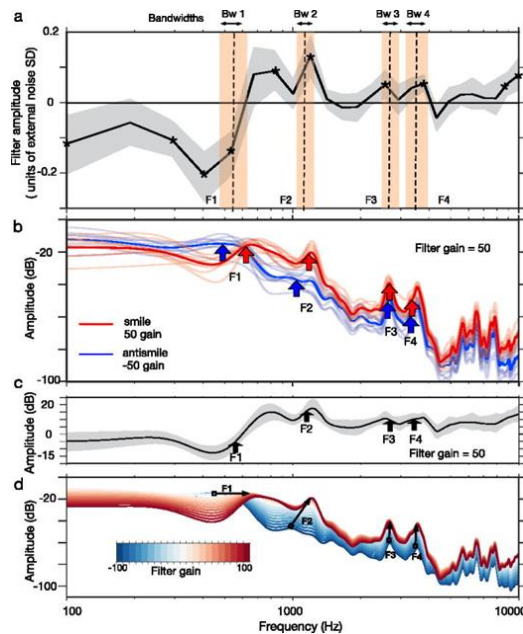
La corrélation inverse peut aussi être utilisée dans la caractérisation et l'étude de paramètres acoustiques (Ponsot, Arias, et al., 2018; Ponsot, Burred, et al., 2018; L. Wang et al., 2022) et

## Matériel et méthodes

notamment pour l'estimation de la représentation d'indices prosodiques (Ponsot, Arias, et al., 2018; Ponsot, Burred, et al., 2018) (figure 12). Dans ces protocoles de corrélation inverse en psychoacoustique, des sons sont modifiés à l'aide de filtres manipulant les caractéristiques spectrales. Contrairement aux études visuelles, du bruit n'est pas ajouté au signal. À la place, ce sont des modulations acoustiques aléatoires qui sont appliquées. Le participant doit ensuite choisir, parmi deux sons, celui qui correspond le mieux à la représentation sonore demandée. Le moyennage des filtres choisis par rapport à ceux non choisis permet d'extraire les paramètres acoustiques associés à la reconnaissance du sourire dans la voix, révélant ainsi la représentation spectrale propre à chaque participant. Cette méthode a permis de développer un modèle robuste du sourire vocal chez les adultes neurotypiques (Ponsot, Arias, et al., 2018). Cette méthode a trois grands avantages repris dans une étude récente (Goupil et al., 2021) :

- Elle permet d'étudier plus de paramètres acoustiques pour une tâche donnée en s'affranchissant des effets de production limités par le petit nombre de stimuli produits par l'expérimentateur ou des acteurs.
- Elle permet de tester sans biais et sans contrainte la représentation d'un sujet sur le paramètre testé.
- Elle permet le calcul d'un modèle dans lequel le sujet compare les exemples qu'on lui présente à sa propre représentation interne du paramètre testé.

À partir de ces essais un modèle de sourire vocal peut être extrait pour ensuite l'appliquer à d'autres sons pour qu'ils sonnent plus « ouverts », simulant l'effet d'une extension des zygomatiques qui étireraient les lèvres. Avec cette méthode il n'y a pas de modulation de la fréquence fondamentale (hauteur du son), on constate plutôt un décalage des résonances (formants) vers le haut (Ponsot, Arias, et al., 2018; Pruvost-Robieux, 2023). L'application du modèle à d'autres sons permet de tester les effets du sourire sur différentes réponses physiologiques.



**Figure 12 : Modèle du sourire vocal à partir de la corrélation inverse Ponsot et al. (2018)**

a : Filtre moyen associé au jugement du sourire vocal, tel que dérivé par la corrélation inverse. Les astérisques indiquent des différences significatives par rapport à 0 (test t bilatéral pour échantillons appariés,  $p < 0,05$ ). Les zones verticales ombrées indiquent comment les quatre premiers formants de la voix s'alignent avec la structure du filtre. b : Lorsque ce filtre ou son opposé est appliqué (ici avec un gain de 50) à la voix originale, il révèle les représentations auditives internes d'une voix souriante et d'une voix non souriante. Les lignes ombrées représentent l'enveloppe spectrale correspondante pour le filtre interne de chaque participant. c : Différence moyenne de l'enveloppe spectrale entre les sons souriants et non souriants pour un gain de filtre de 50. d : Enveloppe spectrale moyenne à travers les participants pour différents gains de filtre, mettant en évidence la transformation globale des enveloppes spectrales lorsqu'on passe des représentations mentales d'une voix fortement non souriante à celle d'une voix fortement souriante. Les zones ombrées représentent les intervalles de confiance à 95 %, calculés à l'aide d'une procédure de Bootstrap.

### 2.2.2. Bruit interne

La méthode de la corrélation inverse offre aussi l'opportunité d'estimer le bruit interne associé à une représentation. En effet, dans toutes les situations il existe un bruit interne, intra-personnel, qui peut biaiser la représentation de la réalité (Faisal et al., 2008). Une manière robuste et stricte d'estimer ce bruit est la méthode de la « double-passe ». Cette méthode consiste à présenter deux fois la même séquence de stimulation et de mesurer la cohérence des réponses entre les 2 présentations chez un même individu. Un observateur idéal, qui ne présenterait pas de bruit interne, répondrait alors 2 fois la même chose. Un observateur réel peut répondre différemment à la seconde présentation de stimuli identiques. À partir de ces différentes réponses, la magnitude du bruit interne peut être estimée (Neri, 2010). Le fait de répéter plusieurs fois les mêmes questions pour des mêmes stimuli permet ainsi d'estimer la variabilité des réponses pour un même individu, cette valeur de bruit prenant aussi en compte

## Matériel et méthodes

la persévération, c'est-à-dire la tendance à plus choisir comme souriant un son présenté en premier ou en second au sein des paires. Dans les études originales de corrélation inverse, les images étaient modifiées par l'ajout d'un bruit qualifié de « bruit externe » et l'objectif de ces études était d'étudier l'influence de ce bruit sur l'utilisation d'indices particuliers pour classer une image. À partir de ce bruit externe, afin d'étudier la fiabilité des résultats dans ces études, un paradigme de double-passe (une partie des essais est répété 2 fois afin de mesurer la cohérence des réponses chez les participants) est mis en place, permettant d'avoir une estimation du bruit interne. Le bruit interne est alors exprimé en déviation standard du bruit externe. Une estimation de la valeur du bruit interne dans un groupe d'adultes neurotypiques est en moyenne, de 1,4 fois la magnitude la variabilité induite par une source externe de bruit, dans la modalité visuelle (Neri, 2010). Il n'existe pas de valeur de référence pour des adultes autistes car peu d'études ont utilisé cette méthode pour estimer le bruit interne dans cette population (Vilidaite et al., 2017).

Dans les études de corrélation inverse dans la modalité auditive, comme mentionné précédemment, ce ne sont pas du bruit qui est ajouté aux sons, mais des modulations du signal acoustique qui sont effectuées. Cela rend difficile l'estimation du bruit interne en fonction du bruit externe. Afin d'avoir une estimation de ce bruit dans ce type de paradigme, il est possible de modéliser un participant linéaire idéal avec toutes les valeurs possibles de bruit (selon la fiabilité des réponses et le biais de réponse dans un 2-AFC) et ainsi d'avoir une valeur de bruit interne en déviation standard du bruit externe à partir de ces essais répétés (Adl Zarrabi et al., 2024).

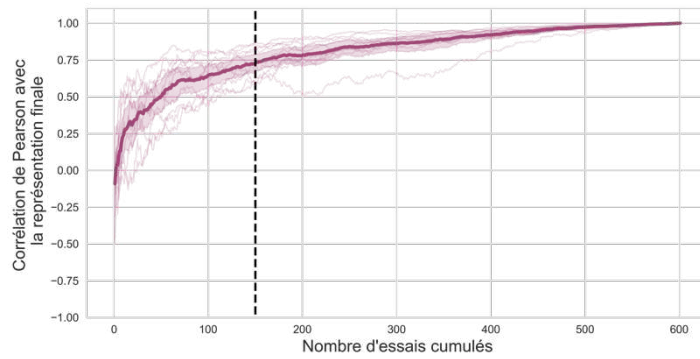
### **2.2.3. Création des sons et décours expérimental**

Afin d'estimer la représentation du sourire vocal dans un groupe d'adultes autistes et un groupe d'adultes neurotypiques, la méthode de la corrélation inverse a été utilisée. Pour cela, un son de base neutre a été choisi. Il s'agissait d'un son /a/ de 500ms prononcé par une voix d'homme et pour lequel le pitch a été aplati, c'est-à-dire que toutes les modulations prosodiques de ce son ont été retirées afin qu'il soit composé des mêmes caractéristiques fréquentielles tout au long des 500 ms, et le rendant le plus « neutre » possible. Les modifications spectrales ont ensuite été appliquées à l'aide de la boîte à outil CLEESE (Burred et al., 2019). Le son a ainsi été modifié sur 25 fréquences différentes entre 100 et 10000 Hz. Les modifications correspondent

## Matériel et méthodes

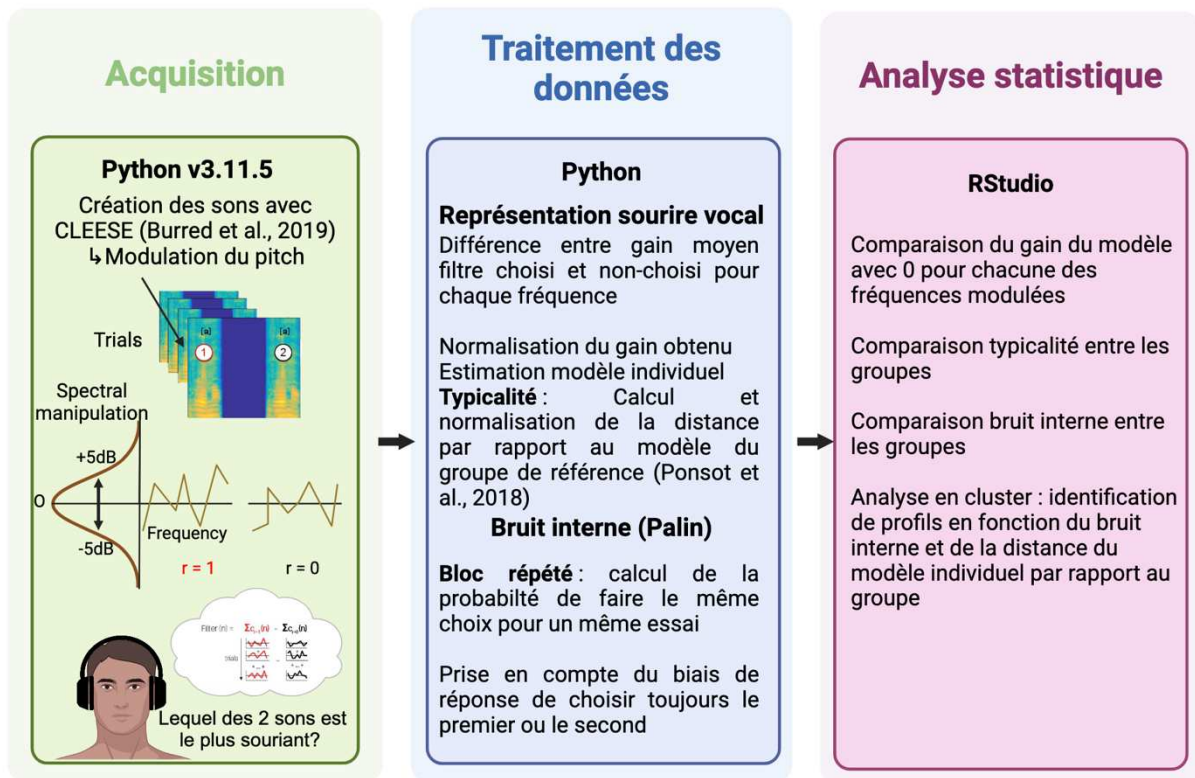
à un gain en amplitude aléatoire de plus ou moins 5dB selon une gaussienne sur chacune de ces fréquences. Un panel de sons a été alors créé et ces sons ont été présentés de manière aléatoire par paire au participant afin qu'il choisisse celui qui lui semblait le plus souriant. À partir des réponses, les modulations du filtre non-choisi comme souriant étaient soustraites de celles du filtre choisi. Le moyennage de tous les essais nous a ainsi permis d'obtenir un *kernel* (i.e. filtre moyen souriant sur chacune des fréquences modifiées) pour chaque participant puis un modèle à l'échelle du groupe (figure 12). Les kernels ainsi obtenus peuvent être appliqués sur le son de base, ce qui permet d'écouter mais aussi d'évaluer les modulations de fréquences et d'amplitude qui sont caractéristiques de la représentation du sourire vocal pour un individu ou un groupe d'individus.

Le nombre d'essais a été réduit, par rapport à la tâche originale de Ponsot et al. (2018), de 600 à 150 essais après une analyse de convergence qui a permis d'affirmer que les résultats étaient fiables avec ce nombre réduit d'essais. Cette analyse de convergence a consisté à calculer les coefficients de corrélation de Pearson entre les gains pour chacune des fréquences après 600 essais et après chacun des 600 essais. Nous avons considéré qu'une corrélation ( $r$  de Pearson) supérieure ou égale à 0,75 était suffisante pour considérer les représentations comme similaires. Après 150 essais les représentations étaient suffisamment corrélées pour être considérées comme semblables (figure 13). Le nombre d'essais proposés aux patients autistes (150) et au groupe contrôle (250) était différent afin que la tâche soit adaptée à chacune des populations. Les comparaisons entre les groupes (autistes et neurotypiques), dans l'étude n°3, ont été faites sur le même nombre de répétitions, en prenant en compte les 150 premiers essais. La tâche était divisée en plusieurs blocs de 50 essais, 4 blocs étaient présentés aux adultes autistes et 6 blocs aux adultes neurotypiques. Deux blocs identiques étaient présentés aux participants à la fin de la séquence afin d'estimer le bruit interne selon la méthode de la double-passe.



**Figure 13 : Résultats analyse convergence RevCor à partir des données de Ponsot et al. (2018)**  
 Chaque ligne fine représente la corrélation entre le modèle après  $n$  essais et le modèle après 600 essais pour chaque participant. La ligne épaisse représente la moyenne. La ligne verticale à 150 essais représente le nombre d'essais choisis.

La figure 14 reprend les différentes étapes d'acquisition et de traitement des données de corrélation inverse.



**Figure 14 : Synthèse de la méthode d'acquisition et de traitement des données de corrélation inverse**

### **2.3. Électromyographie faciale et pupillométrie**

Afin d'étudier la dernière étape de la boucle, la résonance motrice et la contagion émotionnelle, deux techniques d'exploration physiologique ont été utilisées simultanément, l'électromyographie faciale (fEMG) et la pupillométrie, en réponse à des phrases rendues plus ou moins souriantes à partir d'un filtre artificiel.

#### **2.3.1. Principes de l'EMG et de la pupillométrie**

##### 2.3.1.1. Électromyographie faciale (fEMG)

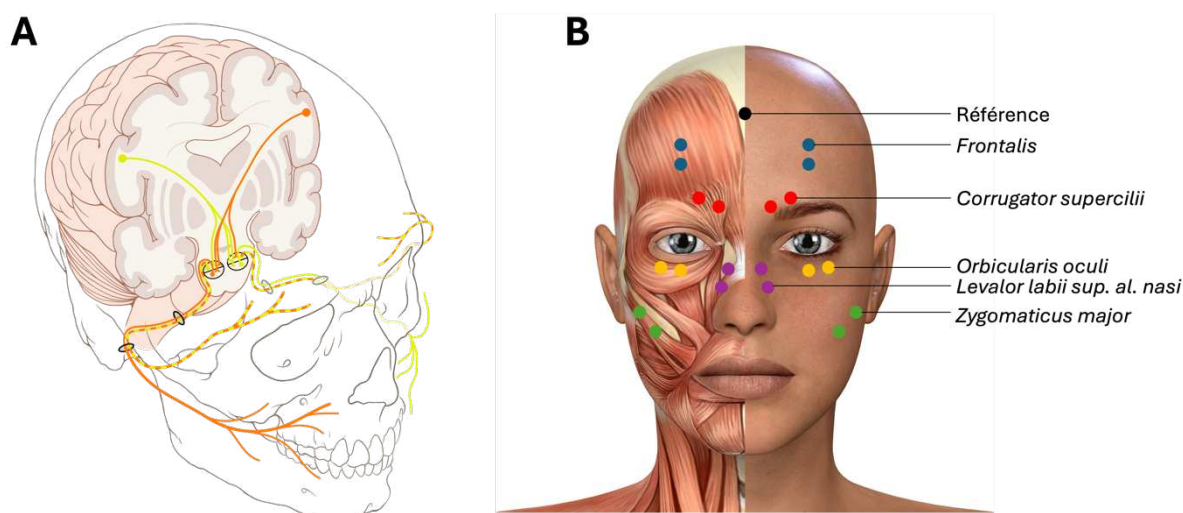
Ce paragraphe est inspiré du chapitre 8 « Électromyographie (EMG) » de Korb et Delplanque dans *Électrophysiologie de la cognition* (Korb & Delplanque, 2013).

L'EMG permet de mesurer le fonctionnement des muscles striés squelettiques. Les muscles squelettiques sont composés de fibres musculaires, chacune constituée de myofibrilles, elles-mêmes formées par un ensemble de sarcomères liés par des disques Z. Dans les sarcomères on retrouve des myofilaments parallèles, avec en alternance des filaments fins d'actine et large de myosine. Lors d'une contraction musculaire, les liens entre les filaments suite à une dépolarisation et à des mouvements d'ions calcium ( $Ca^{2+}$ ) provoquant un coulisement entre les filaments d'actine et de myosine, et donc des disques Z. Tous les sarcomères d'une même fibre se contractent ensemble dans la même direction causant une contraction du muscle entier. Pour observer une contraction musculaire à l'aide de l'EMG plusieurs étapes sont nécessaires. La contraction dépend de la décharge des neurones moteurs (motoneurones) qui innervent les fibres musculaires. Les motoneurones sont constitués d'unités motrices associées à des fibres musculaires par les fibrilles (terminaisons axonales) et formant le nerf moteur. Quand le motoneurone décharge, le potentiel d'action se diffuse à partir de la fibrille causant la libération d'acétylcholine dans la fente synaptique, provoquant la dépolarisation de la myofibrille et donc la contraction musculaire. Une partie de ces changements électriques liés à la contraction musculaire est transmise jusqu'à la peau à travers les fluides et les espaces extracellulaires. Ce sont ces modulations de potentiel qui sont mesurées à l'aide de l'EMG de surface.

L'innervation des muscles de la face se fait en grande partie par le nerf facial (VII), un nerf avec plusieurs rôles (contrôle des muscles de la face, sensations cutanées...). Le nerf facial fait

## Matériel et méthodes

partie des 12 paires de nerfs crâniens qui proviennent du tronc cérébral. Le nerf facial se dissocie ensuite en 5 branches : temporale, frontale, palpébrale, nasale et buccale. Les branches qui nous intéressent ici sont les branches temporales et frontales qui permettent l'innervation de la plupart des muscles de la face permettant d'exprimer des émotions. Ces branches ont un déploiement ipsi- (branche frontale) et contra-latéral (branche temporale et frontale). Dans les études présentées dans cette thèse, les muscles d'intérêt sont le *Zygomaticus major* (ZM) et le *Corrugator supercilii* (CS) qui dépendent respectivement des branches temporale et frontale. L'anatomie des muscles de la face et de l'innervation du nerf facial sont reprises dans la figure 15.



**Figure 15 : Innervation et anatomie des muscles de la face**

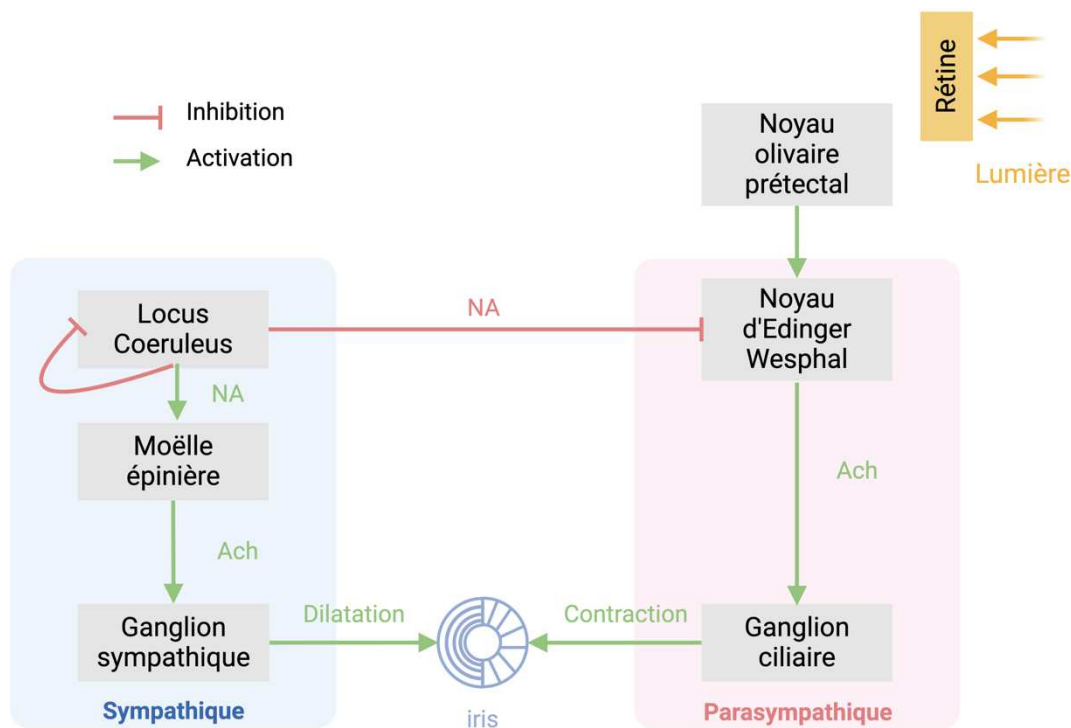
*A : Représentation schématique de la partie motrice du nerf facial. En jaune, les fibres issues de la partie droite du cerveau ; en orange, celles issues de la partie gauche. Noter que la partie inférieure de la face reçoit sa commande uniquement de la partie du cerveau controlatérale alors que la partie supérieure la reçoit des deux côtés B : Placement des électrodes de fEMG selon l'anatomie du visage (Boxtel, 2010; van Boxtel & van der Graaff, 2024)*

### 2.3.1.2. Pupillométrie

La pupille est un indice qui permet de mesurer la réactivité du Système Nerveux Autonome (SNA) dans ses branches sympathique et parasympathique (Beatty & Lucero-Wagoner, 2000). En effet, la dilatation est le reflet d'une activation sympathique due à la contraction du muscle lisse dilatateur de l'iris via de la noradrénaline qui se fixe sur les récepteurs  $\alpha_1$  adrénergique. Quant à la branche parasympathique elle contrôle la constriction de la pupille à travers la contraction du muscle sphincter de l'iris quand l'acétylcholine se fixe sur les récepteurs muscariniques. Les muscles qui permettent la constriction et la dilatation de la pupille sont des

muscles antagonistes qui suivent la loi d'innervation réciproque de Sherrington, l'excitation de l'un s'accompagne de l'inhibition de l'autre (Kanski & Bowling, 2015) (figure 16)

Les réponses pupillaires à des stimuli psychosensoriels sont exclusivement contrôlées par le système sympathique (Wilhelm, 2008). En effet, le système sympathique central inhibe le noyau d'Edinger-Westphal situé dans le mésencéphale, qui appartient au système parasympathique. Cette inhibition s'exerce par deux voies simultanées. La première part du tronc cérébral et atteint le noyau d'Edinger-Westphal via l'hypothalamus. La seconde provient directement du *Locus Coeruleus*, un noyau noradrénergique, et se dirige directement vers le noyau d'Edinger-Westphal.



**Figure 16: Schéma de l'innervation sympathique et parasympathique de la pupille**  
 D'après Wilhelm, conférence pupillométrie d'Uppsala, 2010

La mesure des modulations du diamètre pupillaire informe donc sur l'état interne d'une personne à travers l'activation (ou la désactivation) du système sympathique siège de la réponse autonome et notamment du réflexe photomoteur, de l'éveil, de l'attention et de la réponse aux émotions (Mathôt, 2018). De plus, cette mesure ne nécessite aucun contact physique entre l'expérimentateur et le participant contrairement à l'EEG ou l'fEMG qui nécessitent la pose des électrodes. Cette absence de contact pour effectuer la mesure est un

atout pour étudier ces phénomènes chez des personnes autistes et/ou avec des atypies sensorielles, en particulier tactiles, qui empêchent parfois la pose d'électrodes directement en contact avec la peau.

### 2.3.2. Séquence de stimulation

À partir de la tâche en corrélation inverse, le modèle du sourire vocal de Ponsot et al. (2018) a été appliqué à des phrases pour les rendre souriants ou non-souriants. Les mêmes sons qu'Arias et al. (2021; 2018), ont été utilisés dans ce protocole. Le corpus utilisé était composé de 5 phrases différentes, d'une durée de 1,9s en moyenne, avec un contenu sémantique neutre permettant de tester la prosodie émotionnelle (Russ et al., 2008). Les voix de 10 locuteurs différents ont été utilisées (5 ♀ et 5 ♂). Le tableau 3 reprend les caractéristiques de chacune des phrases.

Contenu sémantique	Nombre d'occurrences	Genre des locuteurs
Je suis en route pour la réunion	4	2 ♀ 2 ♂
L'avion est presque plein	7	4 ♀ 3 ♂
J'aimerais avoir un nouveau réveil	6	3 ♀ 3 ♂
N'oublie pas de prendre une veste	2	1 ♀ 1 ♂
Nous nous arrêterons dans quelques instants	1	1 ♂

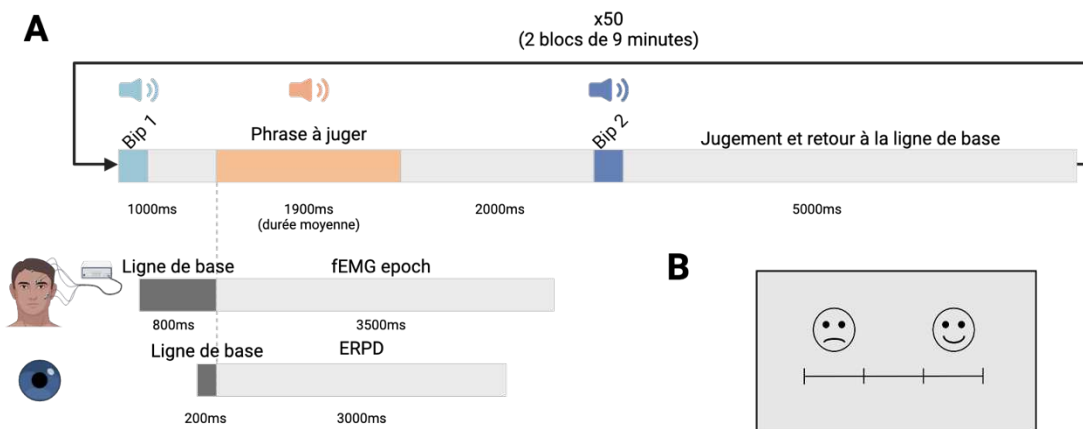
**Tableau 3 : Récapitulatif et caractéristiques des phrases utilisées dans études n°4 et 5**

Ces phrases ont ensuite été filtrées afin de modifier artificiellement, mais de manière contrôlée, la prosodie émotionnelle. En condition souriante, un filtre souriant issu directement la tâche de corrélation inverse a été appliqué, alors qu'il s'agissait de l'inverse arithmétique de ce filtre pour la condition non-souriante. Afin de tester les effets de ces modulations prosodiques sur l'activité musculaire et sur la pupille les phrases modifiées étaient répétées 40 fois pour chacun des filtres et 20 fois non-modifiées (séquence totale de 100 phrases). La présence des phrases non-modifiées avait pour but de rendre la tâche moins monotone et d'éviter les effets d'habituation et d'apprentissage. L'ordre de présentation des phrases était pseudo-aléatoire mais devait répondre à certaines conditions :

- Pas 2 phrases à la suite du même locuteur,
- Pas 2 phrases à la suite avec le même contenu sémantique,
- Pas 3 phrases à la suite avec le même filtre,

- Pas 3 phrases à la suite avec un locuteur du même genre.

La séquence a été divisée en 2 blocs de 50 phrases chacun afin de minimiser les effets d'habituation à la séquence et pour limiter les effets de fatigue sur le diamètre de la pupille (Hopstaken et al., 2015). L'enregistrement de la pupille a induit de nouvelles contraintes dans la mise en place de cette tâche. En effet, la pupille est sensible aux variations de luminosité, et par conséquent au changement d'images projetées. Pour pallier ce problème la même image a été projetée tout le long de la tâche. L'image choisie correspond à une échelle de jugement à 4 niveaux qui permet d'estimer le degré de sourire dans la phrase. Pour indiquer les différentes étapes de la tâche au participant des sons purs ont été utilisés : un premier son (440Hz) pour signifier le début de l'essai et un second (275Hz) après l'écoute des phrases et spécifiant le moment de donner son jugement. Le dérouls temporel de cette expérience est décrit sur la figure 17. La durée de chaque bloc était alors de 9 minutes avec une durée totale de la tâche d'environ 18min.



**Figure 17: Dérouls de la tâche dans le protocole de mesure de pupillométrie et fEMG et échelle de jugement du degré de sourire projetée à l'écran**

Comme pour l'étude en EEG une étude préliminaire à celles présentées dans cette thèse (études n°4 et 5) a été réalisée pour valider le protocole et tester la faisabilité de celui-ci. Cette étude préliminaire utilisait la même séquence de stimulation et le même dérouls expérimental mais une autre échelle de mesure à trois niveaux était utilisée. Un groupe d'une vingtaine d'adultes neurotypiques a été inclus dans cette étude préliminaire, les résultats sont présentés en annexe 3.

### **2.3.3. Acquisition et traitement du signal**

#### 2.3.3.1. Électromyographie faciale (fEMG)

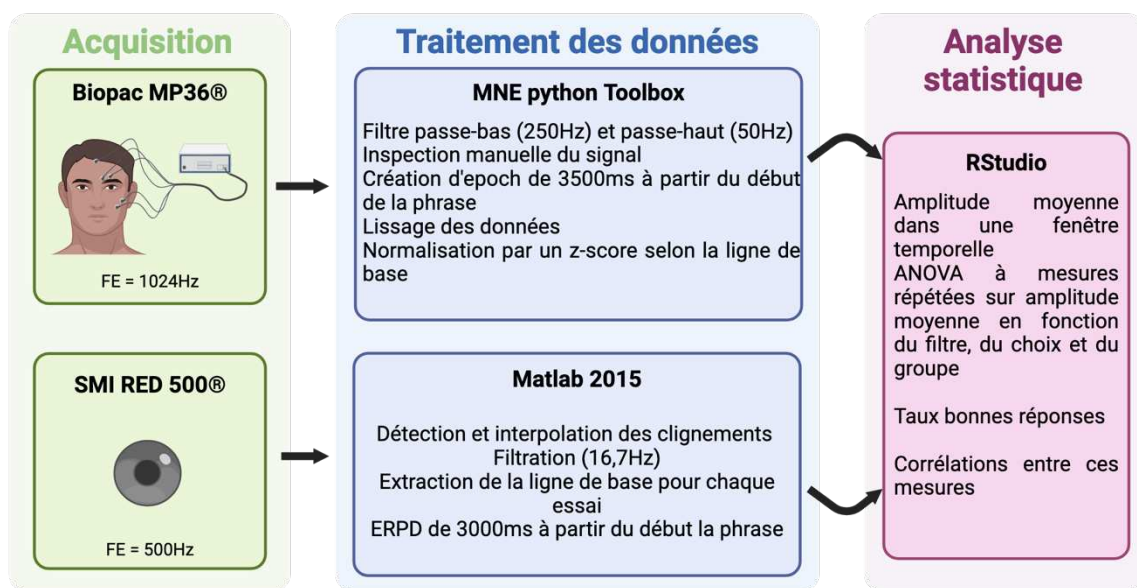
L'EMG facial a été mesuré l'aide du système d'acquisition Biopac MP36® avec des électrodes Ag-AgCl de 4mm (EL254 et EL254S) et le logiciel AcqKnowledge. Afin de mesurer la réactivité faciale émotionnelle à des sons les électrodes ont été placées sur le côté droit du visage (Fridlund & Cacioppo, 1986) de manière à mesurer l'activité du muscle *Zygomaticus major* (ZM) et du *Corrugator supercilii* (CS) (figure 15). Pour mesurer la contraction musculaire, deux électrodes (cathode et anode) étaient placées sur chacun des muscles, séparées d'approximativement 1cm dans le sens du muscle afin de mesurer la différence de polarité entre celles-ci. Pour vérifier le placement des électrodes et pour s'assurer de mesurer l'activité des muscles d'intérêt, en particulier le muscle ZM qui présente un recouvrement important avec d'autres muscles (Cacioppo et al., 2000; Ekman & Friesen, 1975; Rinn, 1984). Avant de commencer le paradigme, pour vérifier le placement des électrodes, il était demandé au participant de produire la voyelle /i/ pour mesurer l'activité du ZM. Une électrode de référence était ajoutée au milieu du front à la racine des cheveux. Le signal a été enregistré avec une fréquence d'échantillonnage de 1024Hz.

La méthode utilisée pour l'analyse et le traitement des données est présentée dans la figure 18.

#### 2.3.3.2. Pupillométrie

Les mesures oculométriques et pupillométriques ont été effectuées avec le système SMI RED 500® avec une fréquence d'échantillonnage de 500Hz. Les systèmes d'eye-tracking sont composés de 2 éléments principaux : un émetteur de lumière infrarouge et une caméra infrarouge. Afin de mesurer les fluctuations du diamètre de la pupille, l'appareil projette une lumière infrarouge ( $\lambda = 870\text{nm}$ ) ; cette lumière est ensuite reflétée par la rétine et la caméra va ainsi récupérer l'information lumineuse qui lui sera transmise. Avec ces informations le système peut ainsi calculer la taille de la pupille (qui est plus sombre que l'iris et la sclérotique) mais aussi la position et les déplacements du regard. Une lumière infrarouge est utilisée car elle correspond à des fréquences qui n'appartiennent pas au spectre visible par l'œil humain, par conséquent cela ne vient pas distraire ou perturber le participant durant le protocole en cours.

Les modulations du diamètre de la pupille ont largement été étudiées dans le cadre du réflexe photomoteur (Croft, 1961; Safran & Roth, 1982). Afin de mesurer l'effet d'une stimulation sur le diamètre de la pupille, il est important de calculer une ligne de base<sup>5</sup>, puis d'évaluer la modulation de l'activité par rapport à cette ligne de base en fonction du moment de la présentation du stimulus (Seropian et al., 2022). Dans le cadre de cette thèse, la mesure de la dilatation pupillaire en réponse aux phrases modifiées artificiellement pour paraître souriantes ou non a été réalisée. Afin d'avoir des données exploitables en pupillométrie, une étape d'extrapolation des données perdues pendant les clignements des yeux est indispensable. Si les clignements sont correctement caractérisés, une interpolation cubique des données est effectuée en utilisant la médiane des valeurs du diamètre avant et après le clignement. Ainsi, les données ne sont pas perdues et d'avantage d'essais peuvent être conservés dans les analyses.



**Figure 18 : Synthèse de la méthode d'acquisition et de pré-traitement des données de pupillométrie et fEMG**

*FE : fréquence d'échantillonnage ; ERPD : Event Related Pupil Dilation*

<sup>5</sup> La ligne de base correspond à la mesure de l'activité, ici pupillométrie, pendant quelques millisecondes avant la présentation d'un stimulus pour contrôler les variations initiales et identifier les modulations induites par le stimulus présenté. Cela permet de normaliser les données et permet une comparaison précise des réponses physiologiques à des stimuli spécifiques.



## **Résultats expérimentaux**

---



## **1. Adaptation neurale à la prosodie émotionnelle**

### **1.1. Avant-propos**

Dans l'autisme une forme de recherche d'immuabilité est observée avec une résistance au changement qui s'accompagne au niveau physiologique d'atypies dans la détection du changement. Les difficultés d'adaptation au contexte et à l'environnement observées dans l'autisme pourraient être liées à des particularités neurophysiologiques. En effet, une mauvaise adaptation à un contexte régulier ne permet pas de percevoir et d'inférer correctement les changements. Cette hypothèse est testée dans le cadre de cette thèse, en réponse à des sons vocaux émotionnels ou neutres. En effet, des différences dans les processus perceptifs de traitement de la voix et de la prosodie émotionnelle pourraient contribuer aux atypies de réponses comportementales aux émotions.

Dans cette partie expérimentale deux études ont ainsi été menées pour caractériser l'adaptation neurale :

- L'étude n°1 a été menée dans un groupe d'adultes neurotypiques, afin d'établir une référence à laquelle comparer les résultats chez des adultes autistes et pour valider le protocole expérimental
- L'étude n°2 s'intéresse à l'adaptation neurale à des sons vocaux, non-vocaux et émotionnels dans l'autisme.

Ces deux études permettront de caractériser l'adaptation neurale en fonction de la nature des sons dans l'autisme et chez l'adulte neurotypique.

## 1.2. Étude n°1 : Adaptation neurale à des sons vocaux et non-vocaux chez l'adulte neurotypique

### Neural repetition suppression to vocal and non-vocal sounds

Heurteloup, C<sup>\*</sup>, **Merchie, A<sup>\*</sup>**, Roux, S., Bonnet-Brilhault, F., Escera, C., & Gomot, M.

<sup>\*</sup>:co-premières autrices

Article publié dans *Cortex* (2022)

**Introduction et objectifs :** L'anticipation et l'adaptation aux stimuli de l'environnement sont essentielles dans la vie quotidienne, notamment, dans les interactions et la communication avec les autres. C'est pourquoi il est intéressant d'étudier l'adaptation neurale à des sons vocaux et non-vocaux. L'adaptation neurale est mesurée au niveau neurophysiologique à travers un indice, la *Repetition Positivity* (RP). Cet indice reflète l'augmentation de l'amplitude des composantes neurophysiologiques en réponse à une augmentation du nombre de répétitions d'un même son.

**Matériel et méthodes :** Les réponses neurophysiologiques à des sons vocaux et non-vocaux ont été mesurées chez 20 adultes neurotypiques à l'aide d'un enregistrement EEG. Deux séquences de stimulation (*Roving*) ont été présentées aux participants avec des trains de 4, 8 ou 16 répétitions d'un même son. Une séquence était constituée de sons uniquement vocaux (8 voyelles /a/ prononcées par des locutrices différentes) tandis que l'autre séquence comportait seulement des sons non-vocaux (sons complexes reprenant les caractéristiques acoustiques des vocaux). Les réponses cérébrales ont été moyennées selon des groupes de répétitions pour mesurer les dynamiques d'adaptation en fonction de la nature des sons (2-4, 5-8, 9-12, 13-16) et les mesures d'amplitude et de latence ont ensuite été réalisées sur la composante P1. Les RP et les *Mismatch Negativity* (MMN) ont aussi été calculées pour étudier la résultante de l'adaptation après de nombreuses répétitions et la détection du changement en fonction de la nature d'un son.

**Résultats :** Une augmentation de l'amplitude de la P1 avec le nombre de répétitions est observée en réponse aux sons vocaux et non-vocaux. Une RP a aussi été mise en évidence avec une amplitude similaire en réponse aux sons vocaux et non-vocaux. Ainsi, les caractéristiques de l'adaptation neurale en réponse à des sons vocaux et non-vocaux sont proches. Néanmoins, la dynamique d'adaptation n'est pas la même pour atteindre une résultante similaire. En effet, 5 à 8 répétitions sont suffisantes pour observer une adaptation en réponse aux sons non-vocaux tandis qu'il faut au moins 9 à 12 répétitions pour les sons vocaux. Enfin, comme supposé, un effet du nombre de répétitions sur l'amplitude de la MMN a été montré. L'amplitude de la MMN s'accroît à mesure que le nombre de répétitions précédant un changement augmente, et cela pour les deux types de sons.

**Discussion :** Cette étude a permis de caractériser pour la première fois l'adaptation neurale à des sons complexes et sociaux avec un protocole adapté. L'adaptation plus lente à des sons vocaux qu'à des sons non-vocaux est surprenante chez des adultes neurotypiques, pour qui la voix est un son plus naturel et familier que les sons complexes. Il semblerait donc que la complexité et la richesse acoustique des sons à encoder ralentissent ce processus : plus un son est riche acoustiquement, plus il faudra de répétitions pour s'y adapter. Enfin, l'effet du nombre de répétitions sur l'amplitude de la MMN montre bien que la détection d'un changement dépend du contexte qui le précède. Les atypies de MMN observées dans l'autisme sont donc à discuter au regard du processus d'adaptation neurale.



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

Journal homepage: [www.elsevier.com/locate/cortex](http://www.elsevier.com/locate/cortex)

## Research Report

## Neural repetition suppression to vocal and non-vocal sounds

Camille Heurteloup<sup>a,1</sup>, Annabelle Merchie<sup>a,1</sup>, Sylvie Roux<sup>a</sup>,  
Frédérique Bonnet-Brilhault<sup>a</sup>, Carles Escera<sup>b,c,d,2</sup> and Marie Gomot<sup>a,\*,2</sup><sup>a</sup> UMR1253 iBrain, INSERM, University of Tours, France<sup>b</sup> Brainlab Cognitive Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, University of Barcelona, Catalonia-Spain<sup>c</sup> Institute of Neurosciences, University of Barcelona, Catalonia-Spain<sup>d</sup> Institut de Recerca Sant Joan de Déu (IRSJD), Catalonia-Spain

## ARTICLE INFO

## Article history:

Received 8 December 2020

Reviewed 31 January 2021

Revised 2 March 2021

Accepted 19 November 2021

Action editor Sonja Kotz

Published online 30 December 2021

## Keywords:

Vocal

Neural adaptation

Repetition positivity

Habituation

Prediction

ERPs

## ABSTRACT

Adaptation to the sensory environment is essential in everyday life, to anticipate future events and quickly detect and respond to changes; and to distinguish vocal variations in congeners, for communication. The aim of the current study was to explore the effects of the nature (vocal/non-vocal) of the information to be encoded, on the establishment of auditory regularities. In electrophysiology, neural adaptation is measured by the 'Repetition Positivity' (RP), which refers to an increase in positive potential, with the increasing number of repetitions of a same stimulus. The RP results from the combined variation of several ERP components; the P1, the first positivity (~100 ms) may reflect the onset of repetition effects. We recorded auditory evoked potentials during a roving paradigm in which trains of 4, 8 or 16 repetitions of the same stimulus were presented. Sequences of vocal and non-vocal complex stimuli were delivered, to study the influence of the type of stimulation on the characteristics of the brain responses. The P1 to each train length, and the RP responses were recorded between 90 and 200 ms, reflecting adaptation for both vocal and non-vocal stimuli. RP was not different between vocal and non-vocal sequences (in latency, amplitude and spatial organization) and was found to be similar to that found in previous studies using pure tones, suggesting that the repetition suppression phenomena is somehow independent of the nature of the stimulus. However, results showed faster stabilization of the P1 amplitude for non-vocal stimuli than for vocal stimuli, which require more repetitions. This revealed different dynamics for the establishment of regularity encoding for non-vocal and vocal stimuli, indicating that the richness of vocal sounds may require further processing before full neural adaptation occurs.

© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\* Corresponding author. UMR1253, INSERM, Université de Tours, Centre de Pédiopsychiatrie, CHRU Bretonneau 2 Boulevard Tonnellé, 37044 TOURS Cedex 09 France.

E-mail address: [gomot@univ-tours.fr](mailto:gomot@univ-tours.fr) (M. Gomot).

<sup>1</sup> Co-First-authors and.

<sup>2</sup> Co-Last authors.

<https://doi.org/10.1016/j.cortex.2021.11.020>

0010-9452/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

We encounter many sensory regularities in everyday life through sensory adaptation, whether simply at the level of word composition, or more complex scene schemas (Turk-Browne et al., 2009). At the cerebral level, information must be processed optimally to select the most relevant events. This requires building and continuously updating a sensory memory trace, following the presentation of a stimulus (Winkler et al., 2009). A repeated stimulus will be considered as regular, leading to the adaptation of the response, through the habituation process. Such storage is an essential property of the sensory systems, in order to process the incoming flow of information. Identifying the probability that the previous stimulus will reappear, enables the prediction of future events (Winkler et al., 2001, 2009). This makes the detection of environmental changes faster, improving our ability to react quickly and optimally. It also underlies our ability to detect modulations of social indices like gaze, facial expressions or voice intonation, and to communicate with others.

One of the most relevant auditory stimuli for our species is voice, which constitutes a very strong social input and tends to be processed as a priority (Whitten et al., 2020). Human beings are considered to be voice experts, because of our experience with and ability to decode such auditory stimuli (Latinus & Belin, 2011). The voice is acoustically richer than other types of auditory stimuli in terms of harmonics, pitch and intensities, with more details to encode. Voice stimuli may therefore require further processing prior to full adaptation. Belin et al. (2004) proposed a neurocognitive model of vocal perception. In brief, part of this model suggests that all auditory stimuli are processed, in a general low-level auditory analysis, in the primary cortex A1, and that vocal stimuli then enter a specific, voice structural analysis, involving other regions close to primary auditory cortex, such as the superior temporal gyrus. This suggests that vocal stimuli require additional processing, involving additional regions, which may be reflected in different dynamics for establishing a regularity for vocal versus non-vocal stimuli.

The neural process underlying habituation is called ‘Repetition Suppression’ (RS). RS refers to neural adaptation—a decrease in neural activity during repeated exposure to the same stimuli. It therefore reflects the formation and continuous updating of sensory memory traces (James et al., 2000; Schacter et al., 2004; Haenschel et al., 2005) and is translated at the cerebral level by a decrease in neural response with an increasing number of repetitions of the same stimulus (neural fatigue or sharpen activity models) (Desimone et al., 1996; Grill-Spector et al., 2006; Segaert et al., 2013). RS was initially studied in animals at the level of the individual neuron and was termed ‘Stimulus Specific Adaptation’ (SSA; see Escera & Malmierca, 2014). SSA refers to the decrease in the response of a single neuron with increasing repetitions of the same stimulus (standard), with no decrease in response to a rare stimulus (deviant) (Ulanovsky, Las, & Nelken, 2003), and has been recently dissociated from prediction error (Parras et al., 2017). SSA has been observed at cortical (Miller, Li, & Desimone, 1991; Ulanovsky et al., 2004) and sub-cortical levels (Ayala & Malmierca, 2013; Pérez-González, Hernandez, Covey, &

Malmierca, 2012), both in visual (Kaliukhovich & Vogels, 2012, 2014; Müller, Metha, Krauskopf, & Lennie, 1999) and auditory regions (von der Behrens, Bäuerle, Kössl, & Gaese, 2009; Nelken, Yaron, Polterovich, & Hershenhoren, 2013). A complementary process of RS, the Repetition Enhancement (RE), reflects the recognition of a stimulus, the anticipation of its appearance, and results in an increase in neural response with repetition (Segaert et al., 2013; Vogels, 2016). These processes can also be explained by the predictive coding model (Friston et al., 2005; Bendixen et al., 2012; Auksztulewicz & Friston, 2016) which proposes that the first presentation of a stimulus triggers the generation of a prediction of future sensory input. When this stimulus is repeated, it is compared to the prediction (Todorovic et al., 2011). According to Bayesian models of perceptual inference, RS would reflect a decrease in computational demand that occurs as the prediction error decreases, due to the match between sensory inputs and expected information (Summerfield et al., 2008). In parallel an Expectation Suppression (ES) effect has been identified, which would correspond to a diminution of the neural activity while the expectation become strongest. ES would co-exist with RS (de Gardelle et al., 2013; Grotheer & Kovacs, 2015) but the two phenomena would be independent as the ES effect occurs slightly later than the RS, as shown in both the visual (Summerfield et al., 2011) and the auditory modalities (Todorovic & Lange, 2012). Finally the RE which indexes an increase in the prediction strength when expected, would also be observed with a short delay, in separate frontal brain areas (Recasens et al., 2015).

The establishment of stimulus regularity has previously been studied in both visual and auditory modalities, in electrophysiology (Haenschel et al., 2005; Costa-Faidella et al., 2011; Recasens et al., 2015; Ferrari et al., 2017) and brain imaging (Fiebach et al., 2005; Gagnepain et al., 2008; Andics et al., 2013; Cacciaglia et al., 2019). It was first explored indirectly by studying the electrophysiological response to the detection of change, i.e., the ‘Mismatch negativity’ (MMN), a negative component, obtained by subtracting the response following the repetition of a standard stimulus from the response evoked by a changing deviant stimulus (Näätänen et al., 1978). MMN indexes deviancy detection, occurring when a stimulus is incongruent with the memory representation of the preceding repeated stimuli (Winkler et al., 2001; Näätänen et al., 2007). In the framework of the predictive coding model, MMN is considered as a marker of error detection, caused by a deviation from a learned regularity (Garrido et al., 2009).

Previous electrophysiological studies in humans, directly investigated auditory regularity encoding through the use of a ‘roving paradigm’, in which a stimulus is repeated a number of times (n), then followed by a new stimulus, which is also repeated n times. This leads to a continuous updating of the memory trace, which is suppressed at the end of each stimulus train (Cowan et al., 1993; Baldeweg et al., 1999, 2004, Haenschel et al., 2005; Costa-Faidella et al., 2011). These studies have highlighted RS phenomena, through the modulation of electrophysiological indices. The repeated presentation of a stimulus results in a decrease of the N1 component, between 90 and 150 ms and an increase in the positive components P1 and P2, reflecting the adaptation of responses. Through comparing the responses to a new stimulus and to

the same stimulus, after a number of repetitions, it is possible to isolate an electrophysiological index of the neural adaptation. The Repetition Positivity (RP), corresponds to a positive deflection between 50 and 250 ms, increasing with the number of repetitions (Baldeweg et al., 2004; Haenschel et al., 2005; Costa-Faidella et al., 2011).

The present study aims to explore the influence of the information to be encoded on the establishment of an auditory regularity. Although adaptation to vocal sound regularities is essential to extract relevant information for social communication, the formation of auditory regularity in the context of vocal stimuli has received little research interest. To our knowledge, most previous studies focused upon pure tones repetitions. Only one study explored the establishment of regularity to more complex semi synthesized sounds of two categories: vowels sounds representing Finnish vowels, and their vowel-like equivalents with increased formants frequencies, that were unfamiliar sounds (Ylinen & Huotilainen, 2007). Using these complex non-natural sounds in a roving paradigm, this study did not show any RP, possibly because of the very low number of participants and of repetitions used (3–4) compared to a minimum of 12 repetitions in other paradigms (Costa-Faidella et al., 2011).

Another study compared RP to vocalizations with different emotional valences, and showed that repetition suppression was increased for positive vocal stimuli as compared to neutral or negative vocal sounds (Pineiro et al., 2017). Authors concluded that positive vocalizations lead to enhanced sensory prediction. However although vocal stimuli might be considered as potent predictors, they also contain rich acoustic information. Whether we become accustomed to repeated vocal information in the same way as to non-vocal auditory stimuli thus remains to be explored.

To study brain correlates of auditory regularity in the context of vocal stimuli, trains of different lengths (4, 8, 16 repetitions) comprising either vocal or non-vocal complex sounds are compared. The different train lengths should allow us to determine the minimum number of repetitions required to elicit RP with more social stimuli. We assume that vocal regularity encoding is related to the phenomenon of RS, and anticipate an increase in the positivity (RP) with increasing number of vocal sound repetitions. Analysis of the RP characteristics (amplitude, latency and organization of the response through brain topography analyses) to vocal and non-vocal auditory stimuli will provide information on the RS effect in each condition, that will thereby be compared independently of the sensory response specific to each stimulus category. The comparison of P1 amplitude modulation with increasing number of repetitions in each condition will allow to estimate the dynamic of the neural adaptation, by providing information about the amount of repeated information necessary to attain a plateau in the response (to reach a full neural adaptation). We hypothesize that neural adaptation is likely to be influenced by the vocal/non-vocal aspect of the stimuli, with more repetitions needed for the P1 amplitude to reach a plateau for vocal stimuli. The involvement of memory traces could be weaker and/or slower for vocal stimuli, since these contain richer acoustic complexity than non-vocal stimuli.

## 2. Materials and method

We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

Previous EEG studies rarely report all the information needed to calculate sample sizes (Larson & Carbine, 2017). The sample size for this exploratory EEG study was thus determined based on previous research in the field. We chose a conventional sample size in ERP research (20 participants), and reported all the information needed for sample size estimation in future studies.

### 2.1. Population

Twenty young adults aged 18 to 30 (mean =  $24 \pm 2$  years) participated in the study (12 female). None had neurological, psychiatric, or metabolic disorders or were under medication at the time of the study. None had a hearing deficit as tested with an audiometer at different frequencies (250, 750 and 1500 Hz). Each participant signed an informed consent form, and the protocol received approval from Ethic Committee (PROSCEA2017/23; ID RCB: 2017-A00756-47).

### 2.2. Stimuli

A total of 8 vocal sounds and 8 non-vocal complex sounds were used to study the possible impact of vocal/non-vocal nature of the information to be encoded. The vocal sounds consisted of the vowel “a” uttered with a neutral prosody by 8 female speakers of different identities, selected from an existing database of vocal sounds validated (on the basis of valence and emotion recognition) on an independent sample of adults ( $n = 16$ ) (Charpentier et al., 2018). Non-vocal stimuli were synthetic complex sounds with acoustic characteristics close to those of the vocal sounds. For this we created complex sounds with a global frequency spectrum similar to that of the voices. Firstly, using speech analysis software (Praat®; see Boersma, 2002), we measured the values of the fundamental frequency (F0) and the first 4 formants of each of the eight voices selected as vocal stimuli. The second step consisted of synthesizing complex sounds using sinewaves of the corresponding frequency values (Adobe Audition® software). Amplitudes of each harmonic were adapted to best fit the voices frequency spectrums. To achieve the necessary attenuations (spectral slope), each time the frequency doubled we applied a decrease of 12 dB to the amplitude of the harmonic (Kreiman & Gerratt, 2012). In both categories, each sound was normalized according to the root mean square of the amplitude, so that all stimuli had the same energy, using Matlab (The MathWorks Inc., Natick, MA, USA); fade-in and fade-out effects of 10% (30 ms) were applied. Global energy was equalized at 65 dB SPL. We obtained 8 vocal sounds with F0 ranging from 190 Hz to 230 Hz, with frequency steps of about 5 Hz, and 8 non-vocal sounds with the same F0s and spectral structure (Table 1; Fig. 1). Sounds are controlled for their ‘speech value’ as the synthesized sounds mimic their natural vocal equivalents in terms of main frequencies; the two categories of

**Table 1 – Frequency composition of each non-vocal sounds based on the identification of F0, F1, F2, F3 and F4 of the corresponding vocal sounds.**

Sound	Fundamental frequency F0	Formant 1	Formant 2	Formant 3	Formant 4
1	190	905	1515	2830	3716
2	199	845	1358	2732	4228
3	205	975	1326	2698	3477
4	210	705	1480	2847	4479
5	215	877	1522	2971	4094
6	220	806	1502	2322	3021
7	226	797	1483	2943	4096
8	229	886	1465	3005	3640

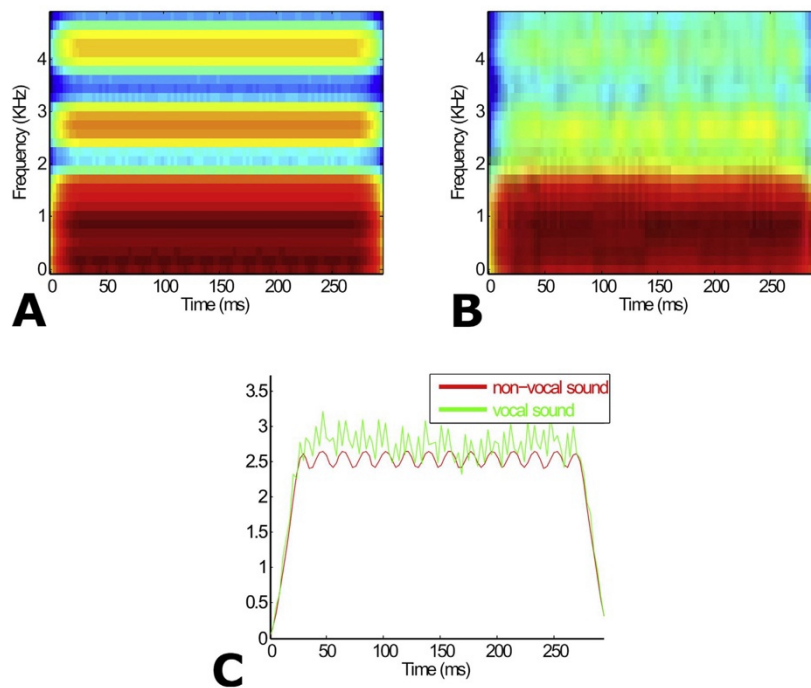
sounds therefore contain the F0, F1 and F2 allowing the phoneme/a/automatic categorization as previously demonstrated (Jacobsen et al., 2004).

### 2.3. Sequences of stimulation

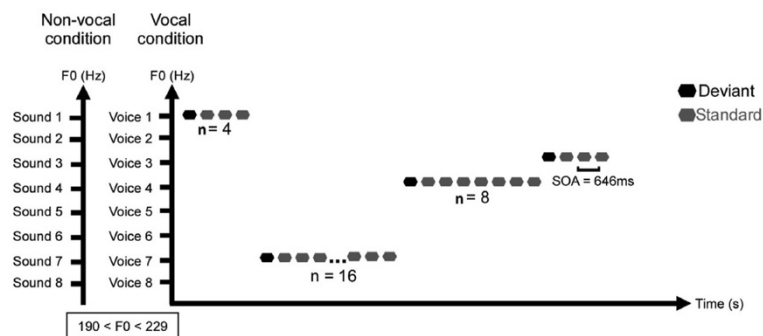
We used a roving paradigm (Baldeweg et al., 2004; Cowan et al., 1993) in which trains of stimulation of different frequencies/voices and different lengths were presented. Each train was composed of the same stimulus, which was repeated  $n$  times (depending on the length of the train), and called standard (S). The first stimulus of each train was

considered as deviant compared to the preceding repeated standard. We used trains of 4, 8 and 16 repetitions. These trains were delivered in a pseudo-random order, with the only constraint that the same stimulus could not be delivered in two consecutive trains.

Two sequences were presented, the first consisting of vocal stimuli and the second consisting of the non-vocal stimuli described above. The order of presentation of the sequences was counterbalanced between subjects. Each stimulus was presented for 300 ms, with a SOA of 646 ms, via two speakers located at 1.2 m from the ears of the participants (Logitech Z-2300). 120 repetitions of each train length were presented for



**Fig. 1 – Physical features of a vocal sounds and of its non-vocal counterpart. (A) spectrogram of the non-vocal sound, (B) spectrogram of the vocal sound, (C) energy of the vocal and the non-vocal sound as a function of time (green: vocal sound, red: non-vocal).**



**Fig. 2 – Roving paradigm used in this study. Trains of four, eight or sixteen repetitions of the same stimulus were pseudo-randomly delivered. The SOA was constant at 646 ms. Vocal condition: run composed of vocal stimuli, the vowel “a”, uttered by 8 different female speakers with progressive increasing of fundamental frequencies ( $190 \leq F_0 \leq 229$ ). Non-vocal condition comprised eight synthetic complex sounds with acoustic characteristics close to those of the vocal sounds, with the same  $F_0$  progressive increase ( $190 \leq F_0 \leq 229$ ).**

each sequence, resulting in a total of 6720 stimuli (3360 per sequence) and a recording time of approximately 1 h. A schematic representation of the roving paradigm used in this study is illustrated in Fig. 2.

#### 2.4. EEG data acquisition and processing

During EEG recording, participants were sitting comfortably in a reclining armchair, located in a sound-attenuated room. Subjects watched a silent movie without subtitles whilst sounds were delivered; they were instructed that they would have to briefly tell the story of the movie at the end of the recording session. This procedure avoided voluntary directing of attention towards the auditory stimulation. The two stimulation sequences were delivered, with a break between, so that the subject could move and relax.

Presentation software (NeuroBehavioral Systems Inc., Berkeley, CA) was used to deliver the stimulation sequences, during which the EEG was recorded from 64 active electrodes (Active-Two Systems Biosemi, The Netherlands) with a sampling rate of 512 Hz. Horizontal and vertical eye movements were monitored using electrodes placed on the left and right outer canthi and below the left eye. An electrode was placed on subject's nose for offline re-referencing. The ELAN software package was used for the analysis of EEG-ERP (Aguera, et al., 2011). The EEG signal was amplified and filtered with a .1 Hz high-pass filter (Butterworth filter, order 1). Artefacts resulting from eye movements were removed by applying Independent Component Analysis (ICA) as implemented in EEG Lab. Blink artifacts were captured into components and selectively removed via inverse ICA transformation. Sixty-four components were examined, and one or two components were removed in each subject, to account for vertical and horizontal movements. Motion artefacts, characterized by high frequency or high amplitude signals, were discarded manually by an experimenter blind from trial type. A 30 Hz low-pass filter was applied (Butterworth filter, order 3), and ERPs were averaged over a 700 ms time window, including a 100 ms pre-stimulus baseline.

The recordings of the two stimulation sequences were carried out separately for each subject, and vocal and non-vocal stimulations were averaged separately.

#### 2.5. Data analysis

##### 2.5.1. Measurements

Responses to each stimulus, in positions 1 to 16, were averaged individually for each sequence, so that the effect of repetition for each position could be observed. Trials presented at the same position, but in trains of different lengths (e.g., stimulus 3 in a train of 4, 8 and 16 repetitions) were also averaged together. This resulted in three times more stimuli from 1 to 4, and twice as many stimuli from 5 to 8, than stimuli from positions 9 to 16. Responses to stimuli in positions 2 to 4 were averaged, as well as those in positions 5 to 8, those in positions 9 to 12 and finally those in positions 13 to 16, in order to obtain more robust responses by increasing the number of trials. For each group of stimuli, the mean number of artifact-free trials was:  $925 \pm 72$  (S2-4),  $820 \pm 63$  (S5-8),  $404 \pm 45$  (S9-12) and  $410 \pm 31$  (S13-16) for the non-vocal sequence and  $912 \pm 81$  (S2-4),  $807 \pm 77$  (S5-8),  $401 \pm 44$  (S9-12) and  $400 \pm 41$  (S13-16) for the vocal sequence.

For the 4 groups of repeated standards, only the P1 component was clearly identified and measured for both conditions. To isolate the RP response, the difference between the average stimuli 13 to 16 (long train) and the averaged stimuli 2 to 4 was calculated for both conditions.

In addition, to isolate MMN responses after each length of train, the differences between the first stimulus (deviant) and last stimulus (standard) of each length of train (4, 8 and 16) were calculated for both conditions. For these stimuli, the mean number of artifact-free trials were:  $307 \pm 27$  (S1),  $104 \pm 9$  (S4),  $103 \pm 10$  (S8) and  $102 \pm 9$  (S16) for the non-vocal sequence and  $302 \pm 29$  (S1),  $100 \pm 9$  (S4),  $102 \pm 11$  (S8) and  $100 \pm 10$  (S16) for the vocal sequence.

The P1 (positive peak at ~ 100 ms) was identified as the first positive deflection occurring between 70 and 120 ms. This

time-window was selected on the basis of studies with similar stimuli (Charpentier et al., 2018; Sheerer et al., 2013). Amplitudes and latencies of the P1 were measured in each participant in a 70–120 ms time window, centered on the peak of the grand mean average of the group. RP and MMN measure time windows were selected on the basis of previous studies with similar stimuli (Bishop et al., 2011; Charpentier et al., 2018; Pinheiro et al., 2017). RP was identified as a positive deflection occurring between 90 and 200 ms and the MMN as a negative deflection occurring in a 130–190 ms time-window. Amplitudes and latencies of the RP and the MMN were measured in a 90–200 ms and a 130–190 ms time window, respectively, centered on the peak of the grand mean average of the group.

### 2.5.2. Statistical analyses

After visual inspection of the responses scalp distributions, and based on previous studies (Haenschel et al., 2005; Costa-Faidella et al., 2011), amplitudes and latencies were analyzed at Fz.

Two-way ANOVAs were performed on P1 amplitudes and latencies, with the Repetition (2–4, 5–8, 9–12 and 13–16) and Condition (vocal vs non-vocal) as within subject factors.

Two-way ANOVAs were performed on MMN amplitudes and latencies with the Repetition (4, 8 and 16) and Condition (vocal vs non-vocal) as within subject factors.

Since the factor Repetition displayed more than 2 levels, a Greenhouse-Geisser correction was applied to correct for potential violations of the sphericity assumption. Additional Bonferroni post-hoc tests were performed to examine the direction of the interactions. The effects sizes are shown as  $\eta^2_p$ .

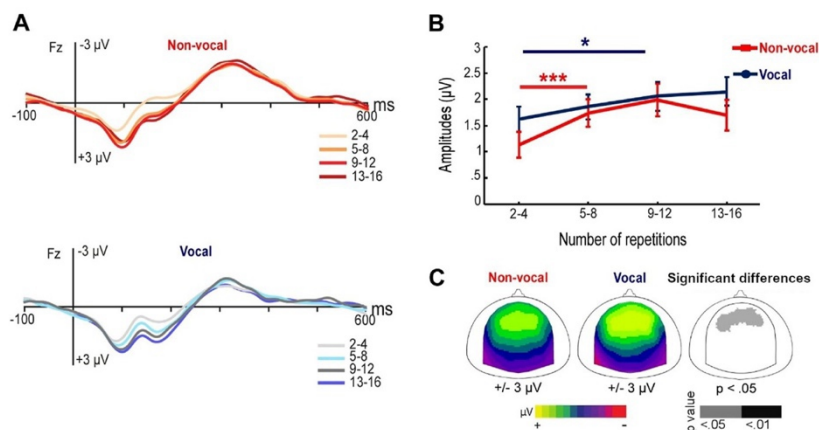
We performed permutation tests based on randomizations (Edgington, 1995) over a 50–300 ms time-window at each electrode and at each time point, to assess difference in the RP between conditions (vocal vs. non-vocal). These analyses provide supplementary information on condition differences,

by confirming peak analyses or by affording additional topographical findings or amplitude statistical comparisons for responses whose peak is barely measurable. Each permutation test involved the random permutation of the values for the 20 pairs of data compared (corresponding to the 20 participants), then the calculation of the sum of squared sums of values in each of the two obtained samples, and finally the computation of the difference between these two statistical values. For each analysis we performed 10,000 such randomizations, to obtain an estimate of the distribution of this difference under the null hypothesis. This distribution was then compared to the actual difference between the values in the two conditions (vocal vs. non-vocal). Correction for multiple comparisons was performed using the statistical-graphical method of Guthrie and Buchwald (1991) which tabulates the minimum number of consecutive time samples that need to be significant in the ERP differences, in order to have a significant effect over a given time window. For the analyses of RP (250 ms: 50–300, i.e., 125 sampling points), the minimum number corresponded to 12 consecutive time points (i.e., 24 ms) with p values below the .05 significance level.

## 3. Results

The grand mean ERPs to standard after 2–4, 5–8, 9–12 and 13–16 repetitions, for vocal and non-vocal conditions are illustrated in Fig. 3A. The P1 component (positive peak at ~ 100 ms) was clearly observed, and modulated by repetition, for both conditions. The mean peak amplitudes and latencies of the P1 for each condition are reported in Table 2.

Auditory evoked potentials were acquired over the entire scalp, revealing a large fronto-central positivity. Analyses were performed at the Fz electrode, where the responses culminated.



**Fig. 3 – (A) Standard ERPs at Fz electrode after 2–4, 5–8, 9–12 and 13–16 repetitions in vocal (top) and non-vocal (below) conditions. (B) P1 amplitude values plotted for 2–4, 5–8, 9–12 and 13–16 repetitions in vocal (blue) and non-vocal (red) conditions (error bars represent SEM); \* $p < .05$ , \*\* $p < .01$ . (C) P1 scalp potential distribution at 100 ms after 13–16 for non-vocal and vocal conditions and Statistical map of the topographical differences calculated using permutation analyses.**

**Table 2 – Summary of P1 mean amplitudes and latencies according to the number of repetitions (m±SEM).**

Number of repetitions	Amplitude ( $\mu\text{V} \pm \text{SEM}$ ) at Fz		Latency (ms $\pm$ SEM) at Fz	
	Non-vocal	Vocal	Non-vocal	Vocal
2–4	1.13 $\pm$ .25	1.62 $\pm$ .24	91.50 $\pm$ 2.52	90.15 $\pm$ 2.31
5–8	1.74 $\pm$ .26	1.86 $\pm$ .24	95.34 $\pm$ 3.25	96.62 $\pm$ 3.52
9–12	1.98 $\pm$ .31	2.05 $\pm$ .28	96.15 $\pm$ 3.12	99.73 $\pm$ 3.03
13–16	1.70 $\pm$ .29	2.15 $\pm$ .27	101.31 $\pm$ 3.21	101.85 $\pm$ 2.33

### 3.1. Difference in repetition effects between conditions

#### 3.1.1. P1 component

A significant repetition effect on P1 amplitudes was found ( $F(3, 57) = 17.635, p < .001, \mathcal{E} = .875, \eta^2 p = .48$ ) due to an increase in P1 amplitude with increased repetition. In addition, a significant condition effect was observed ( $F(1, 19) = 6.287, p = .021, \mathcal{E} = 1.000, \eta^2 p = .25$ ) due to larger amplitudes for the vocal versus non-vocal condition. An interaction between the condition and the number of repetitions was highlighted ( $F(3, 57) = 3.514, p = .021, \mathcal{E} = .848, \eta^2 p = .16$ ).

Post-hoc analyses (Bonferroni) revealed different repetition effects between the two conditions, with a significant increase in P1 amplitude between 2-4 and the three other groups of stimuli (2–4 vs 5–8, 9–12 and 13–16 respectively:  $p < .001, p < .001, p < .001$ ) but not between 5-8 and 9–12, nor between 5-8 and 13–16 repetitions and between 9-12 and 13–16 repetitions ( $p = .45$ ) for the non-vocal condition. For the vocal condition, a significant increase in P1 amplitude was revealed between 2-4 and 9–12 repetitions ( $p = .011$ ) and between 2-4 and 13–16 repetitions ( $p < .001$ ) but not between 2-4 and 5–8 repetitions, 5–8 and 9–12 nor between 9-12 and 13–16. Statistical difference in the P1 amplitude was found for vocal versus non-vocal conditions, for 2–4 ( $p = .003$ ) and 13–16 ( $p = .008$ ) repetitions, with larger amplitudes observed in the vocal condition.

A significant repetition effect was found for P1 latencies ( $F(3, 57) = 7.18, p = .001, \mathcal{E} = .743, \eta^2 p = .27$ ) due to an increase in P1 latencies as repetitions increased. There was no condition effect ( $F(1, 19) = .18, p = .680, \mathcal{E} = 1.000, \eta^2 p = .01$ ) and no interaction between the condition and the number of repetitions ( $F(3, 57) = .69, p = .526, \mathcal{E} = .776, \eta^2 p = .04$ ).

Statistical analyses performed on a larger set of electrodes surrounding Fz (FCz, F1, F2, FC1, FC2), revealed similar results for both P1 amplitude and P1 latency.

The scalp distribution of the P1 remained stable across the two conditions, corresponding to a large fronto-central positivity, increasing with repetition.

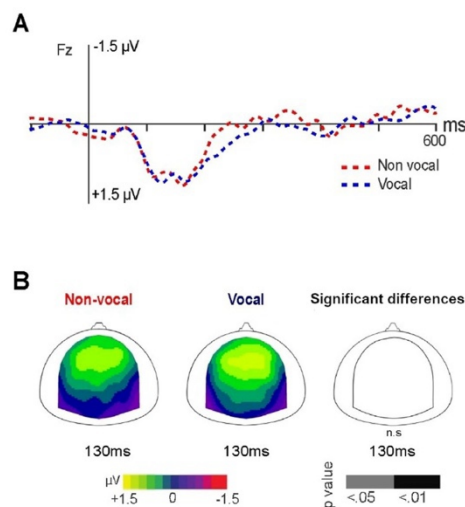
An example of the P1 scalp potential distribution at 100 ms after 13–16 for non-vocal and vocal conditions with the statistical maps of the topographical differences is illustrated in Fig. 3C.

#### 3.1.2. Repetition positivity

To better understand the activity associated with stimulus repetition we isolated the RP by subtracting the average responses to stimuli 2–4 (short trains) from the average response to stimuli 13–16 (long trains). Positive deflections corresponding to the RP were observed between 90 and 200 ms over fronto-central electrodes in both vocal and non-vocal

conditions. The RP grand-average difference waveforms (13–16 minus 2–4 repetitions) for vocal and non-vocal conditions are illustrated in Fig. 4A.

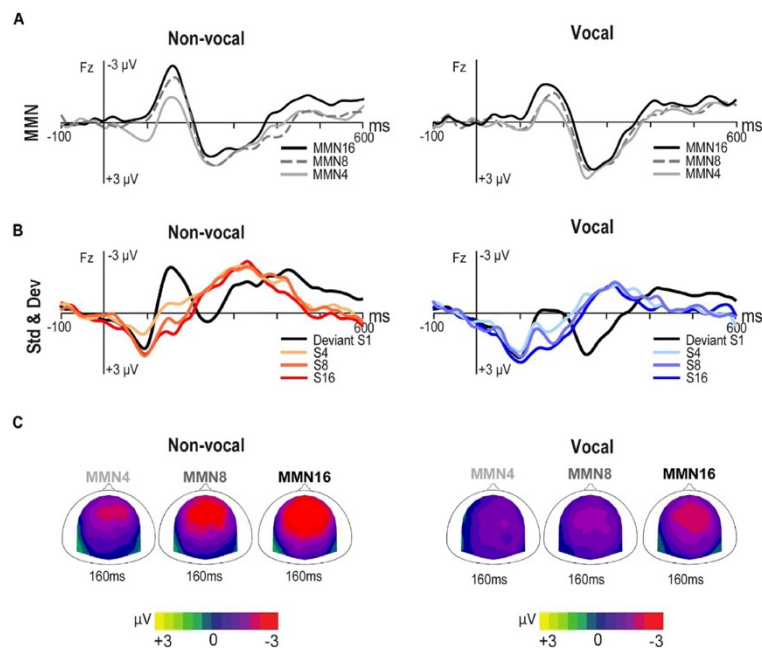
Permutation analyses performed between the two conditions, revealed no difference between RP for vocal and non-vocal sounds, in the reported time-window. The scalp distribution of RP, characterized by a large fronto-central positivity, remained stable across the two conditions. Scalp distributions of RP, for vocal and non-vocal conditions, based on the results of the permutation analyses, are shown in Fig. 4B. Comparison of the responses obtained on mastoid sites, where the negative activity associated with the fronto-central RP is observed, was also performed. Results revealed no significant differences in the amplitude of the mastoid negativity, between vocal and non-vocal conditions. This reverse polarity between the Fz electrode and the mastoids, confirms the location of RP generators in the supra temporal plane, at the level of the auditory regions.



**Fig. 4 – (A) Repetition Positivity (RP) resulting from the difference wave between 13-16 and 2-4 repetitions for vocal (blue) and non-vocal (red) conditions. (B) RP scalp potential distribution at 130 ms for vocal (right) and non-vocal (left) conditions and Statistical map of the topographical differences calculated using permutation analyses.**

**Table 3 – Summary of MMN mean amplitudes and latencies according to the number of repetitions ( $m \pm SEM$ ).**

Number of repetitions	Amplitude ( $\mu V \pm SEM$ ) at Fz		Latency (ms $\pm SEM$ ) at Fz	
	Non-vocal	Vocal	Non-vocal	Vocal
4	-1.60 $\pm$ .47	-1.01 $\pm$ .30	155.56 $\pm$ 2.93	168.75 $\pm$ 3.48
8	-2.41 $\pm$ .41	-1.45 $\pm$ .44	157.81 $\pm$ 3.13	174.71 $\pm$ 3.17
16	-2.91 $\pm$ .42	-1.81 $\pm$ .34	156.84 $\pm$ 2.47	170.01 $\pm$ 3.47



**Fig. 5 – (A) Mismatch Negativity (MMN) difference waves after 4 (grey), 8 (dotted line) and 16 (black) repetitions for non-vocal (left) and vocal (right) conditions. (B) Standard after 4, 8 and 16 repetitions and deviant ERPs at Fz electrode for non-vocal (left) and vocal (right) conditions. (C) MMN scalp potential maps at 160 ms for non-vocal (top) and vocal (below) conditions.**

### 3.2. Effect of repetition on deviance detection

The repetition effect on deviance detection was analyzed after 4, 8 and 16 repetitions. The mean peak amplitudes and latencies of the MMN for each condition are reported in Table 3.

The response to the repeated stimuli in positions 4, 8 and 16 was subtracted from the responses to the deviant stimulus (first stimulus of a train). The MMN grand-average difference waveforms after 4 (S1 minus S4), 8 (S1 minus S8) and 16 (S1 minus S16) repetitions are illustrated in Fig. 5A. The standard ERPs for 4, 8 and 16 repetitions, and the deviant ERPs after 4, 8 and 16 repetitions of the preceding standard, at Fz electrode, for non-vocal and vocal conditions are shown in Fig. 5B.

A significant repetition effect was found on MMN amplitudes ( $F(2, 38) = 12.251, p < .001, \mathcal{E} = .922, \eta^2 p = .39$ ) due to an increase in MMN amplitude with increasing repetitions. No effect of condition ( $F(1, 19) = 1.901, p = .184, \mathcal{E} = 1.000,$

$\eta^2 p = .09$ ) and no interaction between condition and repetition ( $F(2, 38) = .530, p = .556, \mathcal{E} = .845, \eta^2 p = .03$ ) were found.

A significant condition effect was found for MMN latencies ( $F(1, 19) = 8.56, p = .009, \mathcal{E} = 1.000, \eta^2 p = .31$ ) due to an earlier MMN for non-vocal versus vocal stimuli. No effect of repetition ( $F(2, 38) = .208, p = .807, \mathcal{E} = .973, \eta^2 p = .01$ ) and no interaction between condition and repetition ( $F(2, 38) = .627, p = .537, \mathcal{E} = .984, \eta^2 p = .03$ ) were observed for MMN latencies.

The scalp distribution of the MMN, corresponding to a large fronto-central negativity, remained stable between the two conditions. Topographical distributions of MMN observed after 4, 8 and 16 repetitions are presented in Fig. 5C.

## 4. Discussion

This study aimed to explore the possible influence of the information to be encoded (vocal versus non-vocal) on the

establishment of auditory regularity indexed by the repetition positivity.

An increase in the number of repetitions generated a significant increase in positivity at the fronto-central electrodes, between 90 and 200 ms, for both vocal and non-vocal stimuli. Results are consistent with previous studies using pure tones, highlighting an RP between 50 and 250 ms (Baldeweg et al., 2004, Haenschel et al., 2005; Costa-Faidella et al., 2011; Recasens et al., 2015). It should be noted that the latencies observed are later than expected from the literature, a delay that might be due to the nature of the stimuli used. Indeed responses to vocal stimuli have later components latencies than pure tones, possibly explained by a later encoding due to their complexity. Both the scalp topographies, characterized by a large positivity spreading over fronto-central regions, and the shape of the RP curves obtained, were very similar in the two conditions. A permutations analysis revealed no significant differences between conditions. This might be explained by the very close spectral composition between the two types of stimuli used, since the non-vocal stimuli were based on the vocal stimuli. Alternatively, the brain activity associated with RS may display the same organization, regardless of the stimuli used. This perspective is reinforced by the fact that the RP obtained in the present work is also similar to the RP observed in response to pure tones (Haenschel et al., 2005; Costa-Faidella et al., 2011). However, the only other study which attempted to explore regularity encoding by using a roving paradigm and comparing more complex stimuli like vocal and non-vocal sounds (Ylinen & Huotilainen, 2007), did not reveal any RP in response to stimuli repetition. Contrary to our initial assumptions, Ylinen and Huotilainen's negative results would not be related to the low number of repetitions used (2–3 versus 5–6 repetitions) since we showed that 5–8 repetitions are sufficient to observe an effect. The contradictory results could instead, be explained by the use of divers vowels instead of a single vowel—as in our study. Previous studies have shown differences in brain responses between vowels, due to the differences in their acoustic compositions (Obleser et al., 2003; Shestakova et al., 2004), which could have led to inconsistent results.

The observation of a polarity reversal at mastoid electrodes, confirmed the involvement of temporal auditory areas in the RS phenomenon, in response to vocal and non-vocal complex sounds repetition. This is similar to previous observations in response to pure tones (Cooper et al., 2013). Although our methodological approach did not aim to accurately locate the regions generating the observed responses, findings are consistent with Recasens et al. (2015) who used MEG to identify the regions involved in RP to pure tones repetition. Their study allowed to target the regions involved in RS and RE separately, and highlighted the contribution of both temporal regions and anterior frontal insula in these two phenomena.

A few studies have highlighted that RP results from the joint modulation of the P1, N1 and P2 components (increase of P1 and P2 with a decrease of N1) (Baldeweg et al., 1999; Haenschel et al., 2005; Costa-Faidella et al., 2011). Due to the specific pattern of brain potentials elicited by our stimuli, results did not allow for the isolation of these three individual components, for the vocal and non-vocal conditions. However, the latency and morphology of the RP obtained, confirms the joint modulation of components in the same latency range.

Although our results revealed similar RPs for vocal and non-vocal conditions, different dynamics were observed in terms of stimulus adaptation, with different profiles of variation in responses amplitude as repetitions increase for each condition. Statistical analysis of the P1 component revealed an increase in amplitude with repetition, just as previous studies have shown modulation of the canonical ERPs in response to pure tones (Haenschel et al., 2005; Costa-Faidella et al., 2011) and to more complex auditory stimuli (Jacobsen & Schröger, 2001). Moreover, we showed that the effect of repetition no longer changes after 5–8 repetitions for the non-vocal condition, but becomes stable only after 9–12 repetitions for the vocal condition. Based on these observations, the establishment of regularity appears to be faster for non-vocal stimuli, with only 5–8 repetitions required, than for vocal stimuli, where regularity is only fully established after 9 to 12 repetitions. This is consistent with our initial assumption, that adaptation to vocal stimuli would be slower. Moreover it seems that the complexity of voice, which is acoustically richer, requires further processing steps, and subsequently leads to slower neural adaptation. At the basic processing level, the spectral richness of the sounds could explain differences in neural adaptation. Although they have been controlled on several physical features, the sounds used in the present study do not contain exactly the same acoustic information. One mechanistic explanation for the repetition effect is stimulus-specific adaptation, i.e. neurons specialized in processing specific physical features (e.g. frequency) have higher activity when processing identical features of the stimulus and are therefore subject to the so-called “refractory process” (Näätänen et al., 2005). The acoustic richness of vocal sounds in terms of spectral frequencies compared to non-vocal sounds with more focused frequency components would explain the observed difference in electrophysiological responses, as the complexity of the physical features to be encoded could recruit a larger number of neurons and ultimately lead to reduced dimension-specific neural adaptation (or given the observed data, to a neural adaptation that would be slower to take place). This is in line with previous NIRS studies on complex speech structure detection in infants, which found that stimulus complexity influences repetition effects (Gervain et al., 2008; Bouchon et al., 2015).

On the other hand, the human brain is efficient in making predictions and in automatically detecting slight changes in voice, due to its expertise (Latinus & Belin, 2011). And indeed voices are processed faster and more automatically compared to synthetic sounds (Whitten et al., 2020). In the same line, studies have shown that the familiarity of a stimulus has an impact on automatic auditory processing (Jacobsen et al., 2005). Social cues are more familiar to us because they contain more relevant information than simpler stimuli such as the non-vocal sounds used in this study. According to this voice preference, one could have expected a faster or larger neural adaptation for vocal sounds than for non-vocal synthesized unfamiliar sounds. Analyses of ERPs and time-frequency features to repeated vocal emotions (Pinheiro et al., 2017) showed that both induced pre stimulus beta power and RP amplitude were increased for happy vocalizations, suggesting that positive vocalizations lead to strong sensory prediction. Moreover, as described earlier, the repetition positivity includes a predictive coding component (Cacciaglia, et al., 2019) and vocal

sounds, which involve larger neural networks (those encoding for speech, language, social context) than non-vocal sounds, in turn will elicit stronger predictions. These may have affected the RP more strongly as suggested by [Costa-Faidella et al. \(2011\)](#) who showed larger RP in the predictable than in the unpredictable condition. Yet in the present study, although comparison of vocal and non-vocal repetitions showed that neural adaptation takes longer for vocal stimuli, RPs were similar in amplitude.

From the predictive coding perspective, the RP and the P1 modulation we recorded to repeated stimuli, probably reflect the combined effects of both RS, ES and RE. The classical roving paradigm and the ERPs method used do not allow to spatially and temporally distinguish between these different phenomena. However, as RS is thought to be low-level and automatic ([Kouider et al., 2009](#)) and ES more sensitive to attention and stimulus familiarity ([Grotheer & Kovacs, 2014](#)), one can assume that even if vocal stimuli represent potent predictors compared to non-vocal stimulation (which should increase ES), their acoustic complexity would slow down the RS effect, resulting in a decreased global response. Alternatively, the impact of the ES effect might also depend on the social relevance of the vocalization. Based on the present results, we were able to draw conclusions on these different hypotheses. Similar RP were recorded for both non-vocal and vocal sounds, but stimulus adaptation was different, with the latter being faster for non-vocal sounds than for vocal sounds. These observations may indicate that regarding neural adaptation, the acoustic simplicity of a sound takes precedence over its familiarity and/or social value.

According to the present results, a maximum of 9–12 repetitions is sufficient to establish auditory regularity, regardless of the type of sound. This is consistent with the results obtained by [Costa-Faidella et al. \(2011\)](#), who highlighted a maximum effect at 12 repetitions, with pure sounds. It would therefore not be necessary to use trains of stimulation with 36 repetitions, as done by [Haenschel et al. \(2005\)](#). This is confirmed by the RP response, which shows the same morphology found in previous studies ([Costa-Faidella et al., 2011](#); [Haenschel et al., 2005](#)), and also displays a similar amplitude (about 1.5  $\mu$ V), suggesting that the effect no longer evolves between 12 and 36 repetitions.

In addition to the RP analysis, which directly correlates with regularity encoding, we studied the effect of repetition on deviance detection by analyzing the MMN response to each train change. For that, MMN difference waveforms were measured after 4, 8 and 16 repetitions in each condition. Results showed an increase in MMN amplitude with repetition for both vocal and non-vocal complex sounds, as previously shown for tones ([Baldeweg et al., 2004](#); [Garrido et al., 2009](#)). These observations support the fact that deviance detection is impacted by previous regularity encoding. However, one study using pure tones in a roving paradigm, did not reveal an increase in MMN amplitude with an increasing number of preceding standard repetitions ([Cooper et al., 2013](#)). This could be due to the deviant used, which was the last stimulus of each train, varying in duration. Few studies have focused on the impact of long-term memory traces on short-term memory traces formation. A previous MMN study, presented native-vowel and non-native vowel deviants to highlight the

existence of language-dependent memory traces ([Näätänen et al., 1997](#)). Results showed that the MMN elicited by native deviant sounds was larger than the MMN elicited by non-native deviant sounds. Another study using a roving paradigm, studied the impact of long-term memory trace related to native language characteristics on the short-term memory trace formed by phonemes, according to the number of preceding stimulus repetitions ([Huotilainen et al., 2001](#)). Three different types of non-vocal stimuli were used: (1) *pure tones* (2) *prototype-vowels*: complex synthetic sounds based on the frequency composition of 8 vowels and (3) *non-prototype vowels*: complex synthetic sounds based on the same vowels, with their first two formants shifted upwards. Results showed that a reduced number of repetitions was required to produce the most prominent MMN for prototype-vowels than for non-prototype vowels, probably due to stronger long-term memory traces for the former. In the present study we used non-vocal, complex, synthetic sounds based on vocal vowel composition, but also natural-vocal sounds, in order to study the impact of the vocal component. We found no statistical differences in MMN amplitudes, between the conditions. This could be due to inter-individual variability, or to the close composition of the two groups of stimuli. Regarding latencies, [Huotilainen et al. \(2001\)](#) showed a longer MMN latency for non-prototype than for prototype vowels and pure tones, but only after few repetitions. In the present study, an earlier MMN was obtained for the non-vocal condition than the vocal condition, with no effect of repetition. Again, such results could be due to the specificity of vocal sounds, which are more complex to encode and process. In the present study, comparing vocal and non-vocal regularity encoding, findings do not strengthen the link between long-term and short-term memory traces, as in previous studies, but support early discrimination of simplest sounds deviancy. This result is consistent with the earlier establishment of regularity in response to non-vocal than to vocal sounds, and reinforces the link between the process of regularity encoding and deviance detection ([Baldeweg et al., 2004](#); [Garrido et al., 2009](#)).

## 5. Conclusion

The purpose of this study was to explore the possible impact of the nature of the information (vocal versus non-vocal) to be encoded, on the establishment of auditory regularity. Once established, auditory regularity leads to adaptation, a phenomenon known as repetition suppression. We highlighted a Repetition Positivity (RP) between 90 and 200 ms, following repeated trains of vocal and non-vocal complex stimulations, confirming the results of previous studies using simple auditory stimuli. As predicted, repetition suppression seems to be sensitive to the nature/complexity of the stimuli to be encoded. Indeed, there are different underlying dynamics for establishing regularity for vocal and non-vocal stimulations, with more repetition required for responses to vocal sounds to stabilize. Adaptation to the auditory environment would therefore follow a different pattern, depending on the type of stimuli, with voice requiring additional processing due to the more complex composition of vocal stimuli. In terms of social communication, this adaptation is essential to quickly detect

vocal variations and ultimately, to identify the emotional state of others. This high predictive role of vocal emotion is suggested by the study of Pinheiro et al. (2017) which highlighted the impact of positive emotions on neural adaptation. Moreover, one can assume that this adaptation to sensory environment may increase throughout development to become optimal. Deficits in such adaptation may lead to difficulties in everyday life, particularly in detecting relevant environmental changes or in developing adjusted social interactions. Carrying out similar studies in neuro-typical children may help to establish the optimal developmental pathway of repetition suppression, and related neural adaptation. In neuro-developmental disorders such as Autism Spectrum Disorders, where adaptation deficits have been recently reported at the subcortical level (Font-Alaminos et al., 2019), repetition suppression studies would allow us to determine whether difficulties in social interaction and adaptation to change, result from deficits in establishing a regular context.

### Ethics statement

The protocol received approval from Ethics Committee (PRO-SCEA2017/23; ID RCB: 2017-A00756-47). Each participant signed an informed consent form.

### Author contributions

Camille Heurteloup, Carles Escera, Sylvie Roux & Marie Gomot designed the study. Camille Heurteloup & Annabelle Merchie performed data acquisition. Camille Heurteloup, Annabelle Merchie & Marie Gomot were responsible for data and statistical analyses. All authors were involved in preparing and reviewing the manuscript.

### Availability of data and material

The conditions of our ethics approval do not permit public archiving of the data supporting this study. Readers seeking access to this data stimuli should contact the lead author Marie Gomot at the UMR 1253 Inserm, University of Tours. Access will be granted unconditionally to named individuals in accordance with ethical procedures governing the reuse of sensitive data.

Stimuli, procedures and scripts for main analyses are available in open access at <https://osf.io/afr4c/>

### Preregistration

No part of the study was pre-registered prior to the research being conducted.

### Funding

This work was supported by the INSERM (National Institute for Health and Medical Research) Inserm and the ANR (ANR-19-CE37-0022-01). C.H was supported by a grant of the

“Ministère de la recherche”. The funding sources had no involvement in the study design, data collection and analysis, or the writing of the report.

### Declaration of competing interest

All authors declare that they have no conflicts of interest.

### Open Practices

The study in this article earned an Open Materials badge for transparent practices. The conditions of our ethics approval do not permit public archiving of the data supporting this study. Readers seeking access to this data and experimental stimuli should contact the lead author Marie Gomot at the UMR 1253 Inserm, University of Tours. Access will be granted to named individuals in accordance with ethical procedures governing the reuse of sensitive data. There are no specific conditions for applying for data sharing.

### Acknowledgments

We thank all the volunteers for their time and participation in this study and Remy Magne for technical support.

### REFERENCES

- Aguera, P. E., Jerbi, K., Caclin, A., & Bertrand, O. (2011). ELAN: A software package for analysis and visualization of MEG, EEG, and LFP signals. *Computational Intelligence and Neuroscience*, 2011, 158970. <https://doi.org/10.1155/2011/158970>
- Andics, A., Gál, V., Vicsi, K., Rudas, G., & Vidnyánszky, Z. (2013). fMRI repetition suppression for voices is modulated by stimulus expectations. *Neuroimage*, 69, 277–283. <https://doi.org/10.1016/j.neuroimage.2012.12.033>
- Auksztulewicz, R., & Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 80, 125–140. <https://doi.org/10.1016/j.cortex.2015.11.024>
- Ayala, Y. A., & Malmierca, M. S. (2013). Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Frontiers in Neural Circuits*, 6. <https://doi.org/10.3389/fncir.2012.00089>
- Baldeweg, T., Klugman, A., Gruzelier, J., & Hirsch, S. R. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophrenia Research*, 69(2–3), 203–217. <https://doi.org/10.1016/j.schres.2003.09.009>
- Baldeweg, T., Williams, J. D., & Gruzelier, J. H. (1999). Differential changes in frontal and sub-temporal components of mismatch negativity. *International Journal of Psychophysiology*, 33(2), 143–148. [https://doi.org/10.1016/s0167-8760\(99\)00026-4](https://doi.org/10.1016/s0167-8760(99)00026-4)
- Belin, P., Fecteau, S., & Bedard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, 8(3), 129–135. <https://doi.org/10.1016/j.tics.2004.01.008>
- Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, 83, 120–131. <https://doi.org/10.1016/j.ijpsycho.2011.08.003>

- Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, 14(2), 402–416. <https://doi.org/10.1111/j.1467-7687.2010.00990.x>
- Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341–345.
- Bouchon, C., Nazzi, T., & Gervain, J. (2015). Hemispheric asymmetries in repetition enhancement and suppression effects in the newborn brain. *Plos One*, 10(10), Article e0140160. <https://doi.org/10.1371/journal.pone.0140160>
- Cacciaglia, R., Costa-Faidella, J., Zarnowiec, K., Grimm, S., & Escera, C. (2019). Auditory predictions shape the neural responses to stimulus repetition and sensory change. *Neuroimage*, 186, 200–210. <https://doi.org/10.1016/j.neuroimage.2018.11.007>
- Charpentier, J., Kovarski, K., Houy-Durand, E., Malvy, J., Saby, A., Bonnet-Brilhault, F., Latinus, M., & Gomot, M. (2018). Emotional prosodic change detection in autism spectrum disorder: An electrophysiological investigation in children and adults. *Journal of Neurodevelopmental Disorders*, 10, 28. <https://doi.org/10.1186/s11689-018-9246-9>
- Cooper, R., Atkinson, R., Clark, R. A., & Michie, P. T. (2013). Event-related potentials reveal modelling of auditory repetition in the brain. *International Journal of Psychophysiology*, 88, 74–81. <https://doi.org/10.1016/j.ijpsycho.2013.02.003>
- Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between "what" and "when" in the auditory system: Temporal predictability enhances repetition suppression. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 31(50), 18590–18597. <https://doi.org/10.1523/JNEUROSCI.2599-11.2011>
- Cowan, N., Winkler, I., Teder, W., & Naatanen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 909–921. <https://doi.org/10.1037/0278-7393.19.4.909>
- de Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in Extrastriate Visual cortex. *Cerebral Cortex*, 23, 2235–2244. <https://doi.org/10.1093/cercor/bhs211>
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), 13494–13499. <https://doi.org/10.1073/pnas.93.24.13494>
- Edgington, E. S. (1995). *Randomization tests* (3rd ed.). New York, NY: Marcel Dekker.
- Escera, C., & Malmerica, M. S. (2014). The auditory novelty system: An attempt to integrate human and animal research. *Psychophysiology*, 51, 111–123. <https://doi.org/10.1111/psyp.12156>
- Ferrari, V., Codispoti, M., & Bradley, M. (2017). Repetition and ERPs during emotional scene processing: A selective review. *International Journal of Psychophysiology*, 111, 170–177. <https://doi.org/10.1016/j.ijpsycho.2016.07.496>
- Fiebach, C. J., Gruber, T., & Supp, G. G. (2005). Neuronal mechanisms of repetition priming in occipitotemporal cortex: Spatiotemporal evidence from functional magnetic resonance imaging and electroencephalography. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 25(13), 3414–3422. <https://doi.org/10.1523/JNEUROSCI.4107-04.2005>
- Font-Alaminos, M., Cornella, M., Costa-Faidella, J., Hervás, A., Leung, S., Rueda, I., & Escera, C. (2019). Increased subcortical neural responses to repeating auditory stimulation in children with autism spectrum disorder. *Biological Psychology*, 149, 107807. <https://doi.org/10.1016/j.biopsycho.2019.107807>
- Friston, K. (2005). A theory of cortical responses. *Proceedings of the National Academy of Sciences of the United States of America*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- Gagnepain, P., Chetelat, G., Landeau, B., Dayan, J., Eustache, F., & Lebreton, K. (2008). Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imaging. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 28(20), 5281–5289. <https://doi.org/10.1523/JNEUROSCI.0565-08.2008>
- Garrido, M., Kilner, J., Kiebel, S., Stephan, K., Baldeweg, T., & Friston, K. (2009). Repetition suppression and plasticity in the human brain. *Neuroimage*, 48, 269–279. <https://doi.org/10.1016/j.neuroimage.2009.06.034>
- Gervain, J., Macagno, F., Cogo, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37), 14222–14227. <https://doi.org/10.1073/pnas.0806530105>
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, 10(1), 14–23. <https://doi.org/10.1016/j.tics.2005.11.006>
- Grotheer, M., & Kovacs, G. (2014). Repetition probability effects depend on prior experiences. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 34(19), 6640–6646. <https://doi.org/10.1523/JNEUROSCI.5326-13.2014>
- Grotheer, M., & Kovacs, G. (2015). The relationship between stimulus repetitions and fulfilled expectations. *Neuropsychologia*, 67, 175–182. <https://doi.org/10.1016/j.neuropsychologia.2014.12.017>
- Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, 28(2), 240–244. <https://doi.org/10.1111/j.1469-8986.1991.tb00417.x>
- Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 25(45), 10494–10501. <https://doi.org/10.1523/JNEUROSCI.1227-05.2005>
- Huotilainen, M., Kujala, A., & Alku, P. (2001). Long-term memory traces facilitate short-term memory trace formation in audition in humans. *Neuroscience Letters*, 310(2–3), 133–136. [https://doi.org/10.1016/s0304-3940\(01\)02096-1](https://doi.org/10.1016/s0304-3940(01)02096-1)
- Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology*, 38(4), 723–727.
- Jacobsen, T., Schröger, E., & Alter, K. (2004). Pre-attentive perception of vowel phonemes from variable speech stimuli. *Psychophysiology*, 41(4), 654–659. <https://doi.org/10.1111/1469-8986.2004.00175.x>
- Jacobsen, T., Schröger, E., Winkler, I., & Horváth, J. (2005). Familiarity affects the processing of task-irrelevant auditory deviance. *Journal of Neurodevelopmental Disorders*, 17(11), 1704–1713. <https://doi.org/10.1162/089892905774589262>
- James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology: CB*, 10(17), 1017–1024. [https://doi.org/10.1016/s0960-9822\(00\)00655-2](https://doi.org/10.1016/s0960-9822(00)00655-2)
- Kaliukhovich, D. A., & Vogels, R. (2012). Stimulus repetition affects both strength and synchrony of macaque inferior temporal cortical activity. *Journal of Neurophysiology*, 107(12), 3509–3527. <https://doi.org/10.1152/jn.00059.2012>
- Kaliukhovich, D. A., & Vogels, R. (2014). Neurons in macaque inferior temporal cortex show no surprise response to deviants in visual oddball sequences. *Journal of Neuroscience*, 34(38), 12801–12815. <https://doi.org/10.1523/JNEUROSCI.2154-14.2014>
- Kouider, S., Eger, E., Dolan, R., & Henson, R. N. (2009). Activity in face-responsive brain regions is modulated by invisible, attended faces: Evidence from masked priming. *Cerebral Cortex*, 19(1), 13–23. <https://doi.org/10.1093/cercor/bhn048>

- Kreiman, J., & Gerratt, B. (2012). Perceptual interaction of the harmonic source and noise in voice. *The Journal of the Acoustical Society of America*, 131(1), 492–500. <https://doi.org/10.1121/1.3665997>
- Larson, M. J., & Carbine, K. A. (2017). Sample size calculations in human electrophysiology (EEG and ERP) studies: A systematic review and recommendations for increased rigor. *International Journal of Psychophysiology*, 111, 33–41. <https://doi.org/10.1016/j.ijpsycho.2016.06.015>
- Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology*: CB, 21(4), R143–R145. <https://doi.org/10.1016/j.cub.2010.12.033>
- Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, 254(5036), 1377–1379. <https://doi.org/10.1126/science.1962197>
- Müller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science*, 285(5432), 1405–1408. <https://doi.org/10.1126/science.285.5432.1405>
- Näätänen, R., Gaillard, A. W., & Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313–329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, 42(1), 25–32. <https://doi.org/10.1111/j.1469-8986.2005.00256.x>
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385(6615), 432–434. <https://doi.org/10.1038/385432a0>
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 118(12), 2544–2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Nelken, I., Yaron, A., Polterovich, A., & Hershenson, I. (2013). Stimulus-specific adaptation beyond pure tones. *Advances in Experimental Medicine and Biology*, 787, 411–418. [https://doi.org/10.1007/978-1-4614-1590-9\\_45](https://doi.org/10.1007/978-1-4614-1590-9_45)
- Obleser, J., Elbert, T., Lahiri, A., & Eulitz, C. (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Cognitive Brain Research*, 15(3), 207–213. [https://doi.org/10.1016/s0926-6410\(02\)00193-3](https://doi.org/10.1016/s0926-6410(02)00193-3)
- Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C., & Malmierca, M. S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nature Communications*, 8, 2148. <https://doi.org/10.1038/s41467-017-02038-6>
- Pérez-González, D., Hernandez, O., Covey, E., & Malmierca, M. (2012). GABA(A)-mediated inhibition modulates stimulus-specific adaptation in the inferior colliculus. *PLoS One*, 7, 34297.
- Pinheiro, A. P., Barros, C., Vasconcelos, M., Obermeier, C., & Kotz, S. A. (2017). Is laughter a better vocal change detector than a growl. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 92, 233–248. <https://doi.org/10.1016/j.cortex.2017.03.018>
- Recasens, M., Leung, S., Grimm, S., Nowak, R., & Escera, C. (2015). Repetition suppression and repetition enhancement underlie auditory memory-trace formation in the human brain: An MEG study. *Neuroimage*, 108, 75–86. <https://doi.org/10.1016/j.neuroimage.2014.12.031>
- Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: A cognitive neuroscience perspective. *Nature Reviews Neuroscience*, 5(11), 853–862. <https://doi.org/10.1038/nrn1534>
- Scheerer, N. E., Behich, J., Liu, H., & Jones, J. A. (2013). ERP correlates of the magnitude of pitch errors detected in the human voice. *Neuroscience*, 240, 176–185. <https://doi.org/10.1016/j.neuroscience.2013.02.054>
- Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia*, 51(1), 59–66. <https://doi.org/10.1016/j.neuropsychologia.2012.11.006>
- Shestakova, A., Brattico, E., Soloviev, A., Klucharev, V., & Huottilainen, M. (2004). Orderly cortical representation of vowel categories presented by multiple exemplars. *Cognitive Brain Research*, 21(3), 342–350. <https://doi.org/10.1016/j.cogbrainres.2004.06.011>
- Summerfield, C., Monti, J., Trittschuh, E., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, 11(9), 1004–1006. <https://doi.org/10.1038/nn.2163>
- Summerfield, C., Wyart, V., Mareike Johnen, V., & de Gardelle, V. (2011). Human scalp electroencephalography reveals that repetition suppression varies with expectation. *Frontiers in Human Neuroscience*, 5, 67. <https://doi.org/10.3389/fnhum.2011.00067>
- Todorovic, A., & Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 32(39), 13389–13395. <https://doi.org/10.1523/JNEUROSCI.2227-12.2012>
- Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: An MEG study. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 31(25), 9118–9123. <https://doi.org/10.1523/JNEUROSCI.1425-11.2011>
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Neurodevelopmental Disorders*, 21(10), 1934–1945. <https://doi.org/10.1162/jocn.2009.21131>
- Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 24(46), 10440–10453. <https://doi.org/10.1523/JNEUROSCI.1905-04.2004>
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, 6(4), 391–398. <https://doi.org/10.1038/nn1032>
- Vogels, R. (2016). Sources of adaptation of inferior temporal cortical responses. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior*, 80, 185–195. <https://doi.org/10.1016/j.cortex.2015.08.024>
- von der Behrens, W., Bäuerle, P., Kössl, M., & Gaese, B. H. (2009). Correlating stimulus-specific adaptation of cortical neurons and local field potentials in the awake rat. *Journal of Neuroscience*, 29(44), 13837–13849. <https://doi.org/10.1523/JNEUROSCI.3475-09.2009>
- Whitten, A., Key, A. P., Mefferd, A. S., & Bodfish, J. W. (2020). Auditory event-related potentials index faster processing of natural speech but not synthetic speech over nonspeech analogs in children. *Brain and Language*, 207, 104825. <https://doi.org/10.1016/j.bandl.2020.104825>
- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, 13(12), 532–540. <https://doi.org/10.1016/j.tics.2009.09.003>
- Winkler, I., Schroger, E., & Cowan, N. (2001). The role of large-scale memory organization in the mismatch negativity event-related brain potential. *Journal of Neurodevelopmental Disorders*, 13(1), 59–71. <https://doi.org/10.1162/089892901564171>
- Ylinen, S., & Huottilainen, M. (2007). Is there a direct neural correlate for memory-trace formation in audition? *Neuroreport*, 18(12), 1281–1284. <https://doi.org/10.1097/WNR.0b013e32826fb38a>

### 1.3. Étude n°2 : Effet de la prosodie émotionnelle sur l'adaptation neurale à des sons vocaux et non-vocaux dans l'autisme

#### Unravelling neural adaptation to vocal and non-vocal sounds in autism: effect of emotional prosody

*Merchie, A., Bonnet-Brilhault, F., Escera, C., Houy-Durand, E. & Gomot, M.*

*Article publié dans Clinical Neurophysiology (2025)*

**Introduction et objectifs :** Dans l'autisme, une recherche d'immuabilité et une résistance au changement ont été décrites au niveau comportemental. L'intolérance au changement a aussi été observée au niveau neurophysiologique avec des atypies de la détection du changement mises en évidence par l'étude de la MMN. Ces différences de réponse physiologique au changement pourraient être la conséquence d'une adaptation atypique au contexte. La voix joue un rôle social crucial, il est donc essentiel de s'y adapter pour établir une communication efficace. Le contenu prosodique, caractérisé par un grand nombre d'informations à encoder, pourrait aussi moduler cette adaptation. Cette étude a donc pour objectif d'évaluer l'adaptation neurale à des sons vocaux et non-vocaux, mais aussi de déterminer l'effet d'un contenu prosodique émotionnel sur ces processus dans un groupe d'adultes autistes et neurotypiques.

**Matériel et méthodes :** Les réponses neurophysiologiques à des sons vocaux et non-vocaux (neutres et émotionnels) ont été mesurées chez 20 adultes autistes et 20 adultes neurotypiques. Une séquence de *Roving* de 25min a été présentée avec des trains de 4, 8 et 14 répétitions. Afin d'estimer l'existence et la dynamique d'adaptation aux différents sons, les amplitudes des réponses cérébrales moyennées selon des groupes de répétitions (2-4, 5-8, 9-11 et 12-14) et des RP (12-14 – 2-4 répétitions) ont été mesurées.

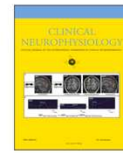
**Résultats :** Aucune différence de RP n'a été observée selon la nature du son (vocale ou non-vocale) dans chacun des deux groupes, reflétant une adaptation identique pour des sons de différentes natures. Cependant, les dynamiques de mise en place de cette adaptation ne sont pas les mêmes. Chez les adultes neurotypiques, cette étude confirme les résultats de l'étude n°1 avec une adaptation plus rapide aux sons non-vocaux qu'aux sons vocaux. Néanmoins, aucun effet du contenu émotionnel n'a été mis en lumière. Au contraire, dans l'autisme, aucune modulation de la réponse n'est observée pour les sons non-vocaux alors qu'une adaptation plus rapide à la voix a été mise en évidence. Chez les personnes autistes, le contenu émotionnel conduit à une adaptation plus lente que pour les sons neutres.

**Discussion :** Cette étude est la première à étudier l'adaptation neurale à des sons de différentes natures avec un paradigme adapté dans l'autisme. Une adaptation neurale atypique est mise en évidence chez les adultes autistes quelle que soit la nature du son, possiblement impliquée dans les difficultés de traitement du changement et la recherche d'immuabilité. De plus, dans l'autisme, une adaptation plus rapide aux sons sociaux et émotionnels qu'aux sons de l'environnement est observée. Ces résultats peuvent être interprétés au regard de la théorie de la motivation sociale (Chevallier et al., 2012). Selon cette théorie, les adultes non autistes sont plus motivés par les stimuli sociaux que les adultes autistes. Ainsi, l'adaptation rapide à la voix chez les autistes refléterait un désintérêt pour celle-ci.



Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: [www.elsevier.com/locate/clinph](http://www.elsevier.com/locate/clinph)

## Unraveling neural adaptation to vocal and non-vocal sounds in autism

Annabelle Merchie<sup>a,b</sup>, Frédérique Bonnet-Brilhault<sup>a,b,c</sup>, Carles Escera<sup>d,e,f</sup>,  
Emmanuelle Houy-Durand<sup>a,b,c</sup>, Marie Gomot<sup>a,b,\*</sup><sup>a</sup> Université de Tours, INSERM, Imaging Brain and Neuropsychiatry iBrain U1253, 37032, Tours, France<sup>b</sup> EXACT, Centre Universitaire de Pédiopsychiatrie, CHRU de Tours, Tours, France<sup>c</sup> Centre Ressource Autisme région Centre-Val de Loire, CHRU de Tours, Tours, France<sup>d</sup> Neuroscience Research Group, Department of Clinical Psychology and Psychobiology, Brainlab - Cognitive, University of Barcelona, Catalonia, Spain<sup>e</sup> Institute of Neurosciences, University of Barcelona, Catalonia, Spain<sup>f</sup> Institut de Recerca Sant Joan de Déu, Santa Rosa 39-57, 08950, Esplugues de Llobregat, Catalonia, Spain

## ARTICLE INFO

## Keywords:

Auditory Evoked Potentials  
Autism  
Neural Adaptation  
Repetition Positivity  
Voice

## ABSTRACT

**Objective:** Autism is linked to a strong need for sameness and difficulties in social communication, associated with atypical brain responses to voices and changes. This study aimed to characterize neural adaptation in autistic adults using a Roving paradigm and assess how vocal vs. non-vocal, as well as neutral vs. emotional sounds, influence this adaptation.**Methods:** Neural adaptation was measured in 20 autistic and 20 non-autistic adults using a Roving paradigm, where sounds were repeated 4, 8, or 14 times. Neural responses and Repetition Positivity (RP) amplitudes were analyzed as indices of adaptation.**Results:** RP amplitudes showed no significant differences between groups for vocal or non-vocal sounds, but adaptation dynamics varied. Non-autistic adults adapted more quickly to non-vocal (5–8 repetitions) compared to vocal sounds (12–14 repetitions). In contrast, autistic adults adapt faster to vocal than to non-vocal sounds. Moreover emotional prosodic content influenced RP amplitude in autistic adults only, suggesting heightened sensitivity to emotional cues in social contexts.**Conclusions:** The study highlights how atypical neural adaptation in autism how emotional content impacts social communication deficits. These insights enhance understanding of autism-related adaptation challenges.

## 1. Introduction

Autism is a neurodevelopmental disorder characterized by two major categories of symptoms: restricted and repetitive behaviors and socio-communicative impairments (DSM-5; American Psychiatric Association, 2022). Atypicalities in processing deviant stimuli – those not conforming to a rule – have been demonstrated consistently in autistic adults (Chen et al., 2020; Kujala et al., 2010, 2007) and children (Charpentier et al., 2018; Gomot et al., 2002). These atypicalities may result from two different complementary processes: difficulties in deviance detection and/or in the process of conforming to a rule in a repetitive context, a process subserved by neural adaptation. Moreover, production and perception of emotional prosody represent a major difficulty in autism (Boucher et al., 2000; McCann and Peppé, 2003; Peppé et al., 2007) (for a review see Asghari et al., 2021), and an atypical electrophysiological response to change in emotional prosody has been

observed (Charpentier et al., 2018). Thus, understanding the neural adaptation process to emotional sounds in this clinical population is essential for assessing its impact on communication and interaction difficulties, and would contribute to the comprehension of both dimensions of autistic symptomatology.

## 1.1. Adaptation and deviance detection in autism

In autism, atypicalities in deviance detection are reflected by differences in latency and amplitude of the mismatch negativity (MMN) and have been repeatedly reported (Chen et al., 2020; Gomot et al., 2002, 2011). Hypotheses about difficulties in change detection in autism have been proposed with an emphasis on the framework of the Bayesian brain (Lawson et al., 2014; Palmer et al., 2017; Pellicano et al., 2022). It has been proposed that the atypical change detection could be the consequence of particular prediction processes and thus difficulties in

\* Corresponding author.

E-mail address: [gomot@univ-tours.fr](mailto:gomot@univ-tours.fr) (M. Gomot).<https://doi.org/10.1016/j.clinph.2024.12.003>

Accepted 8 December 2024

Available online 11 December 2024

1388-2457/© 2024 International Federation of Clinical Neurophysiology. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

adaptation to the context (Gomot and Wicker, 2012).

Studies have reported an absence or a lack of neural adaptation in autism, but these studies used paradigms that were not originally designed for this purpose (i.e. oddball paradigm) (Cary et al., 2023; Gonzalez-Gadea et al., 2015; Guiraud et al., 2011; Jamal et al., 2020; Kolesnik et al., 2019; Martineau et al., 1992; Ruiz-Martínez et al., 2020) or compared only a limited number of repetitions (Cary et al., 2023; Gonzalez-Gadea et al., 2015; Kolesnik et al., 2019; Ruiz-Martínez et al., 2020). These studies also did not compare different categories of sounds or stimuli, and mostly used pure tones. However a suitable paradigm to specifically measure neural adaptation, the “Roving paradigm” has been developed (Cowan et al., 1993). In this type of paradigm, a stimulus is repeated several times and followed by another one also repeated multiple times. The first presentation of a stimulus is referred as a deviant (as in oddball paradigm) and then, standards stimuli are repeated. This paradigm allows to measure neural adaptation thanks to a constant updating of the memory trace (Baldeweg et al., 2004, 1999; Cowan et al., 1993; Haenschel et al., 2005). In terms of Event Related Potentials (ERPs) modulation, the repeated presentation of a stimulus results in a decrease of the N1 component, between 90 and 150 ms and an increase in the positive components P1 and P2, reflecting the adaptation of responses. Roving paradigm also allows for the measurement of the Repetition Positivity (RP), a positive deflection between 90 and 200 ms that is calculated as the difference between many and few repetitions of the same sounds, with an increasing in amplitude as the number of repetitions increases (Baldeweg et al., 2004; Costa-Faidella et al., 2011; Haenschel et al., 2005). In non-autistic adults, comparison between neural adaptation to vocal and non-vocal sounds in a Roving paradigm revealed that regularity encoding for vocal sounds required at least 9 to 12 repetitions. In contrast for non-vocal complex sounds only 5 to 8 repetitions were sufficient to reach a stable adaptation (Heurteloup et al., 2022). In autism, one study attempted to compare neural adaptation to human and non-human sounds and observed a diminished adaptation for both natures of sounds after 2 repetitions of the standard (Ruiz-Martínez et al., 2020), but this study was still using an oddball paradigm, which has been demonstrated to be non-optimal for evidencing neural adaptation processes (Cowan et al., 1993). In autistic children only one study used a Roving paradigm to explore adaptation, focusing on subcortical responses, and showed deficits in neural adaptation in brain responses to pure tones (Font-Alaminos et al., 2020). A review on these adaptation processes at both the behavioral and neural levels in autism and in other neurodevelopmental disorders has recently been published (Merchie and Gomot, 2023). These results need to be confirmed in further studies through measurement of cortical indices but also in response to more complex sounds such as voice, which are essential for effective communication.

### 1.2. Voice and emotional prosody processing in autism

A model of voice processing has been proposed, highlighting a particular pathway (Belin et al., 2004) and a specific brain area (TVA: Temporal Voice Area; Belin et al., 2000). This specificity of processing is also reflected in the specific electrophysiological response to voice: the Fronto-Temporal Positivity to Voice (FTPV) (Bruneau et al., 2013). An atypical FTPV in autistic children has been observed, may actually be driven by an atypical processing of non-vocal sounds (Bidet-Caulet et al., 2017) and by heterogeneity in response to voice and environmental sounds between groups and within the same individual (Latinus et al., 2019). Electrophysiological and neuroimaging results have also confirmed behaviorally, as autistic adults judged non-vocal sounds as more pleasant than vocal sounds, contrary to non-autistic participants (Michel et al., 2023).

A reduced amplitude of the electrophysiological response to emotional prosody in autism has also been observed (Fan and Cheng, 2014; Korpilahti et al., 2006). When looking at the brain connectivity in response to emotional prosody in autistic children for both sadness and

happiness, a non-coherent connection between the TVA and the temporoparietal junction has been observed (Leipold et al., 2022). However other authors reported a positive evolution of responses to prosodic vocal stimuli from childhood to adulthood in autism (Charpentier et al., 2018).

### 1.3. The present study

The aim of the current study was twofold: first to study neural adaptation in an autistic population with a paradigm designed specifically for this purpose, the Roving paradigm, and second, to assess the modulation of this adaptation process according to the social nature of the stimulus by comparing prosodic vocal and non-vocal sounds. This study in autistic adults should inform on both dimensions of the symptomatology: the processing of voice for socio-communicative deficits and the regularity encoding for repetitive and restricted behaviors, but also on their interaction. According to previous findings, an atypical neural adaptation should shed light on deviance detection difficulties and the presence of social content could induce greater difficulties in autistic adults.

## 2. Methods

### 2.1. Population

Twenty-one autistic and 20 non-autistic adults were recruited in the current study. Autistic adults were recruited through the Autism Resource Center of Tours and diagnosed by a trained clinical team according to DSM-5 criteria (American Psychiatric Association, 2022). The diagnosis was confirmed with ADOS-2 and ADI-R (Lord et al., 2000, 1994). Verbal and non-verbal abilities were estimated through the use of selected subtests of the WAIS-IV (Wechsler, 2012) in both groups. Also, the Autism Quotient 50-question survey was proposed to each participant to estimate autistic traits in the entire sample (Baron-Cohen et al., 2001). One participant was discarded because of a doubt about their diagnosis by the clinical team. The group of non-autistic adults reported an absence of any developmental difficulties in language and sensorimotor acquisition. In both groups, no diseases of the central nervous system, infectious or metabolic diseases, epilepsy, or an abnormal audition was reported. Participants had to be off electrogenesis-modifying medication. In the ASD group, most patients were not medicated, and only two participants were receiving a stabilized treatment consisting of Selective Serotonin Reuptake Inhibitors and/or antihistaminic. Each participant (or their legal guardian in the case of an adult under guardianship) signed an informed consent form, and the protocol received approval from the Ethics Committee (PROSCEA 2017/23; ID RCB: 2017-A00756-47). A description of the two groups is presented in Table 1.

### 2.2. Stimuli

The impact of the nature and the emotional content of the stimuli to be encoded was studied by creating and using different categories of

**Table 1**  
Groups characteristics (mean  $\pm$  sd). Verbal and Performance IQ measured with Wechsler intelligence scale (average on 18 autistic and 16 non-autistic adults). ADOS sum of social interaction, communication and restricted and repetitive behavior scores). AQ scores (average on 16 autistic adults).

	Non-autistic	Autistic	Difference
Age	27 $\pm$ 8	31 $\pm$ 2	$p > 0.05$
Gender	13♂ 7♀	17♂ 3♀	$\chi^2 = 1.2; p > 0.05$
Verbal IQ	125 $\pm$ 11	112 $\pm$ 18	$p < 0.05$
Performance IQ	106 $\pm$ 13	98 $\pm$ 14	$p > 0.05$
ADOS	/	12 $\pm$ 6	/
AQ	17 $\pm$ 7	35 $\pm$ 8	$p < 0.05$

sounds. First, two neutral vocal sounds, consisting of the vowel /a/ pronounced with a neutral prosody and two /a/ pronounced with a sad prosody by the same two female speakers were chosen. The selected stimuli had previously been validated and the emotional content (neutral or sad) was well recognized (Charpentier et al., 2018). Moreover, four non-vocal complex synthetic sounds with acoustic characteristics similar to those of the different vocal sounds were created. The objective was to create non-vocal complex sounds with a global frequency spectrum closest to that of the corresponding vocal sounds. To mimic the speech content of the vocal sounds, the synthesized stimuli contain F0, F1 and F2 to provide enough information to automatically categorize the phoneme /a/ (Jacobsen et al., 2004). The same method as in a previously published study was used to synthesize non-vocal sounds (Heurteloup et al., 2022) (see supplementary materials and figure S1 for more details). To ensure acoustic similarity between vocal and synthesized complex non-vocal sounds, the frequencies of the first four formants were compared using a Wilcoxon test. The frequencies of non-vocal complex and vocal sounds were not significantly different ( $W = 201, p = .99$ ). To study adaptation to a same Nature (Vocal or Non-Vocal) or Valence (Neutral or Emotional) of sound, brain responses were averaged for each category.

2.3. Sequence of stimulation

To study the growth of regularity encoding and neural adaptation, a roving paradigm was used (Baldeweg et al., 2004; Cowan et al., 1993) in which several stimulation trains of different sounds and lengths were presented. A train was composed by the same stimulus repeated, in this study, 4, 8 or 14 times. Trains were delivered in a pseudo-random order, the only constraint being that 2 trains composed of the same stimulus could not be consecutive. Each stimulus was presented for 300 ms with a SOA (i.e., Stimulus Onset Asynchrony) of 646 ms via two speakers located at 1.1 m in front of the participants (Logitech Z-2300) with an intensity of 50 dB SPL. Eighty repetitions of the trains of 4 and 8 repetitions were presented and 160 repetitions of the trains of 14 repetitions, resulting in a single sequence of 2304 stimuli. The duration of the protocol was 24 min. A schematic representation of the paradigm is presented in Fig. 1.

2.4. EEG recording and processing

The EEG session took place in a sound-attenuated room where participants were sitting in a reclining armchair. During the delivery of the stimuli, participants watched a silent movie without subtitles, and they

were instructed to answer a question about the film at the end of the session, to keep them awake and avoid directing their attention toward the auditory sequence.

The Presentation software (NeuroBehavioral Systems Inc., Berkeley, CA) was used to deliver the stimuli, while the EEG was recorded with 64-active electrodes according to the extended 10–20 system (Oostenveld & Praamstra, 2001) (ActiveTwo Systems Biosemi, The Netherlands) with a sampling rate of 512 Hz. To monitor horizontal and vertical ocular movements, electrodes were placed on the left and right outer canthi and below the left eye respectively. An external electrode was placed on the nose for offline referencing. To analyze EEG-ERP, the ELAN software was used (Aguera et al., 2011). The EEG signal was amplified and filtered with a 0.1 Hz high-pass filter (Butterworth filter, order 1). After visual inspection, one or two channels were removed, when necessary, for each participant, if excessive noise was present in the signal. An Independent Component Analysis (ICA) was performed with EEGLAB to remove artefacts from ocular movements, especially blinks and saccades which were captured into components and removed with an inverse ICA transformation (Makeig et al., 2004). One or two components in each subject were removed. The motion artifacts (high frequency or high amplitude signals) were removed manually by an experimenter blind to trial type. A 30 Hz low-pass filter was applied (Butterworth filter, order 3), and ERPs were averaged over a 700 ms time window, including a 100 ms pre-stimulus baseline. ERPs from rejected channels were interpolated from the remaining electrodes using spherical spline interpolation.

2.5. Data analysis

Responses to each category of sounds, in each of the positions from 1 to 14, were averaged individually for each sequence, so that the effect of repetition for each position could be investigated. Trials presented at the same position, but in trains of different lengths (e.g., stimulus in position 3 in a train of 4, 8 and 14 repetitions) were also averaged together. This resulted in three times more stimuli from 1 to 4, and twice as many stimuli from 5 to 8, compared to stimuli from positions 9 to 14. Afterwards, responses to stimuli in positions 2 to 4 were averaged, as well as those in positions 5 to 8, those in positions 9 to 11, and finally those in positions 12 to 14, to obtain more robust responses, resulting in four clusters of repetitions. The number of artifacts free trials was kept above 1686 ( $\pm 295$ ) in non-autistic group and 1807 ( $\pm 319$ ) in autistic group (mean  $\pm$  standard deviation).

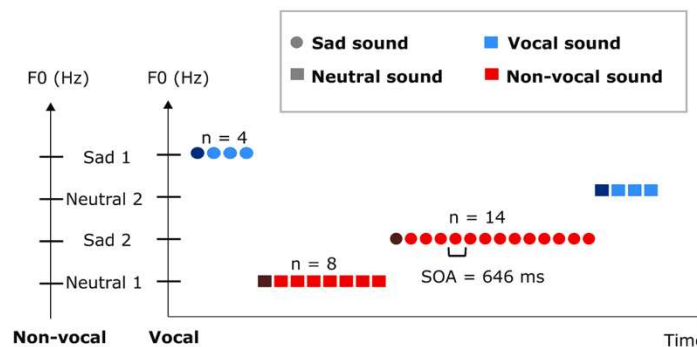


Fig. 1. Roving paradigm and acoustic description of sounds used in this study. (A) Roving paradigm with trains of 4, 8 or 14 repetitions of the same stimulus were pseudo-randomly delivered. The SOA (i.e., Stimulus Onset Asynchrony) was constant at 646 ms. Vocal sounds (blue) were the French vowel /a/ uttered by 2 female speakers with 2 emotional prosodies (neutral and sad). Non-vocal sounds (red) were synthesized as complex sounds with acoustic characteristics close to vocal sounds. (B).

2.5.1. Measurements

For the different groups of repeated sounds and according to the nature of the sound, it was possible to clearly identify several electrophysiological components. N1 was elicited in response to non-vocal complex sounds for the four groups of repetitions. The N100 is the component of choice to investigate neural adaptation. However, in response to vocal sounds, only the P1 (the first positive deflection) component could be identified. These two electrophysiological components were thus measured to study the time course of the establishment of regularity. Based on previous studies (Jordi Costa-Faidella et al., 2011; Haenschel et al., 2005; Heuteloup et al., 2022) and after visual inspection of the responses on the scalp, the analyses of amplitudes and latencies were done over the frontal electrode Fz. For each component of interest (P1 and N100), the peak latency was determined based on the group-averaged ERPs within fixed time windows (P1: 50–100 ms; N100: 100–150 ms). The measure of latency and amplitude was made for each participant in a 40 ms time window, centered on the previously identified peak of interest ( $\pm 20$  ms) of the grand-mean ERPs of the group for each condition (Table 2).

The difference between 12 and 14 and 2 to 4 repetitions was considered as reflecting overall neural adaptation and was calculated as the Repetition Positivity (RP). The RP was identified as a positive deflection between 90 and 200 ms, and the measure of its amplitude was made in this time window according to previous study with similar protocol (Pinheiro et al., 2017).

2.5.2. Statistical analysis

Statistical analyses were conducted on RStudio 4.0.4 (Team, 2021, 2020) with the packages ggplot2 (Wickham, 2016), ggpubr (Kassambara, 2020), dplyr (Wickham et al., 2021), tidyverse (Wickham et al., 2019) and ez (Lawrence, 2016).

First, to evaluate plausible amplitude differences according to group, comparisons of P1 and N100 amplitude for responses to vocal and non-vocal sounds were performed with Student's *t*-test for each cluster of repetitions (2–4, 5–8, 9–11, and 12–14).

A repeated measures ANOVA was performed in each group on the peak amplitude of each component, according to the number of repetitions (2–4, 5–8, 9–11, and 12–14) for each category of sounds (vocal, non-vocal, neutral, and emotional). As the number of levels was greater than two, a Greenhouse-Geisser correction was applied to check the sphericity, and epsilon ( $\epsilon$ ) was reported. To specify repetition effects, FDR-corrected pairwise Student's *t*-tests were performed in case of statistically significant main effects.

Only a qualitative description of neural adaptation dynamics was performed for studying adaptation to vocal sounds according to valence (emotional and neutral) because of the reduced number of trials in each category, which induced greater data variability.

For RP, a repeated measures ANOVA was performed on mean

amplitude in the 90–200 ms time window to evaluate the effect of sound nature (vocal vs non-vocal), valence (emotional vs neutral), and group (ASD vs TD). Then separate analyses were performed within each group to specify the effect. To evaluate if the RP was measurable, a comparison of the mean amplitude against a null amplitude was performed with a Student's *t*-test ( $\mu = 0$ ) within each group for both responses to vocal and non-vocal sounds.

3. Results

3.1. Effect of nature: Vocal vs non-vocal

3.1.1. Difference of amplitude between groups

Student's *t*-test revealed no statistical difference between groups in P1 amplitude to vocal sounds for any clusters of repetitions ( $t_{2,4}(37) = -.75, p = .46; t_{5,8}(34) = -.02, p = 0.98; t_{9,11}(38) = -.19, p = .85; t_{12,14}(37) = -.93, p = .36$ ). There was no N100 amplitude difference between groups either in response to non-vocal sounds ( $t_{2,4}(37) = 1.16, p = .25; t_{5,8}(37) = .85, p = .40; t_{9,11}(29) = .50, p = .62; t_{12,14}(37) = -.37, p = .71$ ).

3.1.2. Adaptation dynamics

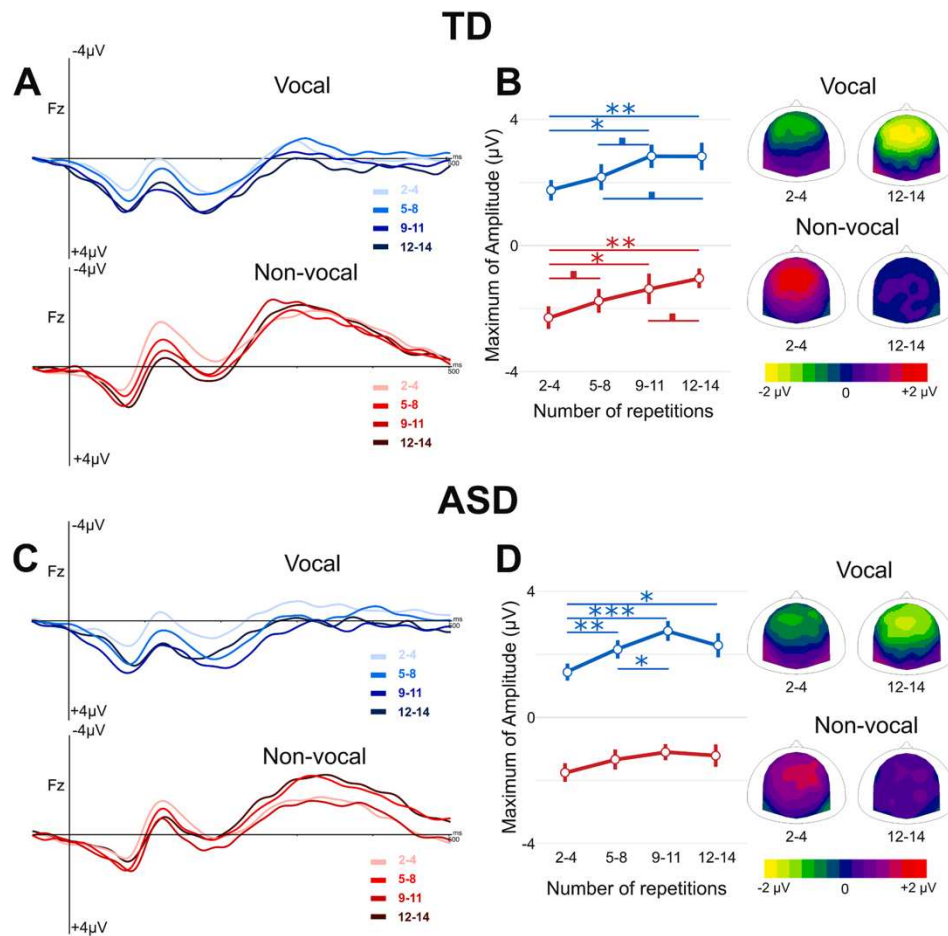
3.1.2.1. Vocal sounds. In the non-autistic group, in response to vocal sounds, a repeated measures ANOVA on P1 amplitude (Fig. 2A, upper part) with repetitions as a within-subject factor revealed a main effect of repetitions ( $F(3,57) = 3.21, p = .03, \eta^2 = .05, \epsilon = .95$ ). Looking at the adaptation pattern with comparisons between repetitions, an increase in P1 amplitude was observed, and it appeared that 12 to 14 repetitions were required to observe a stabilized adaptation (Fig. 2B, upper part). In the autistic group, a main effect of the number of repetitions on P1 amplitude was also demonstrated ( $F(3,57) = 5.81, p = .002, \eta^2 = .08, \epsilon = 0.87$ ) (Fig. 2D, upper part), but a different pattern from the non-autistic group was observed with an increase in P1 amplitude up to 9 to 11 repetitions, followed by a decrease (Fig. 2C, upper part). Statistical pairwise comparisons are presented in Fig. 2B and 2D.

3.1.2.2. Non-Vocal sounds. The same analysis was performed for the response to non-vocal sounds through the measurement of N100 amplitude. In the non-autistic group, a main effect of repetitions was observed ( $F(3,57) = 4.98, p = 0.004, \eta^2 = .07, \epsilon = .88$ ), revealing a decrease in N100 amplitude with repetitions (Fig. 2A, lower part). Nine to eleven repetitions were sufficient to observe the initiation of neural adaptation to non-vocal sounds (Fig. 2B, lower part). In contrast, no modulation of N100 amplitude was observed in the autistic group ( $F(3,57) = 1.78, p = .16$ ) (Fig. 2C, lower part), regardless of the number of repetitions of a non-vocal sound (Fig. 2D, lower part). Statistical pairwise comparisons are presented in Fig. 2B and 2D.

Table 2

Latencies of the maximum of amplitude of P1 and N100 components in ms for each nature of sounds and each repetition groups. The time windows were centered on the maximum of amplitude in the mean group ERPs  $\pm 20$  ms.

	P1							
	2–4		5–8		9–11		12–14	
	TD	ASD	TD	ASD	TD	ASD	TD	ASD
v	64–104	56–96	58–98	54–94	58–98	58–98	54–94	50–90
n	64–104	56–96	50–90	50–90	56–96	54–94	54–94	50–90
s	68–108	58–98	60–100	58–98	62–102	60–100	48–88	50–90
	N100							
	2–4		5–8		9–11		12–14	
	TD	ASD	TD	ASD	TD	ASD	TD	ASD
c	105–145	103–143	101–141	97–137	101–141	101–141	99–139	103–143
cn	99–139	91–131	97–137	99–139	95–135	89–129	97–137	99–139
cs	107–147	105–145	103–143	97–137	103–143	105–145	103–143	103–143



**Fig. 2.** Brain response to vocal and non-vocal sounds according to the number of repetitions in autistic (ASD) and non-autistic (TD) adults. (A) Standard ERPs at Fz electrode after 2–4, 5–8, 9–11 and 12–14 repetitions in response to vocal (blue – top) and non-vocal (red – below) sounds TD adults. (B) P1 amplitude for vocal and N100 amplitude for non-vocal sounds in the TD group plotted for 2–4, 5–8, 9–11 and 12–14 repetitions at Fz (left part) and scalp potential distribution for P1 and N100 after 2–4 and 12–14 repetitions for vocal (blue – top) and non-vocal (red – below) condition. (C) Standard ERPs at Fz electrode after 2–4, 5–8, 9–11 and 12–14 repetitions in response to vocal (blue – top) and non-vocal (red – below) sounds in ASD adults. (D) P1 amplitude for vocal and N100 amplitude for non-vocal sounds in the ASD group plotted for 2–4, 5–8, 9–11 and 12–14 repetitions at Fz (left part) and scalp potential distribution for P1 and N100 after 2–4 and 12–14 repetitions for vocal (top) and non-vocal (below) condition. ·  $p_{FDR-corr} < .1$ ; \*  $p_{FDR-corr} < .05$ ; \*\*  $p_{FDR-corr} < .01$ ; \*\*\*  $p_{FDR-corr} < .001$ .

**3.1.3. Repetition positivity**

Comparisons of RP amplitude in the 90–200 ms time window were performed against a null amplitude to confirm the presence or absence of response for each sound Nature (vocal and non-vocal). Analysis revealed that in the non-autistic group, the RP was different from zero for both vocal ( $t(19) = 2.17, p = .04$ ) and non-vocal ( $t(19) = 4.10, p = .0006$ ) sounds (Fig. 3A). However, in the autistic group, a difference from zero was observed for vocal ( $t(19) = 3.41, p = 0.003$ ) but not for non-vocal ( $t(19) = 1.54, p = .14$ ) sounds (Fig. 3A). This result confirmed the previous analysis, showing no modulation of response amplitude to non-vocal sounds in the autistic group.

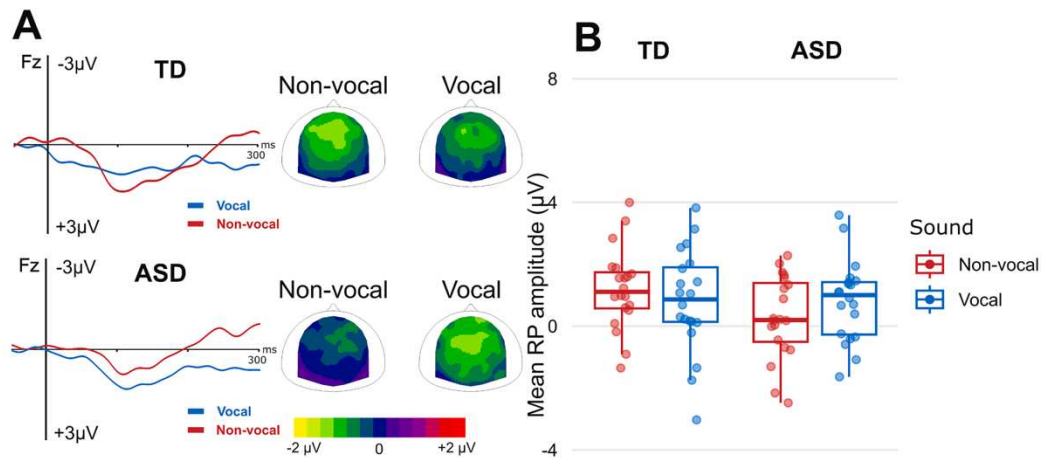
In each group, a comparison of RP amplitude between responses to vocal and non-vocal sounds was performed. This comparison revealed

no significant difference in the non-autistic group ( $t(19) = 1.05, p = .31$ ), nor in the autistic group ( $t(19) = -1.72, p = .10$ ) (Fig. 3B).

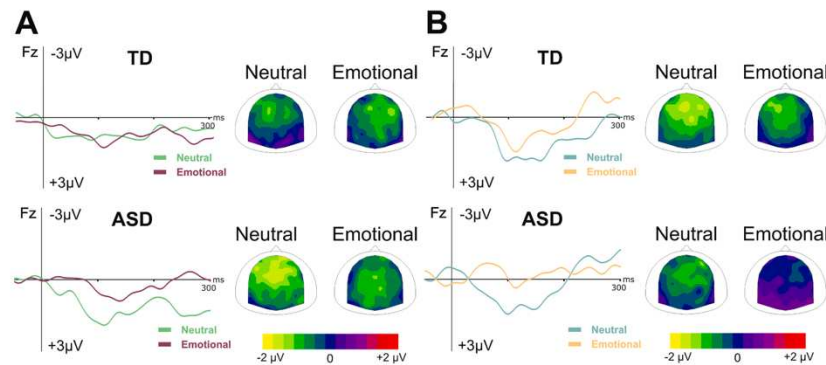
**3.2. Effect Valence: Neural vs emotional**

**3.2.1. Repetition positivity**

A repeated measures ANOVA on mean RP amplitude with Nature (Vocal vs Non-Vocal) and Valence (Neutral vs Emotional) as within factors and Group (ASD vs TD) as the between factor (Figs. 4 and 5) was performed to evaluate neural adaptation according to the emotional content of sounds. This analysis revealed no significant effect of Group ( $F(1,38) = .61, p = .44$ ). Distinct analysis performed for each group showed in the non-autistic group, no effect of Nature ( $F(1,19) = 1.04, p$



**Fig. 3.** Repetition Positivity (RP) to vocal and non-vocal sounds in autistic (ASD) and non-autistic (TD) adults. (A) RP resulting from the difference wave between 12–14 and 2–4 repetitions for vocal (blue) and non-vocal (red) conditions in TD (top) and ASD (below) with RP scalp potential distribution between 90 and 200 ms for both conditions (right part). (B) RP mean amplitude in the 90–200 ms time window for vocal (blue) and non-vocal (red) conditions.



**Fig. 4.** (A) Repetition Positivity (RP) resulting from the difference wave between 12–14 and 2–4 repetitions for vocal neutral (green) and emotional (purple) conditions in TD (top) and ASD (below) with RP scalp potential distribution between 90 and 200 ms for both conditions. (B) Repetition Positivity (RP) resulting from the difference wave between 12–14 and 2–4 repetitions for non-vocal neutral (aqua) and emotional (orange) conditions in TD (top) and ASD (below) with RP scalp potential distribution between 90 and 200 ms for both conditions.

= .32), Valence ( $F(1,19) = .22, p = .64$ ) nor interaction ( $F(1,19) = .87, p = .36$ ) (Fig. 5). In the autistic group no effect of Nature ( $F(1,19) = 3.20, p = .09, \eta^2 = .02$ ) nor interaction between Nature and Valence ( $F(1,19) = .12, p = .74$ ) were observed. However, a significant effect of Valence was revealed ( $F(1,19) = 6.12, p = .02, \eta^2 = .07$ ) with RP amplitude being larger for Neutral (either vocal or non-vocal) than for Emotional sounds ( $t(39) = 2.46, p = .02$ ) (Fig. 5).

Considering the limited number of trials for these categories of sounds (Table S1), the analysis of adaptation patterns was performed and presented in the supplementary materials for vocal sounds according to valence (Emotional and Neutral) in both groups (Figure S2 and Table S2).

#### 4. Discussion

This study is the first, to our knowledge, to evaluate neural

adaptation in response to different natures and valences of sounds in autistic adults using an adapted roving paradigm. Results showed an atypical and unstable regularity encoding in autistic adults, irrespective of the sound category, in comparison to non-autistic adults and a modulation of neural adaptation by sound nature and valence in both groups. Thus, no adaptation was observed for non-vocal sounds in autistic participants, whereas rapid adaptation (5–8 repetitions) was found for non-autistic participants. For vocal sounds, adaptation was faster in autistic (5–8 repetitions) than in non-autistic adults (9–11 repetitions), but it was not maintained over time. Neural adaptation was also sensitive to emotional valence in autistic adults.

##### 4.1. Atypical neural adaptation in autism

The current study revealed a general deficit of sound regularity encoding in autistic adults. This might contribute to the need for

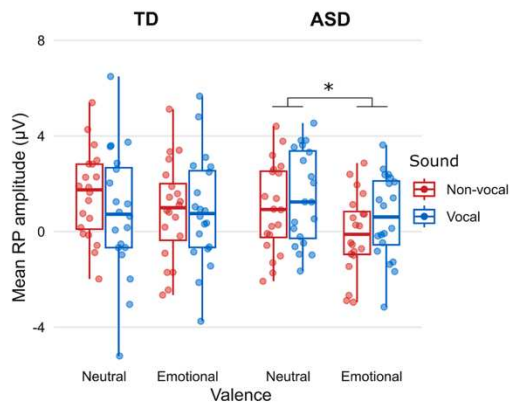


Fig. 5. RP mean amplitude according to sounds valence in autistic (ASD) and non-autistic (TD) adults. RP mean amplitude in the 90–200 ms time window for vocal (blue) and non-vocal (red) conditions according to emotional valence in TD and ASD groups. \* $p < .05$ .

sameness and attachment to routines (Crespo et al., 2013) reported in the clinical descriptions of autism. Difficulties in tolerating changes in autism could also be partly explained by neural adaptation atypicalities (Buhr and Dugas, 2009). Regarding neural response to change in autism, an altered mismatch negativity (MMN) was reported in response to different natures of sounds (Charpentier et al., 2018; Gomot et al., 2002, 2011). In the present study, atypicalities were reported in response to both social and non-social stimuli, indicating that change detection could be altered no matter the content, reflecting a global deficit in regularity encoding processes. These results are compatible with behavioral hyperreactivity to sensory events in autism, as the absence of stabilized adaptation could induce the processing of each incoming stimulus as a new event, even if it is repeated multiple times.

The absence of stabilization of neural adaptation to sounds could also be interpreted as a lack of consolidation of the internal model, in the context of Bayesian theory (Lawson et al., 2014). The current results were supportive of the « weak prior hypothesis in autism » as difficulties in adaptation reflect a weaker influence of priors, leading to an overwhelming environment with continuously new stimulations (Sapey-Triomphe et al., 2021).

The lack of a stabilized internal model could also be discussed in relation to neural noise theory in autism. Indeed, an absence of consolidation of the internal model in the context of adaptation could be compatible with a large endogenous noise in autistic adults (Simmons et al., 2009).

#### 4.2. Reinforcement of effects by sound content

In non-autistic adults a difference in dynamics was observed according to the social nature of sounds with a faster adaptation to non-vocal than to vocal sounds, similarly to our previous study (Heurteloup et al., 2022) but using a single shorter stimulus sequence. Neural adaptation to vocal and non-vocal sounds in autistic adults was different from that of typical participants. Non-vocal sounds could be assimilated to complex environmental sounds. Previous studies reported an atypical processing of vocal sounds in autism (Abrams et al., 2019; Ceponiene et al., 2003; Gervais et al., 2004; Kuhl et al., 2005), but this difference was probably associated with difficulties in processing non-vocal comparison sounds (Bidet-Caulet et al., 2017).

The difference in sound processing according to their nature in both groups could be interpreted in regard to social motivation (Chevallier

et al., 2012), as non-autistic adults are more motivated by social stimuli than autistic individuals. This could explain the longer adaptation to vocal sounds in non-autistic adults, as these social sounds required more extensive processing compared to non-vocal sounds, which contained less information to encode and were therefore less demanding. Conversely, autistic adults would process non-vocal sounds more thoroughly. Additionally, this difference could be the consequence of the saliency of information to be encoded. Non-vocal sounds would be considered more salient in the sequence for autistic people, and vocal sounds for non-autistic people, resulting in slower adaptation processes.

An effect of emotion on neural adaptation has been observed in both autistic and non-autistic adults. The atypical regularity encoding of vocal emotional sounds is in line with difficulties observed in autism with emotional prosody processing (Peppé et al., 2007), and could reflect the non-automatization of emotional processing in autism (Siciliano and Clausi, 2020).

#### 4.3. Limitations

The multiplicity of the tested conditions (vocal, non-vocal, emotional, and neutral) induced the creation of a complex stimulation sequence. However, because the number of trials in the emotional and neutral conditions was not sufficient to conduct a statistically powerful analysis, only a qualitative description of the adaptation pattern was provided. Nonetheless, a longer task with more trials depending on the conditions would not be suitable for an autistic clinical population. Moreover, the main objective of the study was to investigate regularity encoding, and the sequence was created for that purpose but did not allow for the extraction of a robust MMN according to the number of previous trials. This limitation prevents us from answering the question of atypical change detection according to the number of repetitions of the standard stimuli in autism.

To further improve this study, several enhancements could be considered. First, adding behavioral scales that assess hypo- and/or hypersensitivity commonly observed in autism may be beneficial. Tools such as the Dunn Sensory Profile (Kern et al., 2017) or the Glasgow Sensory Questionnaire (Robertson & Simmons, 2013) could help evaluate a potential relationship between sensory profiles and neural adaptation in autism. Adding a measure of Restricted and Repetitive Behaviors (RRB; Bourreau et al., 2009) could provide insight into a potential relationship between RRB and neural adaptation, possibly explaining the need for routine and sameness. Additionally, although brain activity was mapped on the whole scalp, this study focused on a single EEG channel to shed light on auditory evoked potential modulation with neural adaptation. Performing sources localization analyses could help better describe the dynamics of this adaptation, as shown in Recasens et al. (2015) in their MEG study.

#### 4.4. Conclusion

To summarize, this study enabled us to measure neural adaptation to different natures of sounds in a clinical population with a short and simple paradigm, highlighting general difficulties in these processes in autistic adults. This study also revealed difficulties in processing emotions in autism, as well as environmental sounds. These highlighted atypicalities provide information on the two symptomatologic dimensions of autism: restricted and repetitive behaviors with atypical neural adaptation and social communication disorders revealed by difficulties in responding to emotional sounds. In the future, testing more than one emotion could inform our understanding of the processing of emotional prosody in autism.

#### Funding

This work was supported by the INSERM (National Institute for Health and Medical Research). Annabelle Merchie was funded by the

FEANS (Fondation Européenne pour l'Avancement des Neurosciences) and the ANR (ANR-19-CE37-0022-01). Carles Escera was supported by the Maria de Maeztu Center of Excellence CEX2021-001159-M (supported by MCIN/AEI); the 2021SGR-00356 Consolidated Research Group of the Catalan Government, and the ICREA Acadèmia Distinguished Professorship. The funding sources were not involved in the study design, data collection and analysis, or writing of the report.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

We thank all the volunteers for their time and participation in this study, and we thank Zoé Ranty, Lisa Michel and Raphaël Gautier for assistance in data collection.

#### Author Contributions

AM, CE and MG designed the study. AM performed data acquisition. EH-D was responsible for the recruitment and the clinical assessment of patients. AM and MG were responsible for data and statistical analyses. MG was responsible for fundraising. AM and MG wrote the first version of the manuscript. All authors were involved in preparing and reviewing the manuscript.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.clinph.2024.12.003>.

#### References

- Abrams, D.A., Padmanabhan, A., Chen, T., Odrozola, P., Baker, A.E., Kochalka, J., et al., 2019. Impaired voice processing in reward and salience circuits predicts social communication in children with autism. *Elife* 8, e39906. <https://doi.org/10.7554/eLife.39906>.
- Aguera, P., Jerbi, K., Caclin, A., Bertrand, O.E.L.A.N., 2011. A software package for analysis and visualization of MEG, EEG, and LFP signals. *Comput Intell Neuro*, 158970.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. DSM-5-TR. American Psychiatric Association Publishing; 2022. DOI: 10.1176/appi.books.9780890425787.
- Ashgari, S.Z., Farashi, S., Bashirian, S., Jenabi, E., 2021. Distinctive prosodic features of people with autism spectrum disorder: a systematic review and meta-analysis study. *Sci Rep* 11, 1–17. <https://doi.org/10.1038/s41598-021-02487-6>.
- Baldeweg, T., Williams, J.D., Gruzelier, J.H., 1999. Differential changes in frontal and sub-temporal components of mismatch negativity. *Int J Psychophysiol Off J Int Organ Psychophysiol* 33, 143–148. [https://doi.org/10.1016/s0167-8760\(99\)00026-4](https://doi.org/10.1016/s0167-8760(99)00026-4).
- Baldeweg, T., Klugman, A., Gruzelier, J., Hirsch, S.R., 2004. Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophr Res* 69, 203–217. <https://doi.org/10.1016/j.schres.2003.09.009>.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E., 2001. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord* 31, 5–17. <https://doi.org/10.1023/a:1005653411471>.
- Belin, P., Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex 2000;403:4.
- Belin, P., Fecteau, S., Bedard, C., 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn Sci* 8, 129–135. <https://doi.org/10.1016/j.tics.2004.01.008>.
- Bidet-Cautel, A., Latinus, M., Roux, S., Malvy, J., Bonnet-Brilhaut, F., Bruneau, N., 2017. Atypical sound discrimination in children with ASD as indicated by cortical ERPs. *J Neurodev Disord* 9, 13. <https://doi.org/10.1186/s11689-017-9194-9>.
- Boucher, J., Lewis, V., Collis, G.M., 2000. Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *J Child Psychol Psychiatry* 41, 847–857.
- Bourreau, Y., Roux, S., Gomot, M., Bonnet-Brilhaut, F., Barthélémy, C., 2009. Validation of the repetitive and restricted behaviour scale in autism spectrum disorders. *Eur. Child Adolesc. Psychiatry* 18 (11), 675–682. <https://doi.org/10.1007/s00787-009-0028-5>.
- Bruneau, N., Roux, S., Cléry, H., Rogier, O., Bidet-Cautel, A., Barthélémy, C., 2013. Early neurophysiological correlates of vocal versus non-vocal sound processing in adults. *Brain Res* 1528, 20–27. <https://doi.org/10.1016/j.brainres.2013.06.008>.
- Buhr, K., Dugas, M.J., 2009. The role of fear of anxiety and intolerance of uncertainty in worry: an experimental manipulation. *Behav Res Ther* 47, 215–223. <https://doi.org/10.1016/j.brat.2008.12.004>.
- Cary, E., Pacheco, D., Kaplan-Kahn, E., McKernan, E., Matsuba, E., Prieve, B., et al., 2023. Brain signatures of early and late neural measures of auditory habituation and discrimination in autism and their relationship to autistic traits and sensory overresponsivity. *J Autism Dev Disord*. <https://doi.org/10.1007/s10803-022-05866-8>.
- Ceponiene, R., Lepisto, T., Shestakova, A., Vanhala, R., Alku, P., Naatanen, R., et al., 2003. Speech-sound-selective auditory impairment in children with autism: They can perceive but do not attend. *Proc Natl Acad Sci* 100, 5567–5572. <https://doi.org/10.1073/pnas.0835631100>.
- Charpentier, J., Kovarski, K., Houy-Durand, E., Malvy, J., Saby, A., Bonnet-Brilhaut, F., et al., 2018. Emotional prosodic change detection in autism Spectrum disorder: an electrophysiological investigation in children and adults. *J Neurodev Disord* 10, 28. <https://doi.org/10.1186/s11689-018-9246-9>.
- Chen, T.-C., Hsieh, M.H., Lin, Y.-T., Chan, P.-Y.-S., Cheng, C.-H., 2020. Mismatch negativity to different deviant changes in autism spectrum disorders: a meta-analysis. *Clin Neurophysiol* 131, 766–777. <https://doi.org/10.1016/j.clinph.2019.10.031>.
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E.S., Schultz, R.T., 2012. The social motivation theory of autism. *Trends Cogn Sci* 16, 231.
- Costa-Faidella, J., Baldeweg, T., Grimm, S., Escera, C., 2011. Interactions between “what” and “when” in the auditory system: temporal predictability enhances repetition suppression. *J Neurosci* 31, 18590–18597. <https://doi.org/10.1523/JNEUROSCI.2599-11.2011>.
- Cowan, N., Winkler, I., Teder, W., Näätänen, R., 1993. Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *J Exp Psychol Learn Mem Cogn* 19, 909–921. <https://doi.org/10.1037/0278-7393.19.4.909>.
- Crespo, C., Santos, S., Canavarro, M.C., Kielipkowski, M., Pryor, J., Feres-Carneiro, T., 2013. Family routines and rituals in the context of chronic conditions: a review. *Int J Psychol* 48, 729–746. <https://doi.org/10.1080/00207594.2013.806811>.
- Fan, Y.-T., Cheng, Y., 2014. Atypical mismatch negativity in response to emotional voices in people with autism spectrum conditions. *PLoS One* 9, e102471. <https://doi.org/10.1371/journal.pone.0102471>.
- Font-Alaminos, M., Cornella, M., Costa-Faidella, J., Hervás, A., Leung, S., Rueda, I., et al., 2020. Increased subcortical neural responses to repeating auditory stimulation in children with autism spectrum disorder. *Biol Psychol* 149, 107807. <https://doi.org/10.1016/j.biopsycho.2019.107807>.
- Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, L., 2004. Abnormal cortical voice processing in autism. *Nat Neurosci* 7, 801–802.
- Gomot, M., Giard, M.-H., Adrien, J.-L., Barthelemy, C., Bruneau, N., 2002. Hypersensitivity to acoustic change in children with autism: electrophysiological evidence of left frontal cortex dysfunctioning. *Psychophysiology* 39, 577–584.
- Gomot, M., Blanc, R., Cléry, H., Roux, S., Barthelemy, C., Bruneau, N., 2011. Candidate electrophysiological endophenotypes of hyper-reactivity to change in autism. *J Autism Dev Disord* 41, 1078–1087.
- Gomot, M., Wicker, B., 2012. A challenging, unpredictable world for people with autism spectrum disorder. *Int J Psychophysiol J Int Organ Psychophysiol* 83, 240–247.
- Gonzalez-Gadea, M.L., Chennu, S., Bekinschtein, T.A., Rattazzi, A., Beraudi, A., Tripicchio, P., et al., 2015. Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. *J Neurophysiol* 114, 2625–2636. <https://doi.org/10.1152/jn.00543.2015>.
- Guiraud, J.A., Kushnerenko, E., Tomalski, P., Davies, K., Ribeiro, H., Johnson, M.H., 2011. Differential habituation to repeated sounds in infants at high risk for autism. *Neuroreport* 22, 845–849. <https://doi.org/10.1097/WNR.0b013e32834e0bec>.
- Haenschel, C., Vernon, D.J., Dwivedi, P., Gruzelier, J.H., Baldeweg, T., 2005. Event-related brain potential correlates of human auditory sensory memory-trace formation. *J Neurosci Off J Soc Neurosci* 25, 10494–10501. <https://doi.org/10.1523/JNEUROSCI.1227-05.2005>.
- Heurteloup, C., Merchie, A., Roux, S., Bonnet-Brilhaut, F., Escera, C., Gomot, M., 2022. Neural repetition suppression to vocal and non-vocal sounds. *Cortex* 148, 1–13. <https://doi.org/10.1016/j.cortex.2021.11.020>.
- Jacobsen, T., Schröger, E., Alter, K., 2004. Pre-attentive perception of vowel phonemes from variable speech stimuli. *Psychophysiology* 41, 654–659. <https://doi.org/10.1111/1469-8986.2004.00175.x>.
- Jamal, W., Cardinaux, A., Haskins, A.J., Kjelgaard, M., Sinha, P., 2020. Reduced sensory habituation in autism and its correlation with behavioral measures. *J Autism Dev Disord*. <https://doi.org/10.1007/s10803-020-04780-1>.
- Jordi, C.-F., Grimm, S., Slabu, L., Díaz-Santaella, F., Escera, C., 2011. Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials: adaptation in the human auditory system. *Psychophysiology* 48, 774–783. <https://doi.org/10.1111/j.1469-8986.2010.01144.x>.
- Kassambara A. ggpubr: “ggplot2” Based Publication Ready Plots. 2020.
- Kern, J.K., Garver, C.R., Carmody, T., Andrews, A.A., Trivedi, M.H., Mehta, J.A., 2007. Examining sensory quadrants in autism. *Res. Autism Spectr. Disord.* 1 (2), 185–193. <https://doi.org/10.1016/j.rasd.2006.09.002>.
- Kolesnik, A., Begum Ali, J., Gliga, T., Guiraud, J., Charman, T., Johnson, M.H., et al., 2019. Increased cortical reactivity to repeated tones at 8 months in infants with later ASD. *Trans Psychol* 9, 1–11. <https://doi.org/10.1038/s41398-019-0393-x>.
- Korpiolhti, P., Jansson-Verkasalo, E., Mattila, M.-L., Kuusikko, S., Suominen, K., Rytty, S., et al., 2006. Processing of affective speech prosody is impaired in asperger

- syndrome. *J Autism Dev Disord* 37, 1539–1549. <https://doi.org/10.1007/s10803-006-0271-2>.
- Kuhl, P.K., Coffey-Corina, S., Padden, D., Dawson, G., 2005. Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev Sci* 8, F1–F12.
- Kujala, T., Aho, E., Lepistö, T., Jansson-Verkasalo, E., Nieminen-von Wendt, T., von Wendt, L., et al., 2007. Atypical pattern of discriminating sound features in adults with Asperger syndrome as reflected by the mismatch negativity. *Biol Psychol* 75, 109–114. <https://doi.org/10.1016/j.biopsycho.2006.12.007>.
- Kujala, T., Kuuluvainen, S., Saalasti, S., Jansson-Verkasalo, E., von Wendt, L., Lepistö, T., 2010. Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. *Clin Neurophysiol* 121, 1410–1419. <https://doi.org/10.1016/j.clinph.2010.03.017>.
- Latinus, M., Mofid, Y., Kovarski, K., Charpentier, J., Batty, M., Bonnet-Brilhaut, F., 2019. Atypical sound perception in ASD explained by inter-trial (in)consistency in EEG. *Front Psychol* 10.
- Lawrence MA. ez: Easy Analysis and Visualization of Factorial Experiments. 2016.
- Lawson, R.P., Rees, G., Friston, K.J., 2014. An aberrant precision account of autism. *Front Hum Neurosci* 8. <https://doi.org/10.3389/fnhum.2014.00302>.
- Leipold, S., Abrams, D.A., Karraker, S., Phillips, J.M., Menon, V., 2022. Aberrant emotional prosody circuitry predicts social communication impairments in children with autism. *Biol Psychiatry Cogn Neurosci Neuroimaging*. <https://doi.org/10.1016/j.bpsc.2022.09.016>.
- Lord, C., Rutter, M., Le Couteur, A., 1994. Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord* 24, 659–685. <https://doi.org/10.1007/BF02172145>.
- Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., DiLavore, P.C., et al., 2000. The autism diagnostic observation schedule—generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord* 30, 205–223. <https://doi.org/10.1023/A:1005592401947>.
- Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. *Trends Cogn Sci* 8, 204–210. <https://doi.org/10.1016/j.tics.2004.03.008>.
- Martineau, J., Roux, S., Garreau, B., Adrien, J.L., Lelord, G., 1992. Unimodal and crossmodal reactivity in autism: presence of auditory evoked responses and effect of the repetition of auditory stimuli. *Biol Psychiatry* 31, 1190–1203. [https://doi.org/10.1016/0006-3223\(92\)90338-Z](https://doi.org/10.1016/0006-3223(92)90338-Z).
- McCann, J., Peppé, S., 2003. Prosody in autism spectrum disorders: a critical review. *Int J Lang Commun Disord* 38, 325–350. <https://doi.org/10.1080/1368282031000154204>.
- Merchie, A., Gomot, M., 2023. Habituation, adaptation and prediction processes in neurodevelopmental disorders: a comprehensive review. *Brain Sci* 13, 1110. <https://doi.org/10.3390/brainsci13071110>.
- Michel, L., Ricou, C., Bonnet-Brilhaut, F., Houy-Durand, E., Latinus, M., 2023. Sounds pleasantness ratings in autism: interaction between social information and acoustical noise level. *J Autism Dev Disord*. <https://doi.org/10.1007/s10803-023-05989-6>.
- Oostenveld, R., Praamstra, P., 2001. The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112 (4), 713–719. [https://doi.org/10.1016/s1388-2457\(00\)00527-7](https://doi.org/10.1016/s1388-2457(00)00527-7).
- Palmer, C.J., Lawson, R.P., Hohwy, J., 2017. Bayesian approaches to autism: towards volatility, action, and behavior. *Psychol Bull* 143, 521–542. <https://doi.org/10.1037/bul0000097>.
- Pellicano, E., Fatima, U., Hall, G., Heyworth, M., Lawson, W., Lilley, R., et al., 2022. A capabilities approach to understanding and supporting autistic adulthood. *Nat Rev Psychol* 1–16. <https://doi.org/10.1038/s44159-022-00099-z>.
- Peppé, S., McCann, J., Gibbon, F., O'Hare, A., Rutherford, M., 2007. Receptive and expressive prosodic ability in children with high-functioning autism. *J Speech Lang Hear Res* 50, 1015–1028. [https://doi.org/10.1044/1092-4388\(2007\)071](https://doi.org/10.1044/1092-4388(2007)071).
- Pinheiro, A.P., Barros, C., Vasconcelos, M., Obermeier, C., Kotz, S.A., 2017. Is laughter a better vocal change detector than a growl? *Cortex* 92, 233–248. <https://doi.org/10.1016/j.cortex.2017.03.018>.
- Recasens, M., Leung, S., Grimm, S., Nowak, R., Escera, C., 2015. Repetition suppression and repetition enhancement underlie auditory memory-trace formation in the human brain: an MEG study. *Neuroimage* 108, 75–86. <https://doi.org/10.1016/j.neuroimage.2014.12.031>.
- Robertson, A.E., Simmons, D.R., 2013. The relationship between sensory sensitivity and autistic traits in the general population. *J Autism Dev. Disord.* 43 (4), 775–784. <https://doi.org/10.1007/s10803-012-1608-7>.
- Ruiz-Martínez, F.J., Rodríguez-Martínez, E.I., Wilson, C.E., Yau, S., Saldaña, D., Gómez, C.M., 2020. Impaired P1 habituation and mismatch negativity in children with autism spectrum disorder. *J Autism Dev Disord* 50, 603–616. <https://doi.org/10.1007/s10803-019-04299-0>.
- Sapey-Triomphe, L.-A., Temmerman, J., Puts, N.A.J., Wagemans, J., 2021. Prediction learning in adults with autism and its molecular correlates. *Mol Autism* 12, 64. <https://doi.org/10.1186/s13229-021-00470-6>.
- Siciliano, L., Clausi, S., 2020. Implicit vs. Explicit emotion processing in autism spectrum disorders: an opinion on the role of the cerebellum. *Front Psychol* 11.
- Simmons, D.R., Robertson, A.E., McKay, L.S., Toal, E., McAleer, P., Pollick, F.E., 2009. Vision in autism spectrum disorders. *Vision Res* 49, 2705–2739. <https://doi.org/10.1016/j.visres.2009.08.005>.
- Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2021.
- Team Rs. RStudio: Integrated Development for R. RStudio PBC, Boston, MA: 2020.
- Wechsler D. Wechsler Adult Intelligence Scale—Fourth Edition 2012. DOI: 10.1037/t15169-000.
- Wickham, H., 2016. *ggplot2: Elegant graphics for data analysis*. Springer-Verlag, New York.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al., 2019. Welcome to the tidyverse. *J Open Source Softw* 4, 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, H., François, R., Henry, L., Müller, K.dplyr., 2021. *A Grammar of Data Manipulation*.



### 1.4. Synthèse

L'adaptation neurale est atypique dans l'autisme en réponse à la répétition de sons vocaux et non-vocaux. Cependant, cette différence est plus importante pour les sons complexes non-vocaux, ce qui soulève la question de l'adaptation aux sons environnementaux chez les personnes autistes, laquelle semble difficile et pourrait expliquer une hypersensibilité à l'environnement. Un effet du contenu émotionnel a aussi été mis en évidence comme retardant la mise en place de l'adaptation. Dans le cadre de la boucle *Perception – Représentation – Action* cette différence de traitement de la prosodie émotionnelle dès cette première étape pourrait expliquer les atypies dans les étapes qui suivent. En effet, si dès l'étape de l'adaptation neurale aux émotions, le processus est atypique, alors les représentations perceptives des émotions vocales seront plus instables, ce qui entraînera des réponses atypiques à la prosodie émotionnelle.

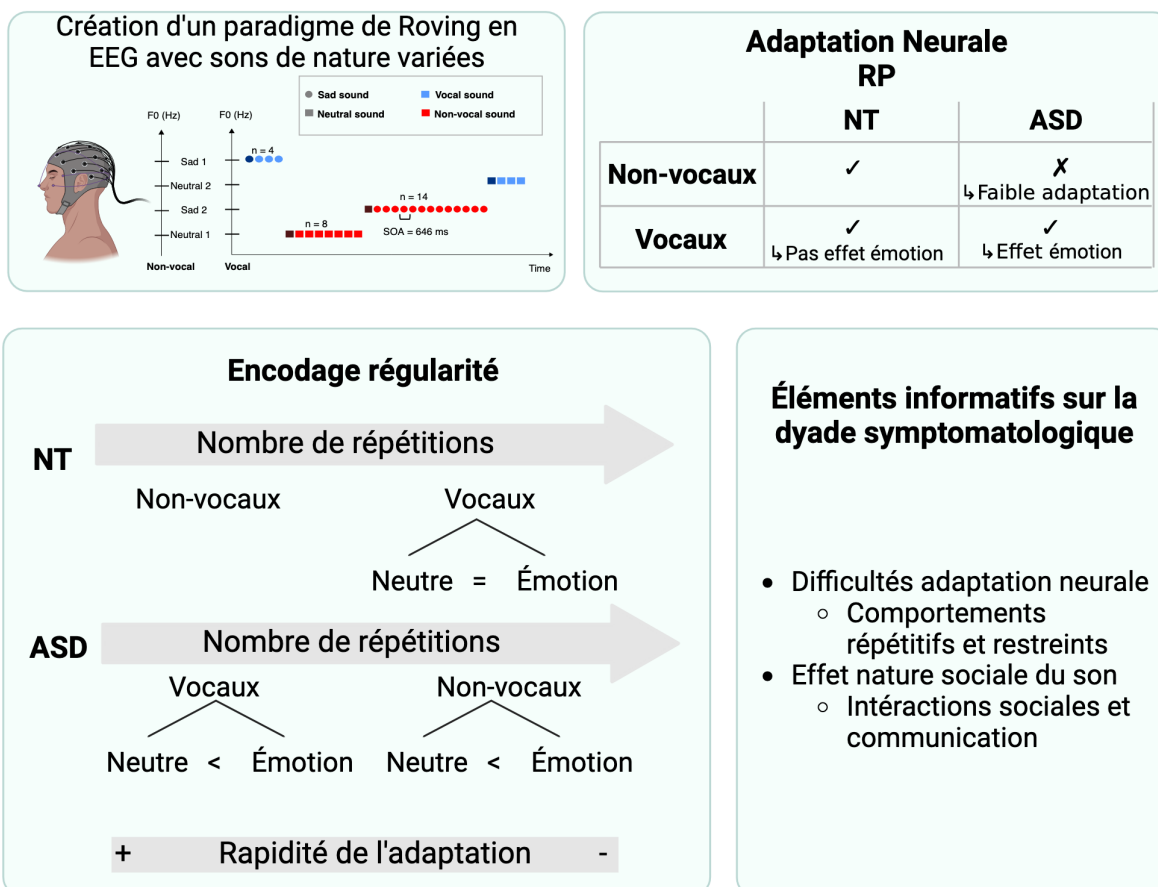


Figure 19 : Schéma de synthèse des études n°1 et n°2 à propos de l'adaptation neurale à la prosodie émotionnelle



## **2. Représentation mentale du sourire vocal**

### **2.1. Avant-propos**

La première partie expérimentale de cette thèse a montré des atypies d'adaptation neurale dans l'autisme plus marquées pour des sons présentant un contenu émotionnel. Il est donc essentiel de se poser la question de l'impact de ces difficultés au niveau de cette première étape de traitement en cherchant à mettre en évidence les caractéristiques des représentations perceptives des émotions vocales. En effet, comme il a été montré dans l'introduction des difficultés de reconnaissance des émotions ont été mises en évidence dans l'autisme. Si les modèles internes de ces émotions ne se basent pas sur les mêmes indices acoustiques alors cela expliquerait en partie ces différences. Une étude en corrélation inverse a été menée dans un groupe d'adultes neurotypiques et un groupe d'adultes autistes afin de déterminer si les représentations perceptives sont différentes selon les groupes, mais aussi, si d'éventuelles différences dans la variabilité des réponses existent.

## 2.2. Étude n°3 : Représentation perceptive du sourire vocal dans l'autisme

### **Intact Representation of Vocal Smile in Autism: A reverse correlation approach**

*Merchie, A., Ranty, Z., Adl Zarrabi, A., Bonnet-Brilhault, F., Houy-Durand, E., Aucouturier, J-J. & Gomot, M.*

*Article en révision dans Research in Autism and Developmental Disorders*

**Introduction et objectifs :** Dans l'autisme, une absence de réponse au sourire social a déjà longuement été décrite, tant à travers les observations cliniques que les études expérimentales. Cette absence de réponse au sourire est d'ailleurs un des premiers signes d'alerte chez le jeune enfant. Cependant, aucune étude n'a cherché à identifier les représentations perceptives associées au sourire dans l'autisme que ce soit dans la modalité visuelle ou auditive. Le sourire vocal a une signature acoustique robuste et partagée entre les adultes neurotypiques (Ponsot et al., 2018). L'objectif de cette étude était d'estimer la ou les représentations perceptives du sourire vocal dans un groupe d'adultes autistes.

**Matériel et méthodes :** 30 adultes autistes et 29 adultes neurotypiques ont réalisé une tâche de Corrélation Inverse (RevCor) dans laquelle ils devaient choisir le son le plus souriant parmi une paire de sons. À partir de ces choix, des caractéristiques acoustiques (modulations d'amplitude et de fréquences) ont pu être extraites afin d'obtenir un modèle du sourire vocal pour chaque participant et pour chacun des groupes. La comparaison au modèle de l'étude originale (Ponsot et al., 2018) a permis d'établir la « typicalité » du modèle. L'ajout d'un paradigme de double-passe à la fin de la séquence de RevCor a permis l'estimation de la variabilité intra-individuelle des réponses (bruit interne). Des corrélations entre les scores cliniques (ADOS, AQ et EQ), la typicalité et le bruit interne ont été effectuées.

**Résultats :** La majorité des adultes autistes (17) a basé sa représentation du sourire vocal sur des caractéristiques acoustiques similaires à celles des participants neurotypiques, et aucune différence dans le niveau de bruit interne n'a été observée entre les groupes. Cependant, les comparaisons entre les groupes ont révélé que la représentation était en moyenne plus typique chez les neurotypiques que chez les participants autistes. Une analyse en cluster a montré que la différence de typicalité est expliquée par un petit sous-ensemble de participants autistes présentant des représentations éloignées de celle des neurotypiques. Cependant, deux autres profils se distinguent : une partie des adultes autistes a un modèle du sourire vocal différent et une autre partie présente un bruit interne élevé. Enfin, dans le continuum des adultes neurotypiques et autistes, une corrélation positive significative entre le score à l'EQ et la typicalité a été mise en évidence.

**Discussion :** La représentation perceptive des émotions vocales semble préservée dans l'autisme pour une grande majorité des adultes, mêmes si une partie du groupe s'est révélée différente. Le lien entre les capacités empathiques et la typicalité du modèle montre que les représentations mentales des émotions se façonnent, entre autres, conjointement à l'empathie. Les difficultés de reconnaissance des émotions vocales dans l'autisme mises en évidence dans de nombreuses études, pourraient donc être liées aux profils avec des représentations différentes qui ont aussi par conséquent un traitement cérébral atypique. L'étude des profils est donc importante pour caractériser ces difficultés de réponse et de reconnaissance des émotions.

**Intact Representation of Vocal Smile in Autism: A reverse correlation approach**

**Authors information**

Annabelle Merchie<sup>a,b</sup>, Zoé Ranty<sup>a,b</sup>, Aynaz Adl Zarrabic, Frédérique Bonnet-Brilhault<sup>a,b,d</sup>,  
Emmanuelle Houy-Durand<sup>a,b,d</sup>, Jean-Julien Aucouturier<sup>c,e</sup> & Marie Gomot<sup>a,b</sup>

a: Université de Tours, INSERM, Imaging Brain & Neuropsychiatry iBraIN U1253, 37032,  
Tours, France

b: Excellence Center for Autism and Neurodevelopmental Disorders, EXAC·T, CHRU de  
Tours, Tours, France

c: Université de Franche-Comté, SUPMICROTECH, CNRS, institut FEMTO-ST, Besançon,  
France

d: Centre Ressource Autisme région Centre-Val de Loire, CHRU de Tours, Tours, France

e: STMS Lab IRCAM, CNRS, Sorbonne Université, Paris, France

**Corresponding author**

Marie Gomot ([gomot@univ-tours.fr](mailto:gomot@univ-tours.fr)), ORCID: 0000-0003-1747-5806

### **Abstract**

Atypical emotional prosody production and perception have been reported in autism. However, it is unclear whether these particularities are associated with unusual mental representations of vocal emotions. The objective of the current study was to explore the mental representation of vocal smile in autistic adults. Twenty-nine autistic (ASD) and 29 neurotypical (NT) adults performed an auditory reverse correlation task, that affords the opportunity to extract acoustic features of mental representation and their variability. Most ASD participants (17) based their representation of vocal smile on similar acoustic features as NT participants and no difference in the level of internal noise was observed. However, comparisons between groups revealed a more typical representation in NT than in ASD. Subsequent cluster analysis revealed that the difference of typicality was explained by a small subset of ASD participants displaying different representations. A correlation analysis also revealed that the typicality was positively correlated with the empathetic level within both groups. While most autistic adults have a preserved mental representation of vocal smiles, a subset shows less robust and typical representations, which is linked to lower levels of empathy. This study highlights that the perception of vocal smiles in autism is more nuanced than previously reported, with empathy playing a substantial role in shaping these mental representations.

### **Highlights**

- Most autistic adults share similar vocal smile representations with neurotypical adults.
- A subset of autistic individuals shows atypical vocal smile representations linked to empathy.
- Vocal smile perception in autism is nuanced, with empathy influencing mental representations.

### **Keywords**

Autism, Prosody, Vocal smile, Reverse Correlation, Internal noise

### 1. Introduction

Prosodic atypicalities are considered as a core central feature of communication profile in autism. In particular, atypicalities in emotional prosody (i.e., vocal emotion) perception and production in children (McCann & Peppé, 2003; Peppé et al., 2007) but also in adults (Hubbard et al., 2017) have been widely studied since the original descriptions of Kanner and Sukhareva (Andronikof & Fontan, 2016; Kanner, 1943). An early hallmark in autism is also a lack of response to (Beall et al., 2008) and production of (Dawson et al., 1990) social smile, one of the most emblematic expression of the human emotional repertoire. Understanding and evaluating mental representation of vocal smile in autism should give keys in the comprehension of prosody atypicalities in autism.

Behavioral studies in autism have shown difficulties in recognition of others prosodic intention (see Leung et al., 2022 for a review). However, pitch discrimination in autism has been described as better than in neurotypical individuals for non-vocal sounds (Bonnell et al., 2003, 2010; Heaton et al., 2008) with also a larger proportion of absolute pitch in the autistic population (Miller, 1999; Rimland & Fein, 1988). Nevertheless, these improved auditory skills in autism are not sufficient for efficient recognition of vocal emotions, suggesting that atypical prosodic processing does not arise from sensory processing difficulties, but rather from subsequent integrative processes enabling the attribution of meaning to specific and voice patterns.

Perceptive representation of several stimulus categories have been extracted from physical features through the reverse correlation technique, historically in the visual modality (Ahumada & Lovell, 1971), but has recently gained in influence in the auditory modality (Adl Zarrabi et al., 2024; Goupil et al., 2021; Merchie et al., 2024; Ponsot, Arias, et al., 2018; Ponsot, Burred, et al., 2018; L. Wang et al., 2022). Smile could be identified from vocal cues (Aubergé & Cathiard, 2003; Tartter, 1980; Tartter & Braun, 1994) and thanks to reverse correlation, a robust acoustic model of vocal smile was revealed, which even led to motor resonance in neurotypical

adults (Arias et al., 2018; Merchie et al., 2024). Vocal smile internal representation is characterized by an upward shift in frequency of formants F1 and F2, and an increase in energy of F2, F3 and F4 compared to neutral voice (Ponsot, Arias, et al., 2018), reflecting the modulation of vocal tract induced by the bilateral contraction of the Zygomaticus major muscles during smile (Wood et al., 2016). Few studies, however, have used this technique to characterize mental representations in autism, neither in the visual nor in the auditory modality. In a visual reverse correlation “Bubbles” task, results suggested that autistic children did not use the same information than neurotypical during identity judgment (Ewing et al., 2018). However, another study demonstrated that to identify positive emotion, contrary to eye tracking studies (Black et al., 2017), ASD children correctly used the information from the eyes region (Song et al., 2012). Only one study used an auditory reverse correlation paradigm in autistic children, and observed similar representations of rising pitch in speech and complex tone, and musical melody contours (L. Wang et al., 2022). No study gets yet interested in the evaluation of acoustic cues used in the classification and representation of vocal emotion in autism.

Both neural and behavioral intra-subject variability have been proposed to contribute to sensory symptoms in ASD (Magnuson et al., 2020; Ward, 2019). Whether this variability also impacts internal representations remains to be explored. Mental representation of specific features could be altered by internal noise leading to difficulty in processing signal. Noise is always present in processing signal through external (Brinkman et al., 2017) or internal noise (Neri, 2010). In reverse correlation studies, the internal noise referred partly to “intrinsic noise in the decision making” and could be estimated in reverse correlation task through the double pass-paradigm (i.e., repetition of 2 blocks of identical stimuli to estimate response variability) (Ponsot, Arias, et al., 2018). Few studies used the double-pass methodology to estimate internal noise in autism, but they reported an increased level during a visual task (Park et al., 2017; Vilidaite et al., 2017) or no difference with an audio reverse correlation protocol compared to neurotypicals (L. Wang

et al., 2022).

The main objective of the current study was to characterize the perceptive representation of vocal smile in a group of autistic adults with an adapted reverse correlation paradigm and to compare it to the robust neurotypical model (Merchie et al., 2024; Ponsot, Arias, et al., 2018), through the calculation of the representation typicality (i.e., standardized distance between representations) taking as a reference Ponsot et al.'s model (2018). This will enable us to determine whether autistic people rely on the same acoustic cues as those of neurotypicals to classify a voice as smiling. This paradigm also allowed to estimate internal noise in the same groups thanks to double-pass methodology. Combination of mental representation typicality and internal noise indices should give the opportunity to extract different profiles in autistic adults and to link these profiles to clinical information. This study with these different measures thus informs on both the typicality and the robustness of the representations of emotional prosody in autistic adults.

## **2. Material and methods**

### **2.1. Population**

Thirty autistic and 29 non-autistic participants participated in the current study. Autistic adults were recruited through the Autism Resource Center of Centre Val de Loire and diagnosed by a trained clinical team according to DSM-5 criteria. Diagnosis was confirmed with ADOS-2 and ADI-R (Lord et al., 1994, 2000). Verbal and non-verbal efficiencies were estimated through the use of selected subtests of the WAIS-IV (Wechsler, 2012) for neurotypicals, while the entire scale was proposed to autistic participants. The Autism Quotient 50-questions (AQ) and the Empathy Quotient 40-questions (EQ) were proposed to all the participants (Baron-Cohen et al., 2001; Baron-Cohen & Wheelwright, 2004). One participant was discarded in the autistic group because of a doubt about diagnosis and another participant who was not able to complete the task. The group of non-autistic adults reported an absence of any developmental difficulties in language and sensorimotor acquisition. In both groups, no disease of central nervous system,

infectious or metabolic diseases, epilepsy, or an abnormal audition was reported. Participant's audition was checked for both ear with a short subjective audiometry test. Every participant (or their legal guardian in case of adult under guardianship) signed an informed consent form, and the protocol received approval from Ethic Committee (PROSCEA 2017/23; ID RCB: 2017-A00756-47). A description of the two groups is presented in table 1.

	NT	ASD	Comparison
<b>Sex</b>	12 ♀ 17♂	7 ♀21♂	$\chi^2 = 1.06$ ; $p = .30$
<b>Age</b>	30.1 ± 12	35.3 ± 13	$t(54) = -1.55$ ; $p = .13$
<b>vIQ</b>	124.2 ± 11 (24)	118.3 ± 20 (25)	$t(40) = 1.30$ ; $p = .20$
<b>nvIQ</b>	106.3 ± 12 (24)	100.7 ± 18 (25)	$t(44) = 1.32$ ; $p = .19$
<b>AQ</b>	16.8 ± 7 (23)	34.9 ± 9 (25)	$t(47) = -8.30$ ; $p < .0001$
<b>EQ</b>	43.2 ± 9 (23)	23.7 ± 13 (25)	$t(45) = 6.31$ ; $p < .0001$
<b>ADOS</b>	/	8.8 ± 5.5	/

*Table 1. Characteristics of Neurotypical (NT) and Autistic (ASD) group (mean ± standard deviation). vIQ : verbal Intelligence Quotient calculated from the WAIS-IV in ASD, and composite score calculated from vocabulary and similarities subtests in NT. nvIQ : non-verbal Intelligence Quotient calculated from the WAIS-IV in ASD, and composite score calculated from matrix reasoning and cubes subtests in NT. ADOS: score calculated as the sum of the subdimensions communication and social interaction behaviors. The values in brackets correspond to the number of values included in the calculation of the mean and standard deviation due to missing data.*

### 2.2.Sound modulations

The same method and stimuli as in Ponsot et al. (2018) have been used to evaluate mental representation of vocal smile in both group. In brief, the French phoneme /a/ uttered by a male speaker with a constant pitch was recorded. To obtain a constant spectral energy a 500ms stationary part of the sound was selected. Then, the sound was modulated, to create multiple variants, in 25 frequencies between 100 and 10,000Hz with random gain values (in dB) from a Gaussian distribution with the Cleese python toolbox (Burred et al., 2019).

### 2.3.Procedure

The current experiment took place in a larger session (~1h30) of testing for most of the

participants. The sequence of stimulation for the reverse correlation task was composed by 200 trials for autistic (~15min) and 300 trials for non-autistic (~20min) divided in blocks of 50 trials. Each trial consisted in a pair of randomly-filtered voice to be compared to estimate the most smiling voice between both presented. Participants had to choose the most “smiling” sound of each presented pair by pressing a key. The last two blocks were identical to evaluate internal noise through the double-pass methodology. Consequently, the number of trials for estimating mental representation of vocal smile was 150 for autistic and 250 trials for non-autistic.

To estimate the minimum number of trials required to obtain a reliable representation of the vocal smile, a convergence analysis was carried out, revealing that 150 trials appeared sufficient for this estimation. However, in order to obtain a robust control representation from the model of the non-autistic participants, we decided to lengthen the sequence by adding 2 blocks of 50 trials each.

### **2.4.Reverse correlation analysis**

#### **2.4.1. Mental representation**

The mental representation of vocal smile was modelled for each participant with the classification image technique. The mean pitch contour of voice classified as “unsmiling” was subtracted from the mean pitch contour of “smiling” voice. Then, the resulting representation was normalized by dividing it by the sum of their absolute value. A model of each participant’s model was then computed and averaged by group. A comparison between NT and ASD’s gain was performed in each modulated frequency.

To estimate the deviation of participant model to the model of an external control group (Controls) (Ponsot, Arias, et al., 2018) the distance between the average group representation and the participant was computed and then normalized in both groups (‘representation typicality’) (see Equation 1). Representations for non-autistic were computed using the same procedure, using only the first 150 trials of each session in order to match the number of trials seen by patients.

$$(1) \text{ distance} = \sum_{f=100}^{10,000} |Gain_{f,controls} - Gain_{f,participant}|$$

$$(2) \text{ typicality} = 1 - \frac{\text{distance} - \min(\text{distance})}{\max(\text{distance}) - \min(\text{distance})}$$

*Equation 1: (1) Calculation of the distance between the Controls model and one participant's model as sum of the absolute difference of model's gains for each of the 25 modulated frequencies (f) between 100 and 10,000 Hz. (2) Calculation of the typicality as the normalization of the distance between 0 and 1 between NT and ASD.*

#### 2.4.2. Internal noise

The same method as in Adl Zarrabi et al. (2024) has been used to evaluate internal noise (expressed in units of the standard deviation of stimulus noise) from response consistency (probability of agreement, p-agree) and response bias (p-int1) between the two repeated blocks using the simulation procedure of Neri (2010). Then for each participant a reverse model was applied to obtain the value of internal noise minimizing the error between the observed and the simulated values for a participant's consistency and bias (Internal Noise). As in Neri (2010) internal noise was estimated in the [0 ; 5 std] range to avoid unreliable large value.

Internal noise was computed using the open-access python toolbox "palin" (<https://github.com/neuro-team-femto/palin>). Data from one autistic participant were removed for internal noise estimation because the repeated block was not completed.

#### 2.5. Statistical analysis

Statistical analyses were conducted on RStudio 4.0.4 (R. C. Team, 2021; Rs. Team, 2020) with the packages ggplot2 (Wickham, 2016), ggpubr (Kassambara, 2020), dplyr (Wickham et al., 2021), tidyverse (Wickham et al., 2019) and factoextra (Kassambara & Mundt, 2017).

In order to compare mental representation and noise between groups, Student's tests were performed for representation typicality, p-agree, p-int1 and internal noise. Pearson's correlation between scores in questionnaires (AQ and EQ) or clinical scores (ADOS score, the sum of communication and social interaction) and the different indices of the reverse correlation results were also performed. A global correlation analysis was performed when scores were available for both groups to reflect the effect of the autistic traits and empathy in a continuum, after

Fisher's z-tests have been performed to confirm that there was no statistical difference between groups in correlation coefficients. As multiple correlations were calculated the p-value was Bonferroni's corrected and the corrected values were reported.

### 2.5.1. Cluster analysis

To evaluate the potential different profiles in participants a cluster analysis based on both mental representation (representation typicality) and noise (internal noise) was performed. Variables were scaled for normalization for both groups. A K-means clustering method was chosen (see Manenti et al., 2024 for a detailed description of the methods to determine the best number of clusters). A qualitative description of the different clusters obtained was presented.

## 3. Results

### 3.1. Reverse correlation analysis

#### 3.1.1. Mental representation of vocal smile

A Student's test has been performed on each frequency between groups (ASD vs NT) to evaluate potential differences in mental representation of vocal smile in acoustic description (Figure 1). As the number of comparisons was quite large (25 frequencies) a Bonferroni correction was applied. No difference between the mental representation in ASD and NT has been revealed.

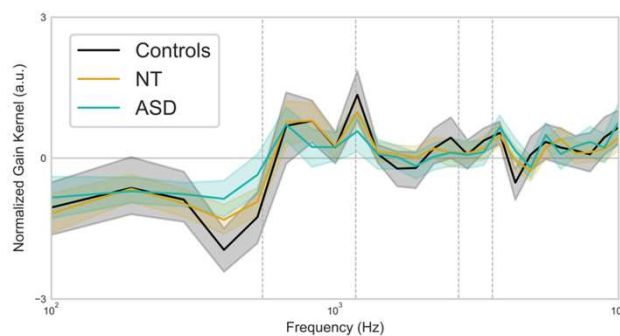


Figure 1. Mental representation of vocal smile in ASD (blue) and NT (orange). The "Controls" group (black) corresponds to data from Ponsot et al. (2018) with 600 trials. Dashed lines represent formants frequencies.

In both group representation of vocal smile was characterized by an upward switch in frequency of F1 and F2, and by an energy increase of F2 and F4.

### 3.1.2. Representation typicality

Representation typicality was calculated according to the distance to the model of the Controls group with 600 trials (Ponsot, Arias, et al., 2018). The typicality of the model appears significantly different between NT and ASD group ( $t(52) = 2.20$  ;  $p < .05$ , see Figure 2a), suggesting that mental representation of the NT was closer to these of the external Control's than the one of the ASD. Nevertheless, participants with autism used relevant frequency modulations to classify a vocal smile, and the typicality differences observed is mainly related to irrelevant frequency modulations ( $< F1$  frequency, 555 Hz).

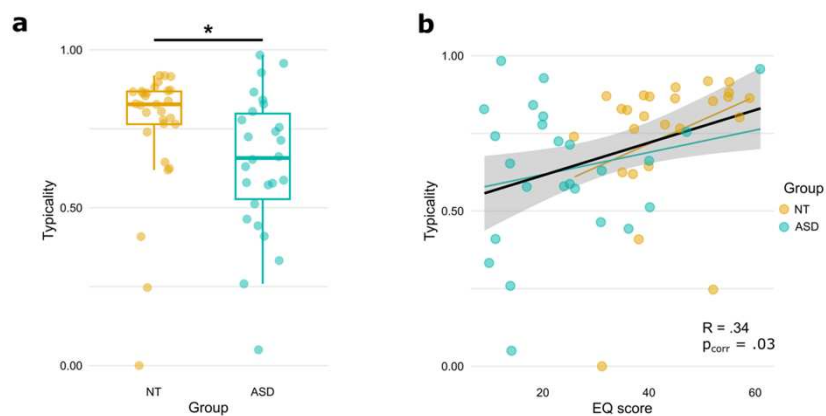


Figure 2. Typicality of the mental representation of vocal smile. a: typicality of the vocal smile representation to Controls in NT (orange) and ASD (blue) group \*  $p > .05$ ; b: correlations between typicality and EQ in ASD and NT groups, the fine colored lines correspond to separated group analysis, the bold black line is the correlation within groups.

To evaluate if a relation between clinical scores and the representation typicality existed in the ASD group, a correlation analysis was performed with the ADOS score but did not reveal any significant result ( $r = -.03$ ,  $p = .91$ ). Neither was the correlation coefficient between typicality

and AQ in NT ( $r = .03$  ( $p = .87$ )), nor in ASD ( $r = -.09$  ( $p = .66$ )) significant. Fisher's z-tests conducted to compare the correlation coefficients between two independent samples did not reveal any statistical difference ( $z = .42$ ,  $p = .68$ ).

The correlation between representation typicality and EQ in NT and in ASD were respectively  $r = .31$  ( $p = .13$ ) and  $r = .21$  ( $p = .32$ ). The Fisher's test revealed no statistical difference between the two correlation coefficients ( $z = .38$ ,  $p = .71$ ). Considering that there was no significant difference in correlation coefficients between groups, a joint analysis of the correlation between AQ or EQ scores and typicality was conducted on the whole population. There was no correlation with AQ score ( $r = -.22$ ,  $p_{\text{corr}} = .26$ ) but a significant correlation between the EQ score and the typicality was identified ( $r = .34$ ,  $p_{\text{corr}} = .03$ , see Figure 2b). The higher the EQ score, the more similar the acoustic smile representation is to that of the Controls group.

### 3.1.3. Internal noise

To evaluate differences between NT and ASD internal noise, Student's tests were performed to compare p-agree, p-int1 and the estimated internal noise. These comparisons revealed no difference for p-agree ( $t(53) = .66$ ;  $p > .1$ ), internal noise ( $t(37) = -1.29$ ;  $p > .1$ , see Figure 3a) and p-int1 was revealed ( $t(54) = 1.79$ ;  $p < .1$ ).

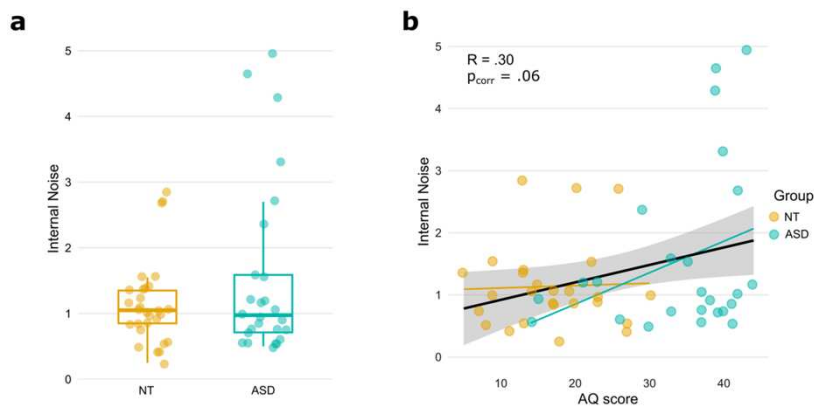


Figure 3. Internal noise estimation. a: internal noise in NT (orange) and ASD (blue) group \*  $p > .05$ ; b: correlations between internal noise and AQ in ASD and NT groups, the fine colored lines correspond to separated group analysis, the bold black line is the correlation within groups.

Then, as for the typicality, correlation analysis between internal noise and ADOS, AQ and EQ were performed with the same method, and none reached significance. No correlation with ADOS score in ASD group was observed ( $r = .04$ ,  $p = .87$ ).

The correlation coefficient between internal noise and AQ in NT was  $r = .04$  ( $p = .85$ ), and in ASD was  $r = .32$  ( $p = .12$ ). No statistical difference between the two correlation coefficients,  $z = -.97$ ,  $p = .33$  was revealed.

The correlation coefficient between internal noise and EQ in NT was  $r = -.43$  ( $p = .03$ ), and in ASD was  $r = -.06$  ( $p = .77$ ), but not different statistically ( $z = -1.29$ ,  $p = .19$ ). Considering that there was no significant difference in correlation coefficients between groups, a joint analysis of the correlation between AQ or EQ scores and internal noise was conducted. A tendency to a positive correlation between the internal noise and the AQ was observed ( $r = .30$ ,  $p_{\text{corr}} = .06$ , see Figure 3b), the higher the AQ score the higher the internal noise estimated, but there was no correlation with EQ score ( $r = -.24$ ,  $p_{\text{corr}} = .18$ ).

### 3.2.Cluster analysis

#### 3.2.1. Clusters definition

The optimal number of clusters was automatically estimated to be three in our sample with the k-means method with a Hopkins statistic of .72 (above the threshold .5) with representation typicality and internal noise (IN) as factors. The profiles were qualified as: *Typical representation – Low IN*, *Atypical representation – Low IN* and *High IN* and are described in table 2. Clusters are presented with an “illustrative” patient for each cluster in figure 4.

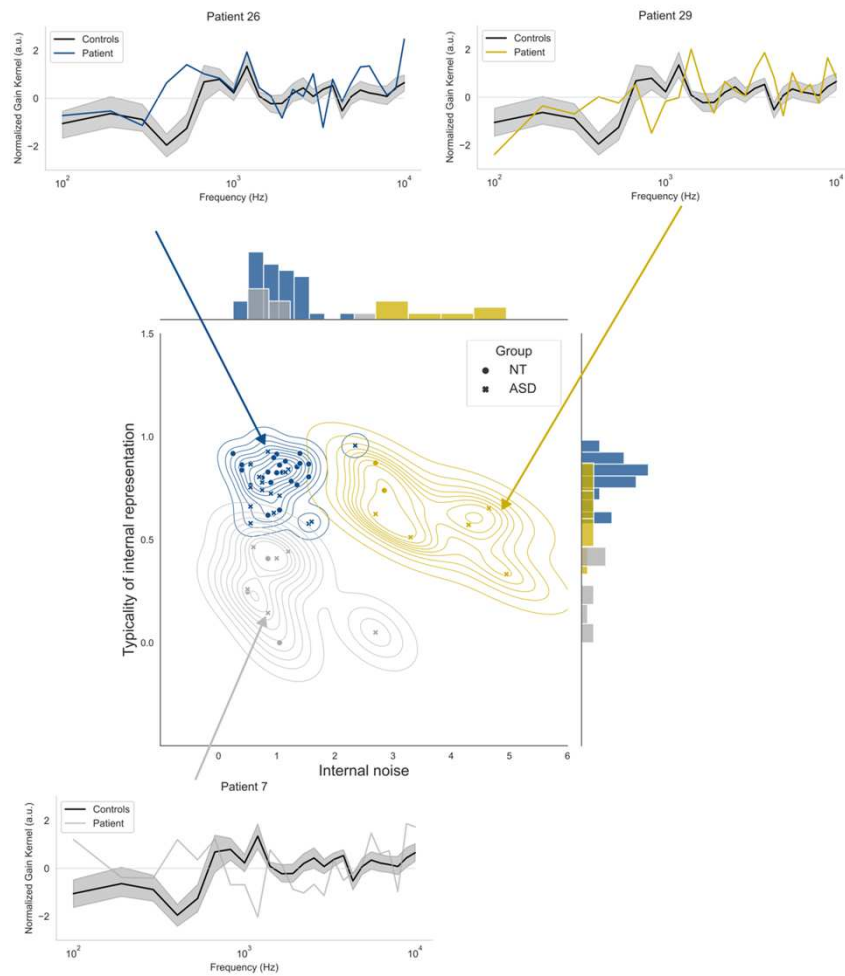


Figure 4. Cluster analysis according to typicality and internal noise in ASD (cross point) and NT (dot point) and an example of a representative patient of each cluster in the chosen color of the clusters. Typical representation – Low IN (blue); Atypical representation – Low IN (grey) and High IN (yellow). The black curve represents the Controls representation of vocal smile. IN: internal noise.

As NT and ASD were considered together in the cluster analysis, the distribution of group in each cluster was compared and no significant association between group and cluster distribution was observed ( $p = .37$  ; Fisher's Exact Test ; FET).

Comparisons of clinical scores for ASD patients were performed between clusters but did not reveal any statistical differences.

### 3.2.2. Qualitative description of clusters

Considering the relatively small size of clusters a qualitative description was performed to evaluate the differences and the clinical implication of these three profiles (table 2). First, the *Typical representation – Low IN* group includes almost the entire group of neurotypical adults (23/29) and a majority of the ASD (17/27) group indicating that as in Ponsot et al. (2018), the internal representation of vocal smile was stable and robust in most NT and ASD adults. However, the fact that a small number of typical adults do not have the same profile of vocal smile and associated internal noise, shows that heterogeneity also exists in the general population. No difference in clinical scores was observed between clusters, but the size of the two alternative clusters (*Atypical representation – Low IN* and *High IN*) did not allow to perform correct statistical analysis.

Group	Typical representation Low IN		Atypical representation Low IN		High IN	
	NT	ASD	NT	ASD	NT	ASD
<b>n</b>	23	17	3	6	3	4
<b>Age</b>	31.4 ± 13.6	32.3 ± 12.1	25.6 ± 1.4	41.4 ± 13.2	24.7 ± 2.8	39 ± 13.2
<b>vIQ</b>	124.4 ± 11.5 (18)	117.4 ± 18.8 (14)	119.7 ± 12.1	117.7 ± 22	127.3 ± 8.6	120.5 ± 30.6
<b>nvIQ</b>	105.9 ± 13.9 (18)	103.8 ± 16.8 (14)	108.7 ± 4.2	95.8 ± 20	106.7 ± 4.6	94.5 ± 23.9
<b>AQ</b>	16.6 ± 7 (20)	33.5 ± 9.0 (16)	15 ± 6.2	35 ± 10.9	19.7 ± 6.5	40.3 ± 2.1
<b>EQ</b>	45.2 ± 8.5 (19)	25.2 ± 13.8 (16)	40.3 ± 10.7	20.5 ± 10.4	33.3 ± 6.7	22.5 ± 13.5
<b>Internal noise</b>	1 ± 0.4	1 ± 0.5	0.8 ± 0.3	1.1 ± 0.8	2.8 ± 0.1	4.3 ± 0.7
<b>ADOS</b>	/	8 ± 5.2	/	10.7 ± 5.5	/	9.3 ± 7.1
<b>Representation typicality</b>	0.8 ± 0.1	0.8 ± 0.1	0.2 ± 0.2	0.3 ± 0.2	0.7 ± 0.1	0.5 ± 0.1

Table 2. Description of clusters in ASD and NT (mean ± standard deviation). The values in brackets correspond to the number of values included in the calculation of the mean and standard deviation due to missing data. IN: internal noise.

Looking at the representations of vocal smile in the *Typical representation – Low IN* revealed that the acoustic cues used to identify a vocal smile were very similar regardless the group

(figure 5a). In contrast, the *Atypical representation – Low IN* models were noisy and did not base their representation of smile on the same acoustic cue. The autistic participants in this group used the inverse information to classify a sound as smiling in comparison to Controls group (figure 5b). Finally, in the *High IN* cluster, acoustic cues were not that different from the Controls’ model but considering the high internal noise in these participants, interpreting their models remains challenging (Figure 5c).

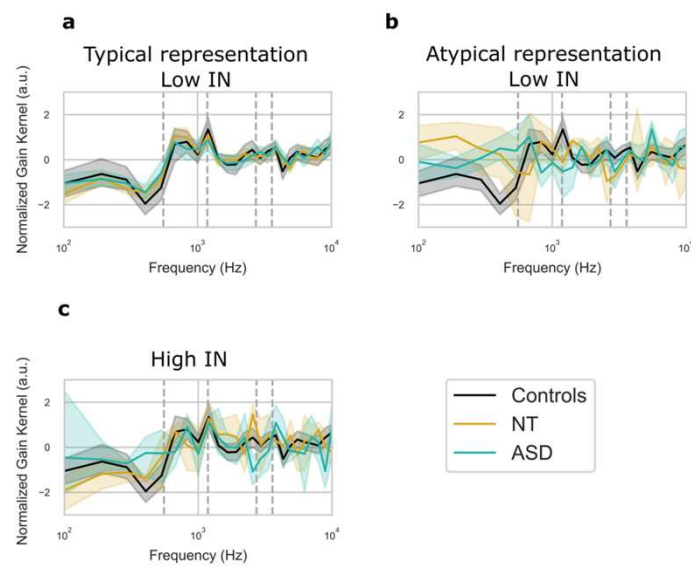


Figure 5. Mental representation of vocal smile in each cluster for NT (orange) and ASD (blue) in comparison to the Controls model (black) with 600 trials. IN: internal noise a: Typical representation and low IN cluster; b: Atypical representation and low IN cluster; c: High In cluster

#### 4. Discussion

This study was the first to characterize mental representation of emotional prosody, a particularly challenging feature, in autism with an emotional auditory reverse correlation paradigm. Results revealed that a large majority of autistic participants based their representation of vocal smile on the same acoustic features as neurotypicals do. However,

differences in the perceptive representation of vocal smile in a subgroup of autistic adults in comparison to neurotypical, but also a larger internal noise in another subgroup of autistic participants, have been observed.

### **4.1.Preserved mental representation of vocal smile**

The majority of autistic participants (17/27) used the same acoustic cues in their mental representation of vocal smile as neurotypical with a robust model (low IN). This result suggests that the representation of the vocal smile is shared by all and preserved in most autistic adults. This is in line with studies that reported no difference in the recognition of emotional prosody neurotypical and autistic participants (Brennand et al., 2011; Brooks & Ploog, 2013; Chevallier et al., 2011; Grossman et al., 2010; Paul et al., 2005). However, some other studies showed difficulties in emotional prosody recognition in autism (Hubbard et al., 2017). Cluster analysis revealed a profile of autistic participants with an atypical representation of vocal smile with limited internal noise, that might contribute to the lower typicality of the internal representation of smile in the ASD group as a whole. This also suggests that difficulties in emotion recognition reported in some studies, might be the consequence of the heterogeneity of the models in some autistic participants.

The current study estimates mental representation of vocal smile based on low-level acoustic analysis and it has been shown that for non-vocal acoustic processing, autistic adults display better skills than neurotypicals (Bonnell et al., 2003, 2010; Heaton et al., 2008). One can hypothesize that studies which report low emotional prosody recognition in autism, used more ecological stimuli and tasks that required higher-level processes than in the present study. The current results demonstrated that the low-level representations of emotional prosody are intact in autism, but one can not exclude that more spontaneous mechanisms as automatic recognition and judgment could be atypical.

In a previous study, even if difficulties in happiness prosody recognition were reported, a positive association between a good recognition of happiness prosody and a better social

adaptation was observed (J.-E. Wang & Tsao, 2015). This relation was not observed between representation typicality and ADOS score in the autism group in the present work, but a positive correlation between the empathetic abilities and the typicality was found. These relationships demonstrated the importance of low-level acoustic processing of emotional prosody in social abilities. Perhaps empathetic abilities, that reflect the ability to understand and share the feelings of other, refine the emotional prosody acoustic processing in neurotypical and autistic adults, or conversely that a typical representation of emotional prosody allows better relations to others and thus develop greater empathy.

### **4.2. Internal noise**

At group level, no difference in internal noise was observed between NT and ASD participants. This lack of difference is compatible with some theories for which noise is not different in autism (Davis & Plaisted-Grant, 2015), but also with the results of a previous reverse correlation study in auditory modality in autistic children in which no difference was shown (L. Wang et al., 2022).

Nevertheless, cluster analysis revealed a subgroup of participant with a *High IN*, mainly composed by autistics (6 ASD vs 3 NT). In this profile the mental representation of vocal smile is similar to Ponsot's results (2018) suggesting that i) unstable noisy representation could lead to both prosody difficulty and hypo- or hyper-sensitivity in autism and ii) the quantity of internal noise is not related to the quality of the representation of vocal smile. Indeed, for some authors a larger internal (neural) noise might facilitate the detection and discrimination of signal (Davis & Plaisted-Grant, 2015). Nonetheless, in this profile the larger IN ( $> 3$ ) were observed in four autistic participants only, a result that should be considered with caution. The heterogeneous results between participants with autism in terms of internal noise levels also reflect the diversity of participants' clinical profiles. Moreover, even in the absence of significant relationships between internal noise level and clinical measures it should be noted that the autistic participants with larger level of intra-individual variability had high ADOS and

AQ scores, reflecting greater autistic symptoms.

Finally, in the context of the Bayesian brain, the difference in internal noise level and mental representation modify the perception of the environment (Haker et al., 2016). The present study offered the possibility to estimate an outline of the mental representation to which the percept is compared, in a Bayesian context, and of the noise that would modulate this representation. In autism the Bayesian theories proposed that the sensory input and the prediction (mental representation) differ in weight, which leads to hypo- or hyper-sensitivity (Brock, 2012; Haker et al., 2016; Lawson et al., 2014; Van de Cruys et al., 2014). The level of noise in the signal and internal noise have been proposed to influence the level of prediction of the sensory input according to the context, and thus noisy inputs would lead to hyposensitivity and precise input to hypersensitivity (Van de Cruys et al., 2017).

### **4.3.Limitations and perspectives**

In the present study, the available clinical measures (ADOS, AQ) did not explain the different profiles. The addition of hyper- and/or hypo-sensitivity assessment, as the Dunn sensory profile (Kern et al., 2007) or the Glasgow Sensory Questionnaire (Robertson & Simmons, 2013) for example, would provide complementary information on the internal noise profiles observed. Moreover, to draw stronger conclusions on the association between clinical measures and the observed profiles, the inclusion of more participant in the alternative experimental profiles (*Atypical representation – Low IN and High IN*) is required, considering, however, that in this study more participants were included than in previous studies of reverse correlations carried in the auditory modality (30 participants in each group versus 10 to 20 in other studies) (Adl Zarrabi et al., 2024; Goupil et al., 2021; Ponsot, Arias, et al., 2018; Ponsot, Burred, et al., 2018). Because the reverse correlation task was a part of a larger protocol, the addition of another task was not possible and thus, no measure of emotional prosody production has been performed. Thought it could have informed on the implication of different representation of vocal emotion on the production. In fact, it has been shown that the measurement of certain acoustic indices

during prosodic productions (frequency of F0 and formants and mean Harmonic to Noise Ratio, among others) enables a precise and sensitive classification according to a diagnosis of autism in children (Briend et al., 2023). Combining this information with the acoustic indices of the mental representation of vocal smile could enable this classificatory analysis to be extended. Musical practice has been shown to enable a finer analysis of the acoustic signal (Molnar-Szakacs & Heaton, 2012), and this could have a beneficial effect on the development of vocal emotion representation in autism (Redondo Pedregal & Heaton, 2021). This finer acoustic analysis might play a role in the analysis of sounds with added social content. Unfortunately, such a detailed analysis of musical practice was not possible in the tested sample due to an insufficient number of musician participants (11/29 in the NT and 8/28 in the ASD group). In the future, it would be interesting to include more musician participants in order to test these potentially beneficial effects on perceptive representations of vocal emotions.

### **5. Conclusion**

To summarize, this study allowed to estimate mental representation of vocal smile, and the internal noise level associated in a clinical population with a short and simple paradigm highlighting fairly preserved processes in autistic adults. This study however revealed different profiles according to both mental representation and internal noise that could contribute to difficulties in recognition and response to emotional prosody in some autistic individuals. Measurement of emotional contagion evoked by prosodic voices in autism should allow to go further in the understanding of prosody atypicalities in the autistic participants who display typical and robust perceptive representation.

## 6. References

- Adl Zarrabi, A., Jeulin, M., Bardet, P., Commère, P., Naccache, L., Aucouturier, J.-J., Ponsot, E., & Villain, M. (2024). A simple psychophysical procedure separates representational and noise components in impairments of speech prosody perception after right-hemisphere stroke. *Scientific Reports*, *14*(1), 15194. <https://doi.org/10.1038/s41598-024-64295-y>
- Ahumada, A., Jr., & Lovell, J. (1971). Stimulus Features in Signal Detection. *The Journal of the Acoustical Society of America*, *49*(6B), 1751–1756. <https://doi.org/10.1121/1.1912577>
- Andronikof, A., & Fontan, P. (2016). Grounia Efimovna Soukhareva: La première description du syndrome dit d'Asperger. *Neuropsychiatrie de l'Enfance et de l'Adolescence*, *64*(1), 58–70. <https://doi.org/10.1016/j.neurenf.2015.07.007>
- Arias, P., Belin, P., & Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. *Current Biology*, *28*(14), R782–R783. <https://doi.org/10.1016/j.cub.2018.05.084>
- Aubergé, V., & Cathiard, M. (2003). Can we hear the prosody of smile? *Speech Communication*, *40*(1–2), 87–97. [https://doi.org/10.1016/S0167-6393\(02\)00077-8](https://doi.org/10.1016/S0167-6393(02)00077-8)
- Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. <https://doi.org/10.1023/B:JADD.0000022607.19833.00>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. <https://doi.org/10.1023/a:1005653411471>
- Beall, P. M., Moody, E. J., McIntosh, D. N., Hepburn, S. L., & Reed, C. L. (2008). Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. *Journal of Experimental Child Psychology*, *101*(3), 206–223. <https://doi.org/10.1016/j.jecp.2008.04.004>
- Black, M. H., Chen, N. T. M., Iyer, K. K., Lipp, O. V., Bölte, S., Falkmer, M., Tan, T., & Girdler, S. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. *Neuroscience & Biobehavioral Reviews*, *80*, 488–515. <https://doi.org/10.1016/j.neubiorev.2017.06.016>
- Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., Burack, J. A., & Mottron, L. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia*, *48*(9), 2465–2475. <https://doi.org/10.1016/j.neuropsychologia.2010.04.020>
- Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *J. Cogn. Neurosci*, *15*,

226–235.

Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, 5(4), 1567–1573. <https://doi.org/10.1016/j.rasd.2011.03.002>

Briend, F., David, C., Silleresi, S., Malvy, J., Ferré, S., & Latinus, M. (2023). Voice acoustics allow classifying autism spectrum disorder with high accuracy. *Translational Psychiatry*, 13(1), 1–8. <https://doi.org/10.1038/s41398-023-02554-8>

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, 28(1), 333–361. <https://doi.org/10.1080/10463283.2017.1381469>

Brock, J. (2012). Alternative Bayesian accounts of autistic perception: Comment on Pellicano and Burr. *Trends in Cognitive Sciences*, 16(12), 573–574. <https://doi.org/10.1016/j.tics.2012.10.005>

Brooks, P. J., & Ploog, B. O. (2013). Attention to emotional tone of voice in speech perception in children with autism. *Research in Autism Spectrum Disorders*, 7(7), 845–857. <https://doi.org/10.1016/j.rasd.2013.03.003>

Burred, J. J., Ponsot, E., Goupil, L., Liuni, M., & Aucouturier, J.-J. (2019). CLEESE: An open-source audio-transformation toolbox for data-driven experiments in speech and music cognition. *PLOS ONE*, 14(4), e0205943. <https://doi.org/10.1371/journal.pone.0205943>

Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, 49(3), 507–517. <https://doi.org/10.1016/j.neuropsychologia.2010.11.042>

Davis, G., & Plaisted-Grant, K. (2015). Low endogenous neural noise in autism. *Autism: The International Journal of Research and Practice*, 19(3), 351–362. <https://doi.org/10.1177/1362361314552198>

Dawson, G., Hill, D., Spencer, A., Galpert, L., & Watson, L. (1990). Affective exchanges between young autistic children and their mothers. *Journal of Abnormal Child Psychology*, 18(3), 335–345. <https://doi.org/10.1007/BF00916569>

Ewing, L., Pellicano, E., King, H., Lennuyeux-Comnene, L., Farran, E. K., Karmiloff-Smith, A., & Smith, M. L. (2018). Atypical information-use in children with autism spectrum disorder during judgments of child and adult face identity. *Developmental Neuropsychology*, 43(4), 370–384. <https://doi.org/10.1080/87565641.2018.1449846>

Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-020-20649-4>

- Grossman, R. B., Bemis, R. H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *J. Speech Lang. Hear. Res. JSLHR*, *53*, 778–793.
- Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian Theories of Autism Spectrum Disorder Help Improve Clinical Practice? *Frontiers in Psychiatry*, *7*. <https://www.frontiersin.org/articles/10.3389/fpsy.2016.00107>
- Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology*, *25*(6), 771–782. <https://doi.org/10.1080/02643290802336277>
- Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder: Affective prosody in ASD. *Autism Research*, *10*(12), 1991–2001. <https://doi.org/10.1002/aur.1847>
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217–250.
- Kassambara, A. (2020). *ggpubr: “ggplot2” Based Publication Ready Plots*. <https://CRAN.R-project.org/package=ggpubr>
- Kassambara, A., & Mundt, F. (2017). Package ‘factoextra.’ *Extract and Visualize the Results of Multivariate Data Analyses*, *76*(2).
- Kern, J. K., Garver, C. R., Carmody, T., Andrews, A. A., Trivedi, M. H., & Mehta, J. A. (2007). Examining sensory quadrants in autism. *Research in Autism Spectrum Disorders*, *1*(2), 185–193. <https://doi.org/10.1016/j.rasd.2006.09.002>
- Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, *8*. <https://doi.org/10.3389/fnhum.2014.00302>
- Leung, F. Y. N., Sin, J., Dawson, C., Ong, J. H., Zhao, C., Veić, A., & Liu, F. (2022). Emotion recognition across visual and auditory modalities in autism spectrum disorder: A systematic review and meta-analysis. *Developmental Review*, *63*, 101000. <https://doi.org/10.1016/j.dr.2021.101000>
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, *30*(3), 205–223. <https://doi.org/10.1023/A:1005592401947>
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685. <https://doi.org/10.1007/BF02172145>

- Magnuson, J. R., Iarocci, G., Doesburg, S. M., & Moreno, S. (2020). Increased Intra-Subject Variability of Reaction Times and Single-Trial Event-Related Potential Components in Children With Autism Spectrum Disorder. *Autism Research, 13*(2), 221–229. <https://doi.org/10.1002/aur.2210>
- Manenti, M., Ferré, S., Tuller, L., Houy-Durand, E., Bonnet-Brilhault, F., & Prévost, P. (2024). Profiles of structural language and nonverbal intellectual abilities in verbal autistic adults. *Research in Autism Spectrum Disorders, 114*, 102361. <https://doi.org/10.1016/j.rasd.2024.102361>
- McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language & Communication Disorders, 38*(4), 325–350. <https://doi.org/10.1080/1368282031000154204>
- Merchie, A., Ranty, Z., Aguillon-Hernandez, N., Aucouturier, J.-J., Wardak, C., & Gomot, M. (2024). 'My face is listening to your smile': Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance. OSF. <https://doi.org/10.31234/osf.io/5crey>
- Miller, L. K. (1999). The Savant Syndrome: Intellectual impairment and exceptional skill. *Psychological Bulletin, 125*(1), 31–46. <https://doi.org/10.1037/0033-2909.125.1.31>
- Molnar-Szakacs, I., & Heaton, P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences, 1252*(1), 318–324. <https://doi.org/10.1111/j.1749-6632.2012.06465.x>
- Neri, P. (2010). How inherently noisy is human sensory processing? *Psychonomic Bulletin & Review, 17*(6), 802–808. <https://doi.org/10.3758/PBR.17.6.802>
- Park, W. J., Schauder, K. B., Zhang, R., Bennetto, L., & Tadin, D. (2017). High internal noise and poor external noise filtering characterize perception in autism spectrum disorder. *Scientific Reports, 7*(1), Article 1. <https://doi.org/10.1038/s41598-017-17676-5>
- Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205–220. <https://doi.org/10.1007/s10803-004-1999-1>
- Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research, 50*(4), 1015–1028. [https://doi.org/10.1044/1092-4388\(2007/071\)](https://doi.org/10.1044/1092-4388(2007/071))
- Ponsot, E., Arias, P., & Aucouturier, J.-J. (2018). Uncovering mental representations of smiled speech using reverse correlation. *The Journal of the Acoustical Society of America, 143*(1), EL19–EL24. <https://doi.org/10.1121/1.5020989>
- Ponsot, E., Burred, J. J., Belin, P., & Aucouturier, J.-J. (2018). Cracking the social code of

speech prosody using reverse correlation. *Proceedings of the National Academy of Sciences*, 115(15), 3972–3977. <https://doi.org/10.1073/pnas.1716090115>

Redondo Pedregal, C., & Heaton, P. (2021). Autism, music and Alexithymia: A musical intervention to enhance emotion recognition in adolescents with ASD. *Research in Developmental Disabilities*, 116, 104040. <https://doi.org/10.1016/j.ridd.2021.104040>

Rimland, B., & Fein, D. (1988). Special talents of autistic savants. In *The exceptional brain: Neuropsychology of talent and special abilities* (pp. 474–492). The Guilford Press.

Robertson, A. E., & Simmons, D. R. (2013). The Relationship between Sensory Sensitivity and Autistic Traits in the General Population. *Journal of Autism and Developmental Disorders*, 43(4), 775–784. <https://doi.org/10.1007/s10803-012-1608-7>

Song, Y., Kawabe, T., Hakoda, Y., & Du, X. (2012). Do the eyes have it? Extraction of identity and positive expression from another's eyes in autism, probed using “Bubbles.” *Brain and Development*, 34(7), 584–590. <https://doi.org/10.1016/j.braindev.2011.09.009>

Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics*, 27(1), 24–27. <https://doi.org/10.3758/BF03199901>

Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *The Journal of the Acoustical Society of America*, 96(4), 2101–2107. <https://doi.org/10.1121/1.410151>

Team, R. C. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.

Team, Rs. (2020). RStudio: Integrated Development for R. In *RStudio, PBC*. <http://www.rstudio.com/>.

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, 121(4), 649. <https://doi.org/10.1037/a0037665>

Van de Cruys, S., Van der Hallen, R., & Wagemans, J. (2017). Disentangling signal and noise in autism spectrum disorder. *Brain and Cognition*, 112, 78–83. <https://doi.org/10.1016/j.bandc.2016.08.004>

Vilidaite, G., Yu, M., & Baker, D. H. (2017). Internal noise estimates correlate with autistic traits. *Autism Research: Official Journal of the International Society for Autism Research*, 10(8), 1384–1391. <https://doi.org/10.1002/aur.1781>

Wang, J.-E., & Tsao, F.-M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, 37, 162–170. <https://doi.org/10.1016/j.ridd.2014.11.013>

Wang, L., Ong, J. H., Ponsot, E., Hou, Q., Jiang, C., & Liu, F. (2022). Mental representations of speech and musical pitch contours reveal a diversity of profiles in autism spectrum disorder. *Autism*, 13623613221111207. <https://doi.org/10.1177/13623613221111207>

Ward, J. (2019). Individual differences in sensory sensitivity: A synthesizing framework and evidence from normal variation and developmental conditions. *Cognitive Neuroscience*, 10(3), 139–157. <https://doi.org/10.1080/17588928.2018.1557131>

Wechsler, D. (2012). *Wechsler Adult Intelligence Scale—Fourth Edition* [Dataset]. <https://doi.org/10.1037/t15169-000>

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <https://ggplot2.tidyverse.org>.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>

Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. (2016). Fashioning the Face: Sensorimotor Simulation Contributes to Facial Expression Recognition. *Trends in Cognitive Sciences*, 20(3), 227–240. <https://doi.org/10.1016/j.tics.2015.12.010>

***Fundings sources***

This work was supported by the INSERM (National Institute for Health and Medical Research). Annabelle Merchie was funded by the FEANS (Fondation Européenne pour l'Avancement des Neurosciences) and the Agence Nationale de la Recherche (ANR-19-CE37-0022-01). The funding sources were not involved in the study design, data collection and analysis, or the writing of the report.

***Competing interests***

All authors declare that they have no conflicts of interest.

***Ethics approval***

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The protocol received approval from Ethics Committee (PROSCEA2017/23; ID RCB: 2017-A00756-47).

***Data and code availability***

Please contact the corresponding author for data and experimental stimuli requests.

***Author's contributions***

Annabelle Merchie, Jean-Julien Aucouturier and Marie Gomot designed the study. Emmanuelle Houy-Durand was responsible for the recruitment and the clinical assessment of patients. Annabelle Merchie and Zoé Ranty performed data acquisition. Annabelle Merchie, Zoé Ranty, Aynaz Adl Zarrabi, Jean-Julien Aucouturier and Marie Gomot were responsible for data and statistical analyses. Annabelle Merchie, Jean-Julien Aucouturier and Marie Gomot wrote the first version of the manuscript. All authors were involved in preparing and reviewing the manuscript.

### 2.3. Synthèse

La représentation perceptive des émotions vocales semble préservée chez une majorité d'adultes autistes, sans lien avec les mesures cliniques, et corrélée avec les compétences empathiques. Il semblerait que l'étape de représentation de la boucle soit préservée pour une majorité des patients. Par ailleurs, un tiers de l'échantillon présentait un profil différent au niveau de la représentation ou de la variabilité des réponses. Ces différences dans les modèles pourraient expliquer les différences dans la reconnaissance des émotions, mais aussi peut-être dans la réponse automatique à celles-ci. Il apparaît donc, que ce ne soit pas des représentations perceptuelles différentes des émotions vocales qui soient en cause dans l'absence de contagion et de résonance émotionnelle dans la modalité auditive. Se pose alors la question de la résonance motrice et de la contagion émotionnelle chez ces patients avec un modèle similaire. Est-ce suffisant pour observer une réponse musculaire faciale congruente avec l'émotion perçue ?

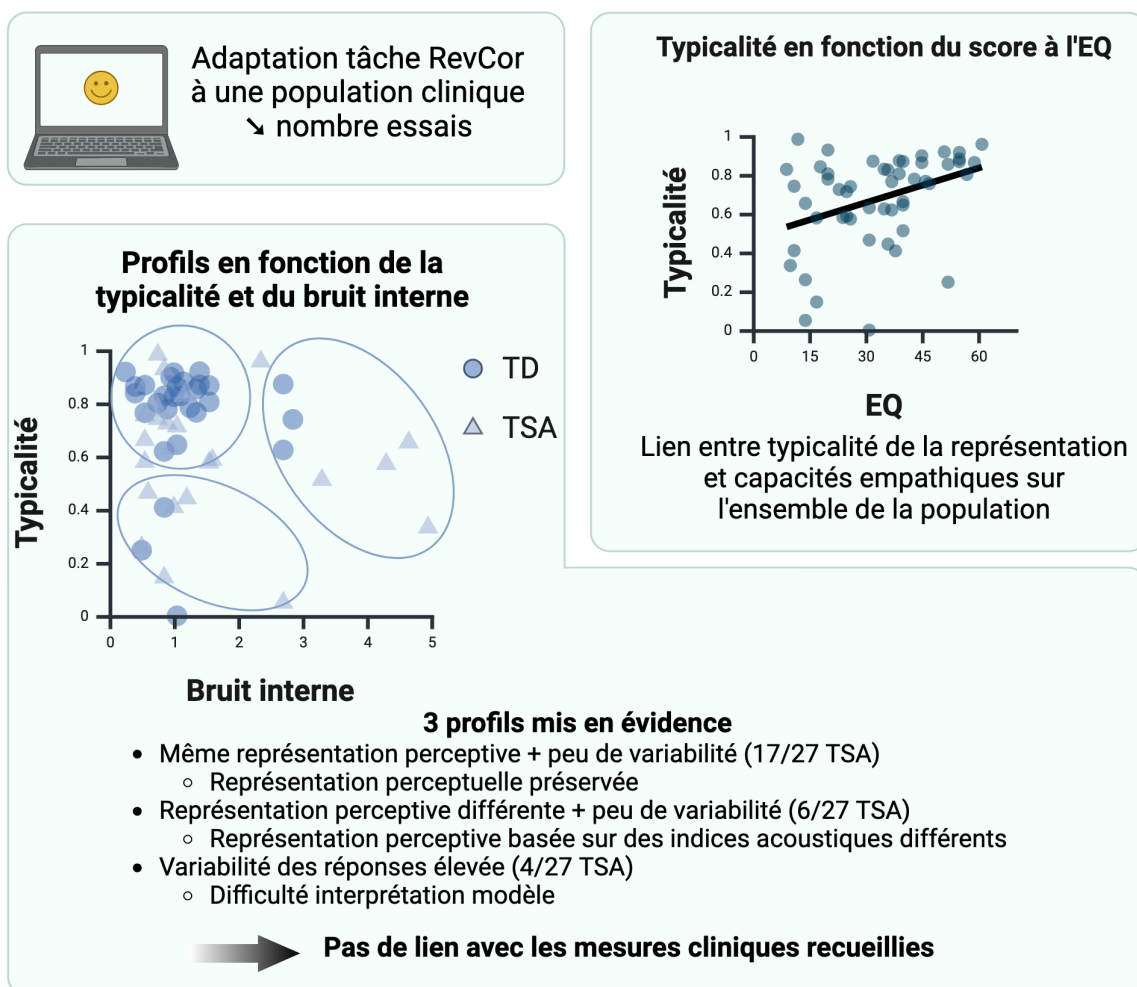


Figure 20 : Schéma de synthèse de l'étude n°3 à propos de la représentation perceptive du sourire vocal dans l'autisme



### **3. Résonance motrice et contagion émotionnelle au sourire vocal**

#### **3.1. Avant-propos**

Un modèle du sourire vocal similaire entre une grande partie des adultes autistes et neurotypiques a été mis en évidence dans l'étude n°3. Des études ont montré que ce modèle du sourire vocal pouvait induire une résonance motrice congruente chez des adultes neurotypiques et chez des aveugles congénitaux (Arias, Bellmann, et al., 2021; Arias et al., 2018). Cependant, ces deux études ne permettent pas, en raison du manque de mesures physiologiques complémentaires attestant de l'état émotionnel, de conclure sur une potentielle contagion émotionnelle en réponse au sourire vocal.

L'objectif de cette dernière partie expérimentale est de mettre en évidence une possible contagion émotionnelle au sourire vocal chez l'adulte neurotypique et autiste à travers la combinaison de mesures des expressions faciales et de la dilatation pupillaire. Une première étude est consacrée à l'étude de ces phénomènes chez l'adulte neurotypique (étude n°4). La seconde étude, vise à explorer la résonance motrice chez les personnes autistes en lien avec la représentation perceptive du sourire vocal (étude n°5) en comparaison à un groupe d'adultes neurotypiques.

### 3.2. Étude n°4 : Résonance motrice et autonome au sourire vocal chez l'adulte neurotypique

#### **Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance**

**Merchie, A.**<sup>\*</sup>, **Ranty, Z.**<sup>\*</sup>, *Aguillon-Hernandez, N., Aucouturier, J-J., Wardak, C. & Gomot, M.*

<sup>\*</sup>: co-premières autrices

*Article sous presse dans Scientific Reports*

**Introduction et objectifs :** Des phrases modifiées artificiellement pour être rendues souriantes, d'après une représentation perceptive du sourire vocal validée chez l'adulte neurotypique, entraînent une résonance motrice congruente (Arias et al., 2018, 2021). Cependant, dans ces deux études seule la réponse motrice était évaluée, ce qui ne permet pas de conclure à une contagion émotionnelle au sourire vocal. L'objectif de cette étude était d'évaluer la contagion émotionnelle au sourire vocal chez l'adulte neurotypique grâce à la mesure combinée de l'activité musculaire faciale et du diamètre de la pupille, un indicateur de l'activité du système nerveux autonome associée à l'état émotionnel.

**Matériel et méthodes :** L'activité musculaire faciale et le diamètre pupillaire ont été mesurés chez 25 adultes neurotypiques en réponse à des phrases rendues artificiellement souriantes ou non-souriantes qui devaient être jugées selon leur « degré de sourire ». Une validation du modèle du sourire vocal dans ce groupe d'adultes a été réalisée à l'aide d'une tâche de corrélation inverse, afin d'estimer un lien potentiel entre le modèle individuel et la contagion émotionnelle.

**Résultats :** En réponse au sourire vocal, une résonance motrice implicite, via un effet du filtre souriant sur le *Zygomaticus major*, et explicite, via un effet du jugement non-souriant sur le *Corrugator supercilii* ont été mises en lumière. De plus, une analyse temporelle de la dynamique de la réponse pupillaire a révélé un effet du contenu émotionnel, puis de la valence et enfin du choix, indiquant un processus de contagion émotionnelle en plusieurs étapes chez l'adulte neurotypique.

**Discussion :** L'évaluation de la réactivité motrice et autonome face aux émotions vocales et à leur interprétation a permis de développer un modèle de la contagion. Cette étude contribue à l'enrichissement des théories des émotions, à propos des boucles de rétroaction entre la perception des émotions et leurs expressions. La réalisation d'une étude similaire dans l'autisme permettra d'informer sur les processus qui conduisent à une résonance motrice atypique.



# OPEN Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance

Annabelle Merchie<sup>1,4</sup>, Zoé Ranty<sup>1,4</sup>, Nadia Aguilon-Hernandez<sup>1</sup>, Jean-Julien Aucouturier<sup>2,3</sup>, Claire Wardak<sup>1</sup> & Marie Gomot<sup>1</sup>✉

The interplay between the different components of emotional contagion (i.e. emotional state and facial motor resonance), both during implicit and explicit appraisal of emotion, remains controversial. The aims of this study were (i) to distinguish between these components thanks to vocal smile processing and (ii) to assess how they reflect implicit processes and/or an explicit appraisal loop. Emotional contagion to subtle vocal emotions was studied in 25 adults through motor resonance and Autonomic Nervous System (ANS) reactivity. Facial expressions (fEMG: facial electromyography) and pupil dilation were assessed during the processing and judgement of artificially emotionally modified sentences. fEMG revealed that *Zygomaticus major* was reactive to the perceived valence of sounds, whereas the activity of *Corrugator supercilii* reflected explicit judgement. Timing analysis of pupil dilation provided further insight into both the emotional state and the implicit and explicit processing of vocal emotion, showing earlier activity for emotional stimuli than for neutral stimuli, followed by valence-dependent variations and a late judgement-dependent increase in pupil diameter. This innovative combination of different electrophysiological measures shed new light on the debate between central and peripheral views within the framework of emotional contagion.

**Keywords** Motor resonance, Emotional contagion, Pupil reactivity, fEMG, Vocal smile

Emotional contagion corresponds to a tendency to infer and react to the sensory, motor, physiological and affective states of others<sup>1</sup>. Three invariable characteristics can be identified: it is primitive, automatic, and implicit. This process enables humans to understand and feel others' emotions, thoughts and intentions<sup>2</sup>. Emotional contagion encompasses both the emotional state (possibly through automatic mapping of emotions<sup>3</sup>) and its expression in the receiver, such as congruent facial expression through the motor resonance process<sup>4</sup>. However, the role of facial output has been widely debated, as it could constitute both a readout of the emotional state and an input for the subjective experience of emotion. Some authors have even argued in favor of the primacy of facial expression over emotional experience<sup>5</sup>, although the effect of facial feedback on subtle emotion remains a controversial issue<sup>6,7</sup>. The hypothesis of a facial feedback, supports more comprehensive theories according to which the emotional state is modulated by feedback from the peripheral bodily sensations<sup>8</sup>, as opposed to central theories that postulate that they would reflect independent components of the emotional response<sup>9</sup>. A complementary theoretical framework suggests that the emotional state would be largely influenced by cognitive appraisal, which is described within the Component Process Model (CPM) as fast and unconscious<sup>10,11</sup>.

Previous studies have explored how facial feedback both accompanies and potentially influences the perception of emotional stimuli, as well as how it modulates *subjective* emotional experience<sup>12</sup>. Additionally, these studies have examined the role of cognitive appraisal (i.e., any perceptual or cognitive process) in shaping facial expressions, through unconscious mimicry<sup>13</sup>. However, in the framework of emotional contagion, studies have never comprehensively described the interplay between facial motor resonance and *objective* markers of emotional states, whether during implicit or explicit appraisal of emotion. The present study proposes to fill this gap by using objective physiological measures of the two components of the emotional response during the judgement of subtle vocal emotions.

A well-known example of emotional contagion is motor resonance to visual smile. It has already been detected in infants<sup>14</sup> and in several species<sup>15</sup>. Because smiling is an important social cue, the automatic response

<sup>1</sup>INSERM, Imaging Brain & Neuropsychiatry iBrain U1253, Université de Tours, Tours 37032, France. <sup>2</sup>FEMTO-ST Institute, CNRS, Université de Bourgogne Franche Comté, Besançon, France. <sup>3</sup>STMS Lab IRCAM, CNRS, Sorbonne Université, Paris, France. <sup>4</sup>These authors contributed equally: Annabelle Merchie and Zoé Ranty. ✉email: gomot@univ-tours.fr

to someone else's smile contributes to the matching the sender's feeling<sup>16,17</sup>, the synchronization of feelings<sup>18</sup> and social reciprocity<sup>19</sup>.

Emotions can also be communicated vocally, through modulations of several acoustic cues (pitch, timbre, speech rate) that constitute emotional prosody. Smiling is a great example of mouth position modulation. It involves the bilateral contraction of *Zygomaticus major* (ZM) muscles<sup>20</sup> that induces mouth retraction, leading to voice-changing vocal tract modifications. Stretching the lips causes a reduction in vocal tract length, and an increase in formants frequency, making the voice brighter<sup>21</sup>. Auditory cues have been shown to be sufficient to identify a smile on the basis of voice alone<sup>22–24</sup>. Vocal smile has a robust mental representation in adults, characterized by an upward shift in the frequency of F1 and F2 formants and an increase in energy in F2, F3 and F4<sup>25</sup>.

Facial motor resonance is thus possible in response to auditory information, as demonstrated by the many studies that have explored reactions to highly salient stimuli such as a baby's cry or the cheering of a crowd from the IADS<sup>26</sup> (IADS: International Affective Digitized Sounds) sound database<sup>27–29</sup>. A contraction of ZM in response to pleasant vocalizations and a relaxation in response to unpleasant sounds have also been observed<sup>29</sup>. Hietanen et al.<sup>27</sup> used vocal expressions of affection (i.e. 'Sarah' pronounced with 10 different emotional expressions, such as astonishment and sadness) that resulted in congruent facial motor resonance.

It has been shown that motor resonance to visual emotional stimuli is unconscious<sup>30,31</sup> and may reflect implicit emotion processing<sup>13</sup>. However, it should be noted that for vocal affect, motor resonance has been elicited only under attentional conditions, while participants were actively evaluating the stimuli they were listening to<sup>31</sup>. This makes it difficult to dissociate implicit and explicit contributions. Arias and colleagues<sup>32,33</sup> designed prosodic filters based on the perceptual representation of the vocal smile that resulted from previous research involving a psychophysical task<sup>25</sup>. By using this filter to create artificially smiling sentences, they showed a congruent facial motor resonance with an increase of ZM activity in response to smiling sentences and an increase of *Corrugator supercilii* (CS) activity in response to sentences rated as unsmiling in sighted adults<sup>33</sup> but also in congenitally blind participants<sup>34</sup>. Interestingly, in these studies ZM activity increased for smiling sentences even when participants did not rate them correctly, suggesting that implicit vocal smile processing elicits a motor resonance that would be at least partly independent of explicit judgement. In contrast, CS activity was directly associated with the rating. This indicates that the paradigm used by Arias and colleagues makes it possible to distinguish between implicit processes and explicit appraisal involved in emotional contagion.

In parallel to facial muscle resonance, other physiological indices reflect the emotional state of the receiver. As an objective<sup>35</sup> and face located index of the Autonomic Nervous System (ANS), pupil reactivity reflects the activation of the locus coeruleus-norepinephrine system<sup>36</sup> and sympathetic activity<sup>37–41</sup> but also the inhibition of the parasympathetic Edinger-Westphal nucleus<sup>42</sup> related to emotional processing.

Prochazkova defined a model of Neurological Mechanisms of Emotional Contagion (NMEC) in which the emotional state of the 'sender' is reflected by the ANS of the 'receiver' to converge in a common physiological (i.e. pupil dilation and motor contagion) and cognitive state<sup>43</sup>.

Pupil reactivity in response to emotion has been widely studied through the visual modality<sup>37,38,44,45</sup>. Greater pupil dilation has been detected in response to emotional facial expressions than in response to neutral facial expressions. In addition, pupil dilation has been shown to be modulated by the ecological character of the stimulus: the more natural (i.e. dynamic rather than static) the stimulus, the more the pupil dilates<sup>38</sup>. The nature of emotions also has an impact on pupil reactivity. For example, happy faces induce a larger pupil dilation compared to sad faces<sup>46</sup>, or conversely according to other authors<sup>38</sup>. Although inconsistent results have been reported, there is broad consensus on the effect of stimulus valence on pupil dilation. In the visual modality, face-to-face interaction generally leads to a synchronization of pupil size. Whether this synchronization is purely related to pupil mimicry<sup>47</sup> or reflects an autonomic response to a salient socioemotional stimulus remains unknown. One way to disentangle these two processes is to study pupil reactivity to non-visual stimuli.

Sounds, including emotional sounds such as vocalizations or voices, also induce pupil dilation<sup>48–50</sup> that is modulated, as is the case with visual stimuli, by authenticity<sup>51</sup>. Strong auditory emotions through vocalizations, whether positive (baby laughing) or negative (a couple arguing), induce pupil dilation<sup>52,53</sup>, with a larger effect for negative than for positive stimuli<sup>39</sup>. In the context of emotional prosody, a difference between negative and positive emotions has been observed, with negative emotional sentences inducing a stronger pupil dilation than positive emotions, in a protocol using very salient emotions from the IADS database<sup>54</sup>, while other studies described the opposite effect. However, the study of pupil reactivity to emotional sounds is still in its infancy and has not yet been extended to ambiguous or subtle emotions.

The main objective of the current study was twofold:

- (i) To distinguish between the different components of emotional contagion (i.e. motor resonance and emotional state) and to gain insight into their interplay.
- (ii) To assess how these components reflect implicit processes and/or a potential explicit appraisal loop that would reinforce vocal emotional contagion.

For this purpose, we investigated facial motor resonance and autonomic reactivity (pupil dilation) to vocal smile. These indices were combined so that the various components of emotional contagion could be identified. To unravel implicit and explicit motor resonance to vocal smile, facial muscle activity (ZM and CS) was measured while participants listened (processing phase) and rated (judgment phase) standardized sentences containing subtle emotions. Using subtle emotions would help prevent saliency effects and estimate the implicit processing through poorly-recognized trials.

## Results

### Validation of vocal smile model

According to an inverse correlation estimate of the vocal smile model<sup>25</sup>, the smiling effect is characterized by an increase and shift of energy in F1 and F2, and by an increase of energy in F3 and F4. Modulations (referred to as Smiling and Unsmiling Filters) were applied on semantically and prosodically neutral sentences (see Supplementary materials).

To confirm that the participants in the current study used the same acoustic cues to characterize the vocal smile as Ponsot et al., the same reverse correlation task was performed<sup>25</sup>. Results revealed that the vocal smile model was identical in this new group of participants. A more detailed analysis of the reverse correlation results is presented in the Supplementary materials (Figure S1).

### Accuracy

In the present study, artificial prosodic modifications were correctly recognized by participants (Smiling and Unsmiling Filters) with an accuracy of  $61 \pm 5\%$  (mean  $\pm$  sd), above 50% ( $t(24) = 11.55, p < .001$ ).

### Motor resonance

Considering that the aim of this work was to compare the muscular activity for Filter, Choice and Response, neutral sentences were removed from the analysis.

Because the two muscles have distinct activity, two separate analyses were performed (see Supplementary materials, Figure S2).

The permutations test revealed significant differences between 3000 and 3400 ms for the Filter condition on *Zygomaticus major* (ZM) activity and between 1500 and 3500 ms for the Choice condition on *Corrugator supercilii* (CS) activity. To standardize the analyses, amplitude measurements were performed over a common 500 ms time window (2900–3400 ms) for each muscle and condition.

#### *Zygomaticus major* (ZM) activity

A repeated-measures ANOVA was performed to explore the effects of Filter (Smiling vs. Unsmiling), Choice (Smile vs. Unsmile) and the interaction between these two factors on mean muscular activity of ZM in the 2900–3400ms time window. The main effect analysis revealed a significant effect of Filter ( $F(1,23) = 4.73, p < .05, \eta^2 = 0.06$ ) (Table 1) (Fig. 1A upper part and Fig. 1C). A Smiling sentence induced higher activity of ZM than an Unsmiling sentence (Fig. 1C). Meanwhile, there was no main effect of Choice ( $F(1,23) = 0.66, p > .05, \eta^2 = 0.007$ ) (Table 1) (Fig. 1A lower part and Fig. 1D) and no interaction between Filter and Choice ( $F(1,23) = 0.28, p > .05, \eta^2 = 0.002$ ). A repeated-measures ANOVA considering the effects of Choice (Smile vs. Unsmile) and Response (Correct vs. Incorrect) did not yield different results (see Supplementary material, Figure S3).

#### *Corrugator supercilii* (CS) activity

The same analysis was performed on CS activity with a repeated-measures ANOVA to explore the effects of Filter (Smiling vs. Unsmiling), Choice (Smile vs. Unsmile) and the interaction between these two factors on the mean muscular activity of CS in the 2900–3400ms time window. There was no main effect of Filter on the mean amplitude activity of CS ( $F(1,23) = 1.88, p > .05, \eta^2 = 0.01$ ) (Table 1) (Fig. 1B upper part and Fig. 1C) and no interaction between Filter and Choice ( $F(1,23) = 3.33, p > .05, \eta^2 = 0.01$ ). In parallel, the main effect analysis showed that Choice ( $F(1,23) = 10.18, p < .01, \eta^2 = 0.09$ ) had a significant effect on the mean amplitude activity of CS (Table 1). An Unsmile rated sentence (regardless of the filter) induced higher activity of the CS muscle than a Smile rated sentence (Fig. 1B lower part and Fig. 1D) (Table 1). A repeated-measures ANOVA considering the effects of Choice (Smile vs. Unsmile) and Response (Correct vs. Incorrect) did not show any effect of Response nor an interaction with Choice (see Supplementary material, Figure S4).

### Pupil emotional reactivity

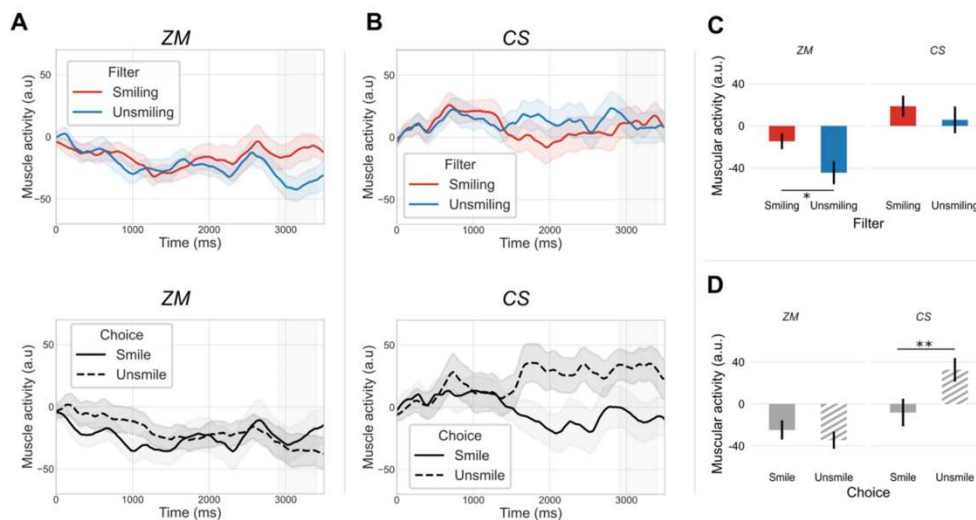
Randomizations tests revealed differences in pupil dilation (PD) between Emotional and Neutral filters from 2100 to 2600 ms. To standardize the analyses, the same time window was used for each condition.

#### Emotional filters

A repeated-measures ANOVA was performed to analyze the effects of Emotional (both smiling and unsmiling) vs. Neutral filters on the mean PD in the 2100–2600ms window. Emotional filters had a significant main effect ( $F(1,24) = 6.61, p < .05, \eta^2 = 0.06$ ) on the mean pupil dilation. Dilation was larger when listening to emotional sentences compared to neutral sentences (Table 2) (Fig. 2A).

Condition	Filter		Choice	
	Smiling	Unsmiling	Smile	Unsmile
ZM	-14.7 $\pm$ 7	-44.6 $\pm$ 11	-24.7 $\pm$ 9	-34.5 $\pm$ 8
CS	18.7 $\pm$ 10	5.8 $\pm$ 13	-8.1 $\pm$ 13	32.5 $\pm$ 11

**Table 1.** Mean muscular activity amplitude in the 2900–3400ms window in response to smiling and unsmiling filtered sentences (a.u.: arbitrary unit.) (mean  $\pm$  standard error of the mean), ZM: *Zygomaticus major*, CS: *Corrugator supercilii*.



**Fig. 1.** Zygomaticus major (ZM) and Corrugator supercilii (CS) activity. A: Effect of filter (upper panel) and choice (lower panel) on ZM; B: Effect of filter (upper panel) and choice (lower panel) on CS; Effect of filter (C) and of the given choice (D) on CS and ZM – mean response amplitude in the window 2900–3400ms (in grey in A and B time series). Shaded areas represent standard error of the mean. \* $p < .05$ ; \*\* $p < .01$ .

		Filter		Choice	
		Neutral	Emotional	Smile	Unsmile
Pupil dilation		0.14 ± 0.03	0.21 ± 0.02	0.17 ± 0.02	0.20 ± 0.02
		Filter response		Filter response	
		Smiling	Unsmiling	Smiling incorrect	Unsmiling correct
Pupil dilation	/	0.18 ± 0.02	0.22 ± 0.02	0.17 ± 0.03	0.20 ± 0.03
				0.19 ± 0.02	0.25 ± 0.02

**Table 2.** Mean pupil dilation in the 2100–2600ms window in response to neutral and emotional sentences, according to the applied filter, choice and response (mm) (mean ± standard error of the mean).

**Filter and choice**

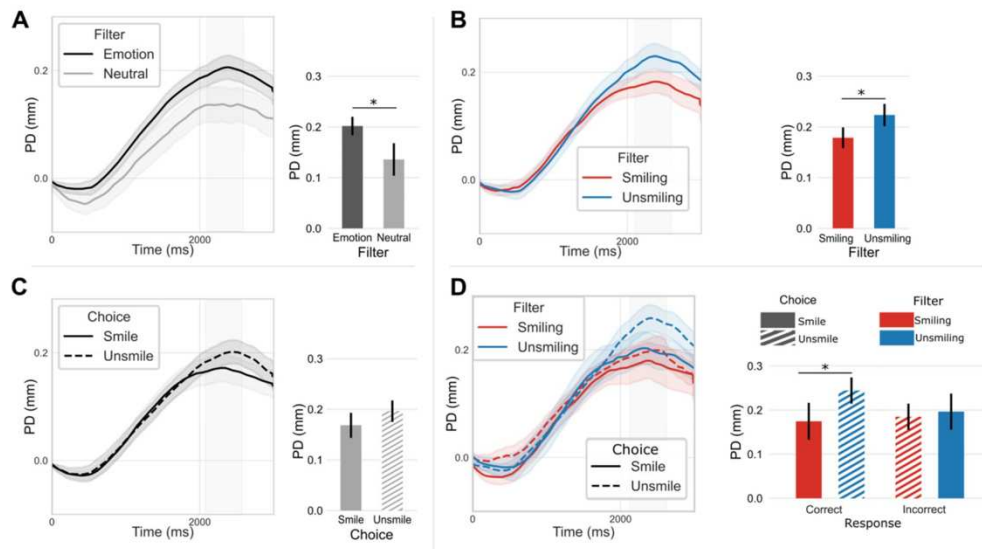
A repeated-measures ANOVA was performed to analyze the effects of Filter (Smiling vs. Unsmiling), Choice (Smile vs. Unsmile) and the interaction between these two factors on mean pupil dilation (PD) in the same time window (2100–2600ms). The main effect analysis showed that Filter had a significant effect ( $F(1,24) = 4.30, p < .05, \eta^2 = 0.02$ ) on mean PD, Smiling sentences induced smaller PD than Unsmiling sentences (Fig. 2B) (Table 2).

No main effect of Choice ( $F(1,24) = 1.93, p > .05, \eta^2 = 0.02$ ) was observed (Table 2) (Fig. 2C) and there was no significant interaction between Filter and Choice ( $F(1,24) = 0.48, p > .05, \eta^2 = 0.004$ ) on mean pupil dilation (Fig. 2D) (Table 2).

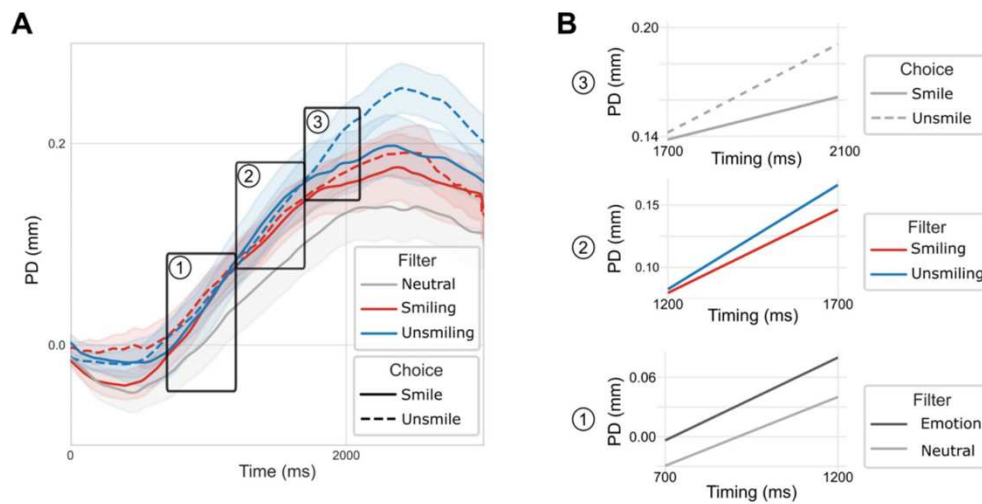
A repeated-measures ANOVA was performed to analyze the effects of Choice (Smile vs. Unsmile), Response (Correct vs. Incorrect) and the interaction between these two factors on mean PD in the same time window (2100–2600ms). No main effects were observed either for the Response ( $F(1,24) = 0.48, p > .1, \eta^2 = 0.004$ ) or Choice ( $F(1,24) = 1.93, p > .1, \eta^2 = 0.02$ ). However, a significant interaction between these two factors was detected ( $F(1,24) = 4.30, p < .01, \eta^2 = 0.02$ ). In order to inspect relevant effects, only comparisons within Correct and Incorrect responses were performed with a Bonferroni correction. This analysis revealed a significant difference between Correct responses as a function of Choice ( $t(24) = 2.45, p_{\text{corr}} < 0.05$ ) (Fig. 2D). No difference was observed between Incorrect responses ( $t(24) = 0.097, p_{\text{corr}} > 0.05$ ).

**Slopes analysis**

Besides analyzing the mean amplitude, a slope analysis was conducted to assess the kinetics of pupil dilation in response to Emotion, Filter, and Choice across different timing conditions Fig. 3A). A difference was observed between Emotional (combined Smiling and Unsmiling) and Neutral sound in the 700–1200ms time window ( $\beta = -2.8e^{-5}$  with Emotional filter as the reference level;  $SE = 1.3e^{-6}$ ;  $p < .001$ ), with the pupil dilation slope



**Fig. 2.** Effect of Filter and Choice on pupil activity. A: Effect of emotional filters on pupil dilation B: Effect of Filter on pupil dilation C: Effect of Choice on pupil dilation (without neutral sentences); D: Effect of Filter, Choice and Response on pupil dilation (neutral sentences excluded). Shaded areas represent standard error of the mean. Barplots represent mean pupil dilation in the window 2100-2600ms (mean  $\pm$  standard error of the mean) (in grey in time series). PD: Pupil Dilation\*  $p < .05$ .



**Fig. 3.** Effect of Filter and Choice on pupil dilation kinetic. A: Effect of Filter and Choice on pupil dilation with Neutral sentences (no matter the judgement for them) with time windows chosen for slopes analysis. B: 1- Effect of emotional content; 2- Effect of Filter; 3- Effect of Choice. PD: Pupil Dilation.

being sharper for Emotional than for Neutral sounds (Fig. 3B-1). Note that in this early time window no difference between the slope of Smiling and Unsmiling Filter was revealed. In the next time window (1200-1700ms) a difference in slope was shown as a function of Filter ( $\beta = 3.4e^{-5}$  with Smiling filter as the reference level;  $SE = 6.8e^{-6}$ ;  $p < .0001$ ), with sharper pupil dilation for Unsmiling sounds than Smiling sounds (Fig. 3B-2). Finally, in the last time window (1700-2100ms), an effect of Choice on pupil dilation slope was observed ( $\beta = 6.3e^{-5}$  with Smile choice as the reference level;  $SE = 2.0e^{-6}$ ;  $p < .0001$ ), with an increasing curve for Unsmile but not for Smile rated sounds (Fig. 3B-3).

## Discussion

To our knowledge, this study was the first to combine, in the same participants, facial EMG (fEMG) and pupillometry in response to vocal smile. The combination of these physiological measures provided a picture of automatic motor resonance/reactivity at different processing stages through different systems (ANS and facial muscular activity) in response to artificially prosodically modified sentences. The combined analysis of these indices revealed that the emotional state is triggered early and thus, is independent of facial motor contagion. This is evidenced by the fact that autonomic reactivity was sharper to emotional stimuli than to neutral stimuli and preceded facial muscular reactivity. Our paradigm also allowed us to examine more rigorously the facilitating influence of explicit judgement on both muscular and autonomic nervous system activity, suggesting an appraisal and a reinforcement of the already implicitly processed and automatically expressed emotion. Upcoming studies should also investigate these processes to ensure the results repeatability considering our small, though 'reasonable' sample size.

Indeed, as in Arias et al. (2018), our results on facial muscular activity made it possible to distinguish between implicit and explicit motor resonance to vocal smile. The *Zygomaticus major* (ZM) reacted to smiling even in the absence of congruent choice, indicating only implicit integration of prosody, whereas *Corrugator supercilii* (CS) activity reflected explicit judgement. The positive correlation between different muscular activities suggests that perception and identification of vocal smile induced congruent implicit and explicit motor resonance. However, even if CS activity reflects the judgement of vocal emotion, no correlation was observed with either accuracy or correct/incorrect responses: correct recognition of emotional prosody was not related to a larger explicit motor resonance.

Consistently, it has been shown that even if expressing facial emotion can help to categorize other's facial<sup>55</sup> and vocal emotions<sup>56</sup>, it is not mandatory to recognize emotions. Patients with Moebius syndrome, characterized by congenital facial paralysis, performed similarly to a control group in a facial expression recognition task<sup>57-59</sup>.

Non-social sounds have been shown to induce congruent emotional responses, such as the contraction of ZM muscle in response to a melody rated as 'happy' and the contraction of CS muscle in response to a melody rated as 'sad'<sup>60</sup>. In the present study, the congruent activity of ZM muscle in response to vocal smile, even when not identified as such (as in Arias et al. studies<sup>33,34</sup>), suggests an automatic emotional facial resonance distinct from music-induced emotional response alone.

The analysis of autonomic reactivity allowed further investigation of emotional contagion. Pupil automatic reactivity would reflect the modulation of emotional state following the perception of a vocal smile. First, pupil reactivity was observed for both emotional and neutral sentences. This systematic dilation reflected low-level sensory processing, cognitive load, and sentence evaluation<sup>61-64</sup>. Second, the artificially modified sentences were sufficiently emotional to induce a larger dilation than unmodified neutral sounds in an early time window, an effect similar to the that observed with natural emotion<sup>46,65</sup>. Third, in a later time window, sound valence influenced pupil reactivity, leading to a larger dilation for negative sentences than for positive sentences. This result has been widely replicated by studies that have focused on the physiological response to sounds<sup>39,53,62</sup>, but only with very salient emotional stimuli. The stimuli included in previous research are highly attention-grabbing, which might lead to confounding effects that prevent purely emotional aspects from being explored. Our results extend these observations to more subtle stimuli: the sentences used conveyed subtle emotions allowing for finer recognition of acoustic cues and thus for reproducing an effect which is closer to that observed during a real communicative situation. Altogether, our findings point towards the existence of emotional contagion to subtle vocal smile. In the last time window, pupil dilation continues for recognized Unsmiling sentences, suggesting that explicit appraisal modulates pupil reactivity for negative emotion. This increasingly detailed analysis of the emotional content of sentences reflected in the activity of the pupil attests that the pupil is a valuable window into the emotional integration processes at the ANS level.

A recent study demonstrated that ZM muscle innervation includes 10% of ANS fibers (vs. 3.9% for CS)<sup>66</sup>. Thus, in our study, implicit ANS response to sounds the emotional content could also be reflected in ZM reactivity. These findings are consistent with the model of Neurological Mechanisms of Emotional Contagion (NMEC) proposed by Prochazkova: in response to a visual emotion, a 'sender' emotion contagion is observed through the ANS on the 'receiver'<sup>43</sup>. This model, according to results of the present study, could thus be generalized to auditory emotion.

The present study time course contributes to enriching theories of emotions by providing information on the feedback loop between perception of emotions, facial expressions, and their experience. This emotional loop raises several questions concerning the historic debate between peripheralist<sup>8</sup> and central theories of emotion<sup>9</sup>. Damasio studied the effect of emotional somatic markers on decision-making<sup>3,67</sup>. Regarding somatic response, emotional motor resonance and pupil reactivity may reinforce each other. This idea is supported by the correlation observed between ZM and CS muscle activities, suggesting a reinforcement of facial muscles. The activation of one of the effectors might thus trigger a congruent response of the whole face. Moreover, although the changes in the pupil cannot be intrinsically perceived, they contribute to global physiological emotional response. Even if there is no feedback from the effect of the smooth muscle opening the iris, this leads to a change in the amount of light and information entering the visual system and processed automatically which

contributes in its turn to changes in the emotional state. Note that this proposal should be interpreted with caution, as no correlation has been reported between muscular and pupil reactivity.

The current study offered the rare opportunity to jointly observe emotional motor resonance and emotional state modulation via the use of different psychophysiological measures. To go deeper into studying these two phenomena, questioning participants about their subjective feelings towards the filters that have been applied could provide further information on emotional contagion. Within the framework of the polyvagal theory<sup>68,69</sup>, the addition of other ANS measures such as heart rate would provide information about the emotion valence of smiling voices, especially their prosocial valence. Finally, to confirm that implicit effects were the consequence of prosodic modulations rather than artificial characteristics, natural smiling and unsmiling sentences could be added.

### Conclusion

This paradigm used in the present study allowed us to control the intensity and acoustic characteristics of emotional prosody content and thus describe emotional contagion through two different physiological indices.

Emotional contagion is the last process of the perception – representation – action loop<sup>70</sup>. Congruent response to visual smile associated with mimicry in the visual domain. This study disentangles emotional states and motor resonance, together with implicit and explicit aspects of emotional contagion through the use of auditory emotional stimuli and the combination of autonomic and facial motor responses. Mimicry of emotions like laughing and smiling is known to be altered in individuals with disorders such as autism<sup>71,72</sup>. The present paradigm, which includes measures of motor and autonomic reactivity to emotional vocal sounds, could be used to investigate clinical populations. In particular, the artificial modification of sounds through the application of a smiling filter could be pushed one step further: sounds could be artificially modified on the basis of individual models of vocal smile. This would allow measurement of emotional contagion to one's own vocal smile, opening the way to the investigation of the individual variability that characterizes many disorders.

### Methods

#### Population

Twenty-five young adults aged 20 to 30 years (mean  $24.0 \pm 2.5$  years) participated in the study (12 females); all were native French speakers. This sample size was chosen according to other fEMG (Vrana et al.,  $n=19$ , Rymarczyk et al.,  $n=30$  and Sato et al.,  $n=29$ )<sup>73–75</sup> and pupillometry (Bufo et al.,  $n=30$ , Legris et al.,  $n=17$ , Ricou et al.,  $n=30$ )<sup>48,76,77</sup> studies' sample size. An *a posteriori* power analysis was performed (see supplementary materials). To be included in this study, participants should not have (history of) neurological, psychiatric and metabolic disorder, nor be under medication at the time of the study. Each participant completed the Empathy Quotient questionnaire (EQ)<sup>78</sup>, which provides an estimate of empathic abilities.

#### Ethics statement

The study was approved by *Comité de Protection des Personnes* CPP Est I (PROCEA2017/23; ID RCB: 2017-A00756-47). All methods in this study were carried out in accordance with the relevant guidelines and regulations, and all data in this study were obtained with informed consent from all subjects and/or their legal guardian(s).

#### Stimuli and experimental design

##### Stimuli and sequence of stimulation

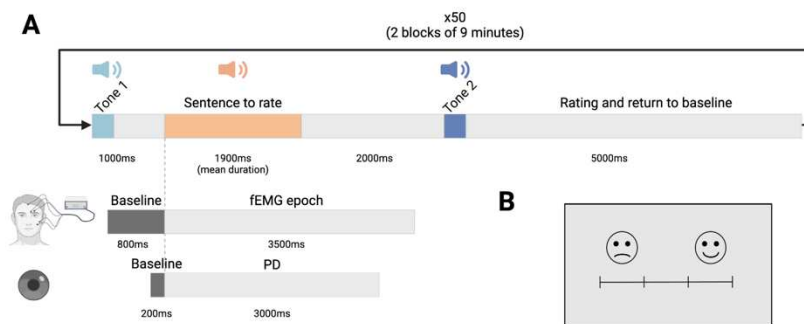
The same corpus as in Arias et al.'s studies<sup>33,34</sup> was used (see Supplementary materials). Only semantically neutral sentences were selected, so as to make sure to test the effect of emotional prosody content<sup>79</sup>. The emotional prosody of the sentences was modified via an algorithm including a Smiling and an Unsmiling filter, created following the model of vocal smile proposed by Ponsot et al.<sup>25</sup>. Validating the vocal smile model in this sample was a prerequisite for the validity of modified emotional sentences.

##### Procedure

Two 300 ms tones, each with a specific frequency, were created using Praat<sup>80</sup> and normalized in energy and faded on in- and output. One tone (440 Hz) was used to signal the beginning of a new trial. The sentence was played 1000 ms later, followed by a silent period of 2000 ms, and finally by the second tone (275 Hz), which indicated the rating time and the return to baseline (lasting 5000 ms). The time course of each trial is represented in Fig. 4A. Because the research question was focused on sound processing, the image of a rating scale displayed on the screen (Fig. 4B) remained unchanged throughout the experiment; successive stages were indicated with tones. Using the same image for the entire time of the experiment allowed us to maintain a constant luminance (around 20 lx), which was a necessary condition to ensure that the observed effects on pupil dilation were a consequence of the auditory task.

The participants were comfortably sitting 70 cm away from the screen (1980 × 1080 px) and the speakers. No chinrest was used. The intensity of the sounds was controlled and maintained between 58- and 62-dB SPL.

The following instruction was given to participants: 'You will hear sentences, and you will be asked to rate how joyfully or unjoyfully the sentence was pronounced. You should not judge the content of the sentence but the way it was pronounced'. For the behavioral data associated to facial EMG and pupil recording, judgements were recorded through a 4-key response pad. The displayed image represented the rating scale: the two right keys corresponded to two levels of Smile intensity and the two left keys corresponded to two levels of Unsmile intensity.



**Fig. 4.** A: Timeline of a trial in the experimental sequence. Tone 1 indicated the beginning of a new trial, tone 2 indicated the time when the subject should rate the sentence. B: Rating scale displayed on screen throughout the duration of the task; PD: Pupil Dilation; fEMG : facial ElectroMyoGraphy.

### Measurements

#### Facial EMG (fEMG) - recording

Facial EMG (fEMG) was recorded with 4-mm Ag-AgCl electrodes EL254 and EL254S, the BIOPAC MP36 acquisition system (BIOPAC® Systems Inc. Goleta, CA, biopac.com) and AcqKnowledge® 4.1 software (biopac.com). In order to measure the activity of the two facial muscles *Zygomaticus major* (ZM) and *Corrugator supercilii* (CS) and following Fridlund and Cacioppo's guidelines<sup>81</sup>, the electrodes were placed on the left side of the participant's face after scrubbing with an abrasive paste to remove traces of moisturizer, makeup or perspiration. Data were recorded with a sampling rate of 1024 Hz. For each muscle, two electrodes separated by approximately 1 cm were used. This allowed us to measure the potential difference between the anode and cathode, placed in the same position on each participant's face. A reference electrode was added to the center of the forehead at the hairline.

#### Pupillometry - recording

Pupil data were recorded using a SMI RED 500® system, synchronized with BIOPAC MP36®, at a 500 Hz sampling rate using an infrared light ( $\lambda = 870$  nm) to detect and measure pupil dilation (PD) in millimeters (mm)<sup>82</sup>. Data were recorded only after a correct 5-point calibration with a deviation degree of less than 1° (mean: 0.59°/0.58°).

### Data analysis

#### Behavioral data analysis

For all analysis, the acoustical modification of the sentences is referred to as Filter (Smiling vs. Unsmiling) and the rating of the degree of smiling in the sentences is referred to as Choice (Smile vs. Unsmile). The congruence between Filter and Choice is referred to as the Response, which can be Correct (Smiling rated as Smile and Unsmiling rated as Unsmile) or Incorrect (Smiling rated as Unsmile and Unsmiling rated as Smile).

Accuracy was calculated as the number of Correct responses divided by the sum of all Smiling and Unsmiling sentences (80 for each participant), to reflect the good rating of both filters (Eq. 1). Accuracy was compared to chance level (50%) to evaluate the recognition of emotional prosodic modulation.

$$Accuracy = \frac{nSmiling_{smile} + nUnsmiling_{unsmile}}{nSmiling + nUnsmiling} \quad (1)$$

Equation Formulae Accuracy formula;  $nSmiling_{smile}$ : number of trials for smiling sentences rated as smile;  $nUnsmiling_{unsmile}$ : number of trials for Unsmiling sentences rated as unsmile;  $nSmiling$ : total of trials for the Smiling Filter;  $nUnsmiling$ : total of trials for the Unsmiling Filter.

#### fEMG - processing

Analysis and pre-processing of fEMG data were performed with MNE-toolbox on Python 3.9.4<sup>83</sup>. Data were filtered with a 50 Hz IIR high-pass filter and a 250 Hz low-pass filter. Subsequently, significant muscular activity artifacts were manually removed after visual inspection of the recordings. Data were then segmented into epochs, with an 800 ms pre-stimulus baseline and a duration of 3500 ms, starting from the beginning of each sentence. The absolute value of the muscular activity was smoothed and averaged using a 300 ms sliding window, followed by z-score normalization based on the baseline of the current trial. On average,  $20 \pm 11\%$  (mean  $\pm$  sd) of trials per participant were removed due to excessive motion artifacts. One participant was excluded from the fEMG group analysis because of excessive noise in the signal, but data from this participant were retained for pupil and behavioral analysis.

#### Pupillometry – processing

Pre-processing and processing of raw pupil data were performed using in-house MATLAB (R2015b; MathWorks) scripts. The first step involved detecting and interpolating artifacts caused by blinks and brief signal losses. Blink detection was achieved using a velocity-based algorithm<sup>83,84</sup> supplemented by manual detection. Interpolation for missing values was performed with a cubic interpolation with the median value. For each trial, a 200 ms baseline before the beginning of a sentence playing was extracted. Pupil Dilation (PD) was calculated as pupil diameter increase in mm in response to stimulus presentation and choice relative to the baseline of the trial. Pupillary response was measured during 3000 ms after the start of sentence playback. Each trial was visually inspected, and  $34 \pm 22\%$  (mean  $\pm$  sd) of trials per participant were removed due to excessive artifacts.

#### Statistical analysis

Statistical analyses were conducted with Rstudio 4.2.2<sup>85,86</sup> with the packages ggplot2<sup>87</sup>, ez<sup>88</sup>, tidyverse<sup>89</sup> and dplyr<sup>90</sup>.

For each physiological measurement (i.e., muscle activity and pupil dilation), the effects of different factors (Filter, Choice, Response) were estimated through the mean amplitude of electrophysiological measures within selected time windows. To assess the differences between conditions in a psychophysiological relevant time window, randomizations ( $n = 10,000$ ) were performed across conditions (Smiling vs. Unsmiling and Smile vs. Unsmile)<sup>91</sup> with a Guthrie-Buchwald correction<sup>92</sup>. For the fEMG data randomizations, a down-sampling to 512 Hz was applied to the signal using the MNE-toolbox<sup>93</sup>.

Within the randomizations, a common 500 ms window was selected, in the significant time period, in which average amplitude was measured for different conditions in each physiological measure. To identify the significant effect in the selected time window, a repeated-measures ANOVA (Filter x Choice) was performed on mean muscular activity or PD amplitude. To assess the effect of Response and interaction with Choice, a repeated-measures ANOVA of the mean amplitude of muscle activity and PD was performed with post-hoc pairwise comparisons with Bonferroni correction to further specify significant interaction effects within Correct and Incorrect Response.

In addition to the mean amplitude analysis, an analysis of slopes was performed in order to evaluate the kinetics of pupil dilation in response to Emotion, Filter and Choice as a function of Timing. A linear mixed model was performed with Emotion, Filter and/or Choice and Timing as fixed factors and with subjects as random factor and Emotion, Filter and/or Choice as random slopes with the lmer R library<sup>94</sup>. This analysis was performed by taking into account different time windows after signal visual inspection: 700 to 1200 ms for Emotion, 1200 to 1700 ms for Filter and 1700 to 2100 ms for Choice effect.

Finally, a Pearson correlation matrix was computed between physiological measures, rating responses, and EQ scores to explore the relationships among all measures. A Bonferroni correction was applied for multiple comparisons. Results are presented in Supplementary materials (Figure S5).

#### Data availability

Code and data for processing and statistical analysis are publicly available in an OSF project (<https://osf.io/6rus4/>).

Received: 17 May 2024; Accepted: 30 September 2024

Published online: 23 October 2024

#### References

- Hatfield, E., Cacioppo, J. T. & Rapson, R. L. *Emotional Contagion* (Cambridge University Press, 1993). <https://doi.org/10.1017/CBO9781139174138>
- Premack, D. & Woodruff, G. Does the chimpanzee have a theory of mind? *ISTEX*. **1**, 515–526 (1978).
- Damasio, A. R. & Blanc, M. *L'erreur De Descartes: La Raison Des  motions* (O. Jacob, 1995).
- Carsten, T., Desmet, C., Krebs, R. M. & Brass, M. Pupillary contagion is independent of the emotional expression of the face. *Emot. Wash. DC*. **19**, 1343–1352 (2019).
- Izard, C. E. *The Face of Emotion*. xii, 468 (Appleton-Century-Crofts, East Norwalk, CT, US, (1971).
- Coles, N. A. et al. A multi-lab test of the facial feedback hypothesis by the many smiles collaboration. *Nat. Hum. Behav.* **6**, 1731–1742 (2022).
- Soussignan, R. Regulatory function of facial actions in emotion processes. In *Advances in psychology research*, Vol. 31, 173–198. (Nova Science Publishers, Hauppauge, NY, US, 2004).
- James, W. What is an emotion? *Mind*. **9**, 188–205 (1884).
- Cannon, W. B. The James-Lange theory of emotions: A critical examination and an alternative theory. *Am. J. Psychol.* **39**, 106–124 (1927).
- Scherer, K. R. Appraisal considered as a process of multilevel sequential checking. In: *Appraisal processes in emotion: Theory, methods, research*. New York, NY, US: Oxford University Press; 2001. p. 92–120. (Series in affective science).
- Scherer, K. R. Emotions are emergent processes: they require a dynamic computational architecture. *Philos. Trans. R Soc. B Biol. Sci.* **364**, 3459–3474 (2009).
- Hess, U. & Blairy, S. Facial mimicry and emotional contagion to dynamic emotional facial expressions and their influence on decoding accuracy. *Int. J. Psychophysiol.* **40**, 129–141 (2001).
- Dimberg, U., Thunberg, M. & Elmehed, K. Unconscious facial reactions to emotional facial expressions. *Psychol. Sci.* **11**, 86–89 (2000).
- Simpson, E. A., Murray, L., Paukner, A. & Ferrari, P. F. The mirror neuron system as revealed through neonatal imitation: Presence from birth, predictive power and evidence of plasticity. *Philos. Trans. R. Soc.* **369**, 20130289 (2014).
- Burrows, A. M., Waller, B. M., Parr, L. A. & Bonar, C. J. Muscles of facial expression in the chimpanzee (Pan troglodytes): Descriptive, comparative and phylogenetic contexts. *J. Anat.* **208**, 153–167 (2006).
- Olszanowski, M., Wr bel, M. & Hess, U. Mimicking and sharing emotions: A re-examination of the link between facial mimicry and emotional contagion. *Cogn. Emot.* **34**, 367–376 (2020).
- Rychlowska, M. et al. Functional smiles: tools for love, sympathy, and war. *Psychol. Sci.* **28**, 1259–1270 (2017).

18. Shore, D. M. & Heerey, E. A. The value of genuine and polite smiles. *Emotion*. **11**, 169–174 (2011).
19. Hess, U. & Bourgeois, P. You smile–I smile: Emotion expression in social interaction. *Biol. Psychol.* **84**, 514–520 (2010).
20. Wood, A., Rychlowska, M., Korb, S. & Niedenthal, P. Fashioning the face: Sensorimotor simulation contributes to facial expression recognition. *Trends Cogn. Sci.* **20**, 227–240 (2016).
21. Ohala, J. J. The acoustic origin of the smile. *J. Acoust. Soc. Am.* **68**, S33–S33 (1980).
22. Aubergé, V. & Cathiard, M. Can we hear the prosody of smile? *Speech Commun.* **40**, 87–97 (2003).
23. Tarter, V. C. Happy talk: Perceptual and acoustic effects of smiling on speech. *Percept. Psychophys.* **27**, 24–27 (1980).
24. Tarter, V. C. & Braun, D. Hearing smiles and frowns in normal and whisper registers. *J. Acoust. Soc. Am.* **96**, 2101–2107 (1994).
25. Ponsot, E., Arias, P. & Aucouturier, J. J. Uncovering mental representations of smiled speech using reverse correlation. *J. Acoust. Soc. Am.* **143**, EL19–EL24 (2018).
26. Bradley, M. M. & Lang, P. J. The international affective digitized sounds: affective ratings of sounds and instruction manual (Technical Report No. B-3). University of Florida. *NIMH Cent. Study Emot. Atten. Gainesv. FL* **30** (2007).
27. Hietanen, J. K., Surakka, V. & Linnankoski, I. Facial electromyographic responses to vocal affect expressions. *Psychophysiology*. **35**, 530–536 (1998).
28. Larsen, J. T., Norris, C. J. & Cacioppo, J. T. Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*. **40**, 776–785 (2003).
29. Verona, E., Patrick, C. J., Curtin, J. J., Bradley, M. M. & Lang, P. J. Psychopathy and physiological response to emotionally evocative sounds. *J. Abnorm. Psychol.* **113**, 99–108 (2004).
30. Dimberg, U., Thunberg, M. & Grunedal, S. Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cogn. Emot.* **16**, 449–471 (2002).
31. Dimberg, U. & Thunberg, M. Rapid facial reactions to emotional facial expressions. *Scand. J. Psychol.* **39**, 39–45 (1998).
32. Lindström, R. et al. Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders. *Clin. Neurophysiol.* **129**, 2411–2420 (2018).
33. Arias, P., Belin, P. & Aucouturier, J. J. Auditory smiles trigger unconscious facial imitation. *Curr. Biol.* **28**, R782–R783 (2018).
34. Arias, P., Bellmann, C. & Aucouturier, J. J. Facial mimicry in the congenitally blind. *Curr. Biol.* **31**, R1112–R1114 (2021).
35. Babiker, A., Faye, I. & Malik, A. Pupillary behavior in positive and negative emotions. In *IEEE International Conference on Signal and Image Processing Applications* 379–383 (IEEE, Melaka, Malaysia, 2013). <https://doi.org/10.1109/ICSIPA.2013.6708037>
36. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function. Adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
37. Bradley, M. M., Miccoli, L., Escrig, M. A. & Lang, P. J. The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*. **45**, 602–607 (2008).
38. Aguilón-Hernández, N. et al. The pupil: A window on social automatic processing in autism spectrum disorder children. *J. Child. Psychol. Psychiatry* **61**, 768–778 (2020).
39. Babiker, A., Faye, I., Prehn, K. & Malik, A. Machine Learning to differentiate between positive and negative emotions using pupil diameter. *Front. Psychol.* **6**, (2015).
40. Beatty, J. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull.* **91**, 276–292 (1982).
41. Hess, E. H. & Polt, J. M. Pupil size in relation to mental activity during simple problem-solving. *Science* **143**, 1190–1192 (1964).
42. Samuels, E. R. & Szabadi, E. Functional neuroanatomy of the noradrenergic locus Coeruleus: Its roles in the regulation of Arousal and autonomic function part II: Physiological and pharmacological manipulations and pathological alterations of Locus Coeruleus activity in humans. *Curr. Neuropharmacol.* **6**, 254–285 (2008).
43. Prochazkova, E. & Kret, M. E. Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neurosci. Biobehav. Rev.* **80**, 99–114 (2017).
44. Bradley, M. M., Codispoti, M., Cuthbert, B. N. & Lang, P. J. Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion* **1**, 276–298 (2001).
45. Geangu, E., Hauf, P., Bhardwaj, R. & Bentz, W. Infant pupil diameter changes in response to others' positive and negative emotions. *PLoS ONE*. **6**, e27132 (2011).
46. Burley, D. T. & Daughters, K. The effect of oxytocin on pupil response to naturalistic dynamic facial expressions. *Horm. Behav.* **125**, 104837 (2020).
47. Aktar, E., Rajmakers, M. E. J. & Kret, M. E. Pupil mimicry in infants and parents. *Cogn. Emot.* **34**, 1160–1170 (2020).
48. Legris, E. et al. Relationship between behavioral and objective measures of sound intensity in normal-hearing listeners and hearing-aid users: A pilot study. *Brain Sci.* **12**, 392 (2022).
49. Cherng, Y. G., Baird, T., Chen, J. T. & Wang, C. A. Background luminance effects on pupil size associated with emotion and saccade preparation. *Sci. Rep.* **10**, 15718 (2020).
50. Zekveld, A. A., Koelewijn, T. & Kramer, S. E. The pupil dilation response to auditory stimuli: Current state of knowledge. *Trends Hear.* **22**, 233121651877717 (2018).
51. Cosme, G. et al. Pupil dilation reflects the authenticity of received nonverbal vocalizations. *Sci. Rep.* **11**, 1–14 (2021).
52. Jin, A. B., Steding, L. H. & Webb, A. K. Reduced emotional and cardiovascular reactivity to emotionally evocative stimuli in major depressive disorder. *Int. J. Psychophysiol.* **97**, 66–74 (2015).
53. Partala, T. & Surakka, V. Pupil size variation as an indication of affective processing. *Int. J. Hum. Comput. Stud.* **59**, 185–198 (2003).
54. Jürgens, R., Fischer, J. & Schacht, A. Hot speech and exploding bombs: Autonomic arousal during emotion classification of prosodic utterances and affective sounds. *Front. Psychol.* **9**, (2018).
55. Lewis, M. B. The interactions between botulinum-toxin-based facial treatments and embodied emotions. *Sci. Rep.* **8**, 14720 (2018).
56. Vilaverde, R. F. et al. Inhibiting orofacial mimicry affects authenticity perception in vocal emotions. *Emotion*. **24** (6), 1376–1385 (2024).
57. Calder, A. J., Keane, J., Cole, J., Campbell, R. & Young, A. W. Facial expression recognition by people with mobius syndrome. *Cogn. Neuropsychol.* **17**, 73–87 (2000).
58. Rives Bogart, K. & Matsumoto, D. Facial mimicry is not necessary to recognize emotion: Facial expression recognition by people with Moebius syndrome. *Soc. Neurosci.* **5**, 241–251 (2010).
59. Bate, S., Cook, S. J., Mole, J. & Cole, J. First report of generalized face processing difficulties in möbius sequence. *PLoS One*. **8**, e62656 (2013).
60. Bullack, A., Büdenbender, N., Roden, I. & Kreutz, G. Psychophysiological responses to happy and sad MusicA replication study. *Music Percept.* **35**, 502–517 (2018).
61. Ferencova, N., Visnovcova, Z., Bona Olexova, L. & Tonhajzerova, I. Eye pupil – a window into central autonomic regulation via emotional/cognitive processing. *Physiol. Res.* **70**, S669–S682 (2021).
62. Oliva, M. & Anikin, A. Pupil dilation reflects the time course of emotion recognition in human vocalizations. *Sci. Rep.* **8**, 4871 (2018).
63. van der Wel, P. & van Steenbergen, H. Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychon. Bull. Rev.* **25**, 2005–2015 (2018).
64. Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J. & Kramer, S. E. The eye as a window to the listening brain: neural correlates of pupil size as a measure of cognitive listening load. *NeuroImage*. **101**, 76–86 (2014).

65. Kuchinke, L., Schneider, D., Kotz, S. A. & Jacobs, A. M. Spontaneous but not explicit processing of positive sentences impaired in Asperger's syndrome: Pupillometric evidence. *Neuropsychologia* **49**, 331–338 (2011).
66. Tereshenko, V. et al. Axonal mapping of the motor cranial nerves. *Front. Neuroanat.* **17**, (2023).
67. Soussignan, R. Duchenne smile, emotional experience, and autonomic reactivity: A test of the facial feedback hypothesis. *Emot. Wash. DC* **2**, 52–74 (2002).
68. Porges, S. W. The polyvagal theory: New insights into adaptive reactions of the autonomic nervous system. *Cleve Clin. J. Med.* **76**, S86–S90 (2009).
69. Sorinas, J., Ferrández, J. M. & Fernandez, E. Brain and body emotional responses: Multimodal approximation for valence classification. *Sensors* **20**, 313 (2020).
70. Chartrand, T. L. & Bargh, J. A. The chameleon effect: The perception–behavior link and social interaction. *J. Personal. Soc. Psychol.* **76**, 893–910 (1999).
71. Beall, P. M., Moody, E. J., McIntosh, D. N., Hepburn, S. L. & Reed, C. L. Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. *J. Exp. Child. Psychol.* **101**, 206–223 (2008).
72. McIntosh, D. N., Reichmann-Decker, A., Winkielman, P. & Wilbarger, J. L. When the social mirror breaks: Deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Dev. Sci.* **9**, 295–302 (2006).
73. Vrana, S. R. & Gross, D. Reactions to facial expressions: Effects of social context and speech anxiety on responses to neutral, anger, and joy expressions. *Biol. Psychol.* **66**, 63–78 (2004).
74. Rymarczyk, K., Biele, C., Grabowska, A. & Majczynski, H. EMG activity in response to static and dynamic facial expressions. *Int. J. Psychophysiol.* **79**, 330–333 (2011).
75. Sato, W., Fujimura, T. & Suzuki, N. Enhanced facial EMG activity in response to dynamic facial expressions. *Int. J. Psychophysiol.* **70**, 70–74 (2008).
76. Bufo, M. R. et al. Autonomic tone in children and adults: Pupillary, electrodermal and cardiac activity at rest. *Int. J. Psychophysiol.* **180**, 68–78 (2022).
77. Ricou, C., Rabadan, V., Mofid, Y., Aguillon-Hernandez, N. & Wardak, C. Pupil dilation reflects the social and motion content of faces. *Soc. Cognit. Affect. Neurosci.* **19**, nsae055 (2024).
78. Baron-Cohen, S. & Wheelwright, S. The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* **34**, 163–175 (2004).
79. Russ, J. B., Gur, R. C. & Bilker, W. B. Validation of affective and neutral sentence content for prosodic testing. *Behav. Res. Methods.* **40**, 935–939 (2008).
80. Boersma, P. Praat, a system for doing phonetics by computer. *Glott Int.* **5**, 341–345 (2002).
81. Fridlund, A. J. & Cacioppo, J. T. Guidelines for human electromyographic research. *Psychophysiology*. **23**, 567–589 (1986).
82. Steinhauer, S. R., Bradley, M. M., Siegle, G. J., Roeklein, K. A. & Dix, A. Publication guidelines and recommendations for pupillary measurement in psychophysiological studies. *Psychophysiology*. **59**, e14035 (2022).
83. Kret, M. E. & Sjak-Shie, E. E. Preprocessing pupil size data: Guidelines and code. *Behav. Res. Methods* **51**, 1336–1342 (2019).
84. Nyström, M. & Holmqvist, K. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behav. Res. Methods.* **42**, 188–204 (2010).
85. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2022).
86. Team, R. Boston, MA. RStudio: Integrated Development for R. in *RStudio, PBC* (2020).
87. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
88. Lawrence, M. A. *Ez: Easy Analysis and Visualization of Factorial Experiments*. (2016).
89. Wickham, H. et al. Welcome to the Tidyverse. *J. Open. Source Softw.* **4**, 1686 (2019).
90. Wickham, H., François, R., Henry, L. & Müller, K. *Dplyr: A Grammar of Data Manipulation*. (2021).
91. Voeten, C. C. *Permutest: Permutation Tests for Time Series Data*. (2022).
92. Guthrie, D. & Buchwald, J. S. Significance testing of difference potentials. *Psychophysiology*. **28**, 240–244 (1991).
93. Gramfort, A. et al. MNE software for processing MEG and EEG data. *NeuroImage*. **86**, 446–460 (2014).
94. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: tests in Linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).

### Acknowledgements

We thank all the volunteers for their time and participation in this study. We would like to thank Dennis Wobrock for his help and technical support in setting up the protocol and processing the data. We would like to thank Marta Manenti for her help in proofreading and correcting the manuscript.

### Author contributions

A.M., N.A.-H., C.W., J.-J.A. and M.G. designed the study. A.M. and Z.R. performed data acquisition. A.M., Z.R., C.W., J.-J.A. and M.G. were responsible for data and statistical analyses. A.M., Z.R., C.W. and M.G. wrote the first version of the manuscript. All authors were involved in preparing and reviewing the manuscript.

### Funding

This work was supported by the INSERM (National Institute for Health and Medical Research), the ANR (ANR-19-CE37-0022-01), the Region CVL and the FEANS (Fondation Européenne pour l'Avancement des Neurosciences). The funding sources were not involved in the study design, data collection and analysis, or the writing of the report. All authors declare that they have no conflicts of interest.

### Declarations

#### Competing interests

The authors declare no competing interests.

#### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-74848-w>.

**Correspondence** and requests for materials should be addressed to M.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024



### 3.3. Étude n°5 : Résonance motrice au sourire vocal chez l'adulte autiste

#### **Autism : when my face is not listening to your smile**

*Merchie, A., Ranty, Z., Wardak, C., Aguillon-Hernandez, N., Bonnet-Brilhault, F., Houy-Durand, E., Aucoeur, J-J., & Gomot, M.*

*Article soumis à Nature Human Behavior*

**Introduction et objectifs :** Dans l'autisme, une absence de réponse au sourire sociale a été largement décrite cliniquement et expérimentalement. Dans l'étude n°4, il a été montré que les représentations perceptives du sourire vocal étaient similaires entre une partie des adultes autistes et neurotypiques. Cependant, aucune étude n'a cherché à mesurer la réponse au sourire vocal dans l'autisme en relation avec les représentations perceptives de celui-ci. L'objectif de cette étude était d'évaluer si, en réponse au sourire vocal, une résonance motrice était observée dans l'autisme, en lien avec les représentations perceptives.

**Matériel et méthodes :** L'activité musculaire faciale (*Zygomaticus major*, ZM, et *Corrugator supercilii*, CS) a été mesurée chez 20 adultes autistes et 20 adultes neurotypiques en réponse à des phrases rendues artificiellement souriantes ou non-souriantes qui devaient être jugées selon leur « degré de sourire ». La représentation perceptive du sourire vocal a été estimée pour chaque participant autiste et neurotypique à l'aide de la même tâche de Corrélation Inverse que dans l'étude n°3.

**Résultats :** Chez les adultes neurotypiques, les résultats concordent avec ceux de l'étude n°4, montrant un effet implicite du filtre sur le ZM et un effet explicite du jugement sur le CS. Chez les personnes autistes, malgré un modèle du sourire vocal proche de celui des adultes neurotypiques et une bonne reconnaissance des émotions présentées, aucune résonance motrice n'a été mise en lumière. Chez les adultes autistes, un effet de l'écoute de la phrase a pris le pas sur une potentielle résonance motrice liée à l'émotion perçue ou au jugement émis.

**Discussion :** Ces résultats remettent en question l'idée que des représentations atypiques mènent inévitablement à une reconnaissance émotionnelle déficiente, et mettent en avant la nécessité d'une compréhension plus approfondie et nuancée du traitement émotionnel dans l'autisme. Cette étude amène aussi la question d'atypies de la cognition incarnée dans l'autisme, avec des représentations internes qui semblent intactes mais qui ne permettent pas une activation cohérente des schémas moteurs correspondants.

**Autism : When my face is not listening to your smile**

**Author names and affiliations :**

Annabelle Merchie<sup>1,2</sup>, Zoé Ranty<sup>1,2</sup>, Claire Wardak<sup>1,2</sup>, Nadia Aguillon-Hernandez<sup>1,2</sup>,  
Frédérique Bonnet-Brilhault<sup>1,2,3</sup>, Emmanuelle Houy-Durand<sup>1,2,3</sup>, Jean-Julien Aucouturier<sup>4,5</sup>, &  
Marie Gomot<sup>1,2\*</sup>

1: Université de Tours, INSERM, Imaging Brain & Neuropsychiatry iBrain U1253, 37032,  
Tours, France

2: Excellence Center for Autism and Neurodevelopmental Disorders, EXAC·T, CHRU de  
Tours, Tours, France

3: Centre Ressource Autisme région Centre-Val de Loire, CHRU de Tours, Tours, France

4: Université de Franche-Comté, SUPMICROTECH, CNRS, institut FEMTO-ST, Besançon,  
France

5: STMS Lab IRCAM, CNRS, Sorbonne Université, Paris, France

Annabelle Merchie : [annabelle.merchie@inserm.fr](mailto:annabelle.merchie@inserm.fr)

Zoé Ranty : [zoe.ranty@univ-tours.fr](mailto:zoe.ranty@univ-tours.fr)

Claire Wardak : [claire.wardak@univ-tours.fr](mailto:claire.wardak@univ-tours.fr)

Nadia Aguillon-Hernandez : [nadia.aguillon-hernandez@univ-tours.fr](mailto:nadia.aguillon-hernandez@univ-tours.fr)

Frédérique Bonnet-Brilhault : [frederique.brilhault@univ-tours.fr](mailto:frederique.brilhault@univ-tours.fr)

Emmanuelle Houy-Durand : [E.HOUY-DURAND@chu-tours.fr](mailto:E.HOUY-DURAND@chu-tours.fr)

Jean-Julien Aucouturier : [aucouturier@gmail.com](mailto:aucouturier@gmail.com)

Marie Gomot : [gomot@univ-tours.fr](mailto:gomot@univ-tours.fr)

\* : Corresponding author: Gomot Marie: Université de Tours, Inserm, iBrain U1253, Centre  
de Pédiopsychiatrie, CHRU Bretonneau 2 Boulevard Tonnellé, 37044 Tours, Cédex 09, France

Email: [gomot@univ-tours.fr](mailto:gomot@univ-tours.fr) – Phone: [\(+33\)247278664](tel:+33247278664)

**Abstract**

The absence of response to social smile is one of the early hallmarks of autism. This study explores how autistic (n = 20) and neurotypical (n = 20) adults process and respond to vocal smile through the measurement of motor resonance. Despite autistic adults recognizing and perceptually representing vocal smiles similarly to neurotypicals, they did not show typical motor resonance when listening to emotional voice. In contrast, neurotypicals displayed facial muscle activity aligned with perceived emotions. Results suggest that while autistic individuals can cognitively recognize vocal emotions based on accurate acoustic characteristics, they may experience a disconnect in the physical embodiment of these emotions. This difference in motor resonance could influence social interactions by reducing emotional reciprocity in autism. These findings highlight the importance of developing therapeutic interventions that support emotional embodiment in autistic individuals.

**Keywords** : autism, motor resonance, emotional prosody, embodiment, vocal smile

## 1. Introduction

Atypical facial expressions, described as reduced or flat, have been repeatedly reported in autism<sup>1</sup>. Such differences in the production and perception of facial expressions can hinder the ability to convey emotions and intentions, which are crucial for social interaction. As a result, autistic individuals may not exhibit typical responses to emotional cues<sup>2</sup>, leading to misunderstandings and the perception of being less engaged or empathetic<sup>3</sup>. Similarly, their limited facial expressiveness can make it challenging to form social connections, as neurotypical individuals may find it harder to interpret the emotional states of autistic people<sup>4,5</sup>. Building on these challenges in facial expressiveness, individuals with autism may also experience difficulties in producing and processing emotional prosody and exhibit atypical responses to these vocal cues.

Atypicalities in the production and perception of emotional prosody have been extensively documented in autism, through both clinical reports and experimental studies. Prosody production is often described as monotonous or exaggerated in autism<sup>6,7</sup>, to the extent that in the ADOS (Autism Diagnostic Observation Schedule), atypical prosody is one of the hallmarks for diagnosis in the language and communication dimension<sup>8</sup>. These atypical patterns can negatively impact social interactions and hinder the ability to convey or understand emotional nuances in speech. In experimental settings, difficulties in the recognition of emotional prosody have also been reported in both autistic children<sup>9</sup> and adults<sup>10</sup>. However, some studies have reported that prosody identification can be similar to that of neurotypicals<sup>11,12</sup>, and others indicate it may simply be less accurate<sup>13,14</sup> (but see Leung et al.<sup>15</sup> for a comprehensive review on emotional prosody perception and recognition in autism).

In the visual modality, a lack of motor facial resonance (i.e., the automatic and often unconscious activation of facial muscles in response to emotional expressions in others<sup>16</sup>) has

been observed in individuals with autism when responding to various social cues, such as yawns<sup>17–19</sup> and facial expressions<sup>2,17,20–22</sup>. While neurotypical individuals demonstrate yawning contagion in response to auditory stimuli<sup>17,23</sup>, this phenomenon is absent in autistic children<sup>24</sup>, suggesting that sensory modality does not affect behavioral contagion in either neurotypical or autistic individuals. However Magnée et al. demonstrated that motor resonance was atypical for autistic adults in response to visual-only emotion while no differences were observed between autistic and neurotypical responses for audio-visual emotional stimuli<sup>21</sup>. This finding raises questions about the facilitating effect of prosodic content on motor resonance in autism.

Among social cues, smiling is recognized as one of the most effective triggers of emotional resonance<sup>25</sup>. In individuals with autism, difficulties in both the response to<sup>20</sup> and production of social smiles<sup>26</sup> have been reported in the visual modality. Smiles are also identifiable in the auditory modality (i.e., the vocal emissions produced when the face smiles, referred as vocal smile in this manuscript), characterized by a robust acoustic signature in neurotypical adults<sup>28</sup>. Given that recognition of emotional prosody in autism varies across studies, it could be hypothesized that perceptive representations of vocal emotion can differ across participants, as it has been suggested for the representation of facial expression<sup>32</sup>. However in a recent study, we found that neurotypical individuals and most autistic participants shared a similar internal representation of vocal smiles<sup>27</sup>. In neurotypical and congenitally blind individuals, motor resonance in response to vocal smiles has been observed, leading to congruent motor responses<sup>29–31</sup>. However, to date, no studies have specifically investigated facial motor resonance in response to purely emotional prosodic cues, such as vocal smiles, in autism.

Given these insights, the objective of the current study was to investigate facial motor resonance in response to vocal emotions in autistic adults. Specifically, we measured facial muscle (*Zygomaticus major* and *Corrugator supercilii*) responses to sentences artificially uttered in a

smiling or non-smiling prosody using facial electromyography (fEMG) in both autistic and neurotypical adults. As in a passive listening task no motor resonance could be observed<sup>33</sup>, participants were asked to judge the prosody of the spoken phrases. To ensure that any observed difficulties in facial motor responses could be attributed to differences in emotion perception and recognition, vocal smile perceptual representations were also estimated in both groups using a reverse correlation task<sup>28</sup>.

## 2. Results

### 2.1. Representation of vocal smile and Accuracy

Representations of vocal smile were similar between groups with analogous acoustic frequency and amplitude features (Figure 1A). A statistical difference between groups was observed for only one frequency, 533 Hz ( $t(39) = 2.2$ ;  $p < .05$ ).

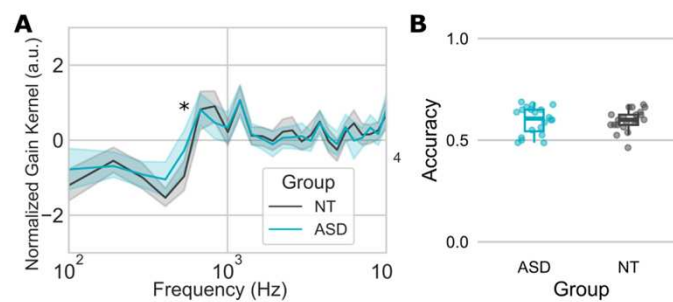


Figure 1: Representations of vocal smile from the Reverse Correlation task and Accuracy in the EMG task. A: averaged filter underlying the evaluation of vocal smile in ASD (blue) and NT (grey), as derived with reverse-correlation. Asterisks indicate significant differences between the two groups (two-tailed; paired sample t-tests;  $p < .05$ ); B: Accuracy in the EMG task in ASD (blue) and in NT (grey).

Accuracy during the EMG task was above the chance (0.5) level in both groups (NT  $t(20) = 8.83$ ;  $p < .001$ ; and ASD  $t(19) = 6.49$ ;  $p < .001$ ). There was no significant difference between the NT and ASD groups in Accuracy on the EMG task ( $t(36) = -.07$ ,  $p > .9$ ) (Figure 1B).

The correlation analysis revealed no significant correlation between the Kernel Distance and Accuracy when the NT and ASD groups were combined ( $t(39) = -.67$ ;  $p > .05$ ). Additionally,

no significant correlation was found within the individual groups, neither for the ASD group ( $t(18) = -1.08$  ;  $p > .05$ ), nor for the NT group ( $t(19) = .29$  ;  $p > .05$ ).

No significant correlations were identified between EMG accuracy and AQ in the combined sample (ASD and NT) ( $t(34) = 1.59$ ,  $p_{\text{corr}} > .05$ ), nor with EQ ( $t(34) = -0.70$ ,  $p_{\text{corr}} > .05$ ). Separate analysis in each group showed no significant correlations between EMG accuracy and AQ or EQ

In ASD, a larger distance to the model of vocal smile was observed in male than in female ( $W=16$  ;  $p < .05$ ), while no such difference was observed in NT ( $W = 36$  ;  $p > .05$ ). However, no difference according to the gender was observed for the Accuracy in the EMG task either in NT ( $W = 48$ ;  $p > .1$ ) nor in ASD ( $W = 51$ ;  $p > .1$ ).

### **2.1. Motor resonance**

#### **2.1.1. Muscle activity**

To evaluate the effect of Listening and Processing of sentences on muscular activity between Groups and within Muscles a repeated measures ANOVA was performed on mean muscular activity during Listening to sentences (0-1900 ms) and Processing (1900-4000 ms), referred as Timing in the ANOVA. The ANOVA revealed a main effect of Muscle ( $F(1,39) = 9.44$ ,  $p < .01$ ,  $\eta^2 = .07$ ) with a larger contraction of the CS than ZM, and of Timing ( $F(1,39) = 5.70$ ,  $p < .05$ ,  $\eta^2 = .02$ ) (Figure 2A-B). These effects of Timing were specified by the significant interaction between Timing and Group ( $F(1,39) = 4.27$ ,  $p < .05$ ,  $\eta^2 = .02$ ) driven by the larger relaxation of muscles in the ASD group only during the processing of sentences ( $p_{\text{corr}} < .001$ ) (Figure 2C).

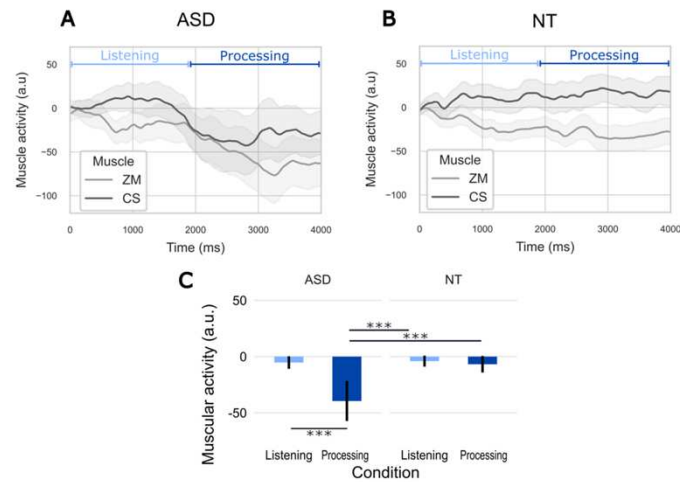


Figure 2: Muscular activity during Listening and After hearing sentences (Timing) . A: Effect of Condition on ZM and CS in ASD; B: Effect of Timing on ZM and CS in NT; Effect of Timing (C) on muscular activity in ASD and NT – mean response amplitude in selected time windows: 0-1900 ms (Listening – light blue) and 1900-4000 ms (Processing – dark blue). Shaded areas and error bars represent standard error of the mean.\*\*\*  $p < .001$

As a main group difference is observed during the second time period (Processing phase), repeated measured ANOVAs were conducted on a chosen 500 ms window (2900-3400 ms) during this phase and in accordance to previous study<sup>31</sup>.

An omnibus repeated measures ANOVA was conducted to analyze mean muscular activity, considering Group and Muscle as, respectively between-subject and within-subject, factors, in the 2900-3400 ms time window. Although muscular activity was smaller in ASD than in NT, no statistical difference was observed between groups ( $F(1,42) = 3.34, p < .1, \eta^2 = .05$ ). A significant effect of the muscle was observed, indicating larger muscular strength for CS than for ZM muscles ( $F(1,42) = 6.40; p < .001, \eta^2 = .09$ ). Based on these results, the analysis was split by muscle to evaluate effects of Filter and Choice independently. Group was used as a between-subject factor in the main ANOVAs, followed by separate analyses within each group to specify the effects of Filter and Choice independently.

### 2.1.2. Zygomaticus major activity

A repeated measures ANOVA was conducted to analyze mean muscular activity of the ZM in the 2900-3400 ms time window considering Group as between, and Filter and Choice as within factors. No main effects were found for any of the factors ( $p > .05$ ) and no interaction, but a difference for Filter almost reached significance ( $F(1,42) = 3.95$ ;  $p = .052$ ;  $\eta^2 = .01$ ).

Even though, no Group and Filter interaction was significant separate analyses within each group were performed to specify these effects considering the observed effect on time series (Figure 3A-B), the trend of the Filter effect and our previous study in NT showing an effect of Filter<sup>31</sup>. In the ASD group, no difference was found for Filter ( $F(1,19) = .35$ ;  $p > .05$ ), Choice ( $F(1,19) = .72$ ;  $p > .05$ ), or their interaction ( $F(1,19) = .41$ ;  $p > .05$ ) (Figure 3A-C). In the NT group, a significant main effect of Filter was revealed ( $F(1,20) = 5.27$ ;  $p < .05$ ,  $\eta^2 = .07$ ), with greater activity for smiling compared to unsmiling sentences (Figure 3B-C). No main effect of Choice ( $F(1,20) = .07$ ,  $p > .05$ ) or interaction between Filter and Choice ( $F(1,20) = .004$ ,  $p > .05$ ) was observed.

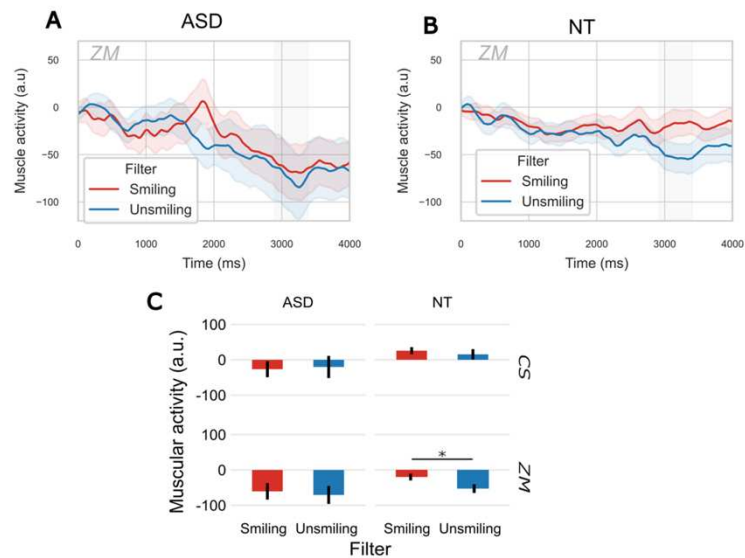


Figure 3: Muscular activity (ZM and CS) in response to Filter. A: Effect of Filter on ZM in ASD; B: Effect of Filter on ZM in NT; Effect of Filter (C) on CS and ZM in ASD and NT – mean response amplitude in the window 2900-3400ms (in grey in A and B time series). Shaded

areas and error bars represent standard error of the mean. \*  $p < .05$

### 2.1.3. Corrugator supercilii activity

As for the ZM activity, a repeated measures ANOVA was conducted to analyze mean muscular activity of the CS in the 2900-3400 ms time window considering Group as between, and Filter and Choice as within factors. No main effects were found for any of the factors ( $p > .05$ ), but significant interactions were observed between Group and Choice ( $F(1,42) = 5.34; p < .05; \eta^2 = .01$ ) and between Filter and Choice ( $F(1,42) = 4.55; p < .05; \eta^2 = .009$ ). Post-hoc pairwise comparisons revealed that the interaction between Group and Choice was due to the larger activity of the CS for Unsmile Choice in NT, which was difference from Smile Choice ( $p_{\text{corr}} < .05$ ) and larger than the CS for Unsmile and Smile Choice in ASD ( $p_{\text{corr}} < .05$ ). The interaction between Filter and Choice resulted from a statistical trend for the choice on Smiling filtered sentences, with greater activity in response to an Unsmile choice ( $p_{\text{corr}} = .1$ ), due to the larger activity in NT for Unsmile Choice.

Separate analyses within each group were performed to specify these effects. In the ASD group, no differences were found for Filter ( $F(1,19) = .12; p > .05$ ), Choice ( $F(1,19) = .48; p > .05$ ), nor their interaction ( $F(1,19) = 1.41; p > .05$ ) (Figure 4A-C). In the NT group, a significant main effect of Choice was revealed ( $F(1,20) = 5.60; p < .05; \eta^2 = .05$ ), with greater activity for sentences rated as Unsmile compared to Smile (Figure 4B-C). No main effect of Filter ( $F(1,20) = 1.40; p > .05$ ) or interaction between Filter and Choice ( $F(1,20) = 3.84; p > .05$ ) was observed.

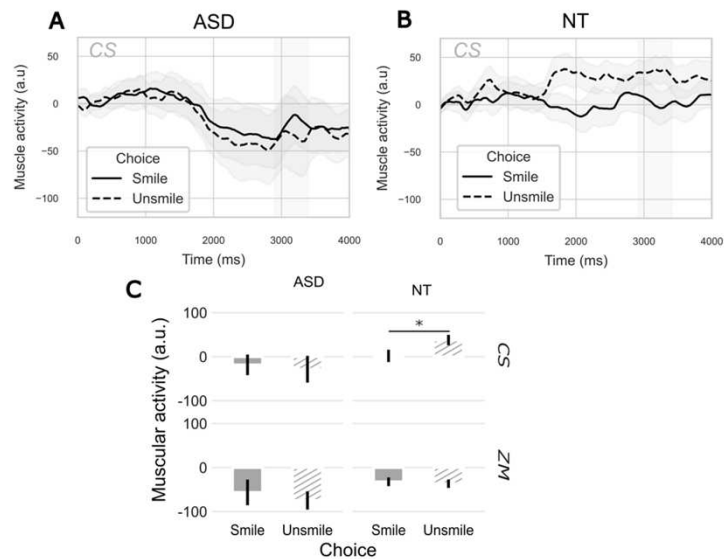


Figure 4: Muscular activity (ZM and CS) in according to Choice. A: Effect of Choice on ZM in ASD; B: Effect of Choice on ZM in NT; Effect of Choice (C) on CS and ZM in ASD and NT – mean response amplitude in the window 2900-3400ms (in grey in A and B time series). Shaded areas and error bars represent standard error of the mean. \*  $p < .05$

### 2.2. Correlation between measures

Previous analyses thus revealed that ZM activity was mainly modulated by filters, whereas the CS activity was sensitive to choice in NT. To simplify the expression of the effect of Filter and Choice on muscular activity, the difference between the Smiling and Unsmiling ZM response was computed as  $ZM_{filter}$ , and the difference between the Unsmile and Smile CS response was computed as  $CS_{choice}$ . Correlations between muscular activities were performed within each group. This analysis revealed no significant correlation between  $ZM_{filter}$  and  $CS_{choice}$  in either the NT group ( $t(19) = .38$ ;  $p_{corr} > .05$ ), or in the ASD group ( $t(18) = .64$ ;  $p_{corr} > .05$ ). Indicating that the activation of one effector did not trigger a fully congruent motor resonance.

No significant correlations were found in the NT group between  $ZM_{filter}$  and AQ or EQ, nor between  $CS_{choice}$  and AQ or EQ (all  $p_{corr} > .05$ ). Similarly, in the ASD group, neither  $ZM_{filter}$  nor  $CS_{choice}$  correlated significantly with AQ or EQ (all  $p_{corr} > .05$ ).

### 3. Discussion

The current study was the first to assess motor resonance to vocal emotion while controlling for perceptual representations of emotional prosody in a group of autistic adults.

Although vocal emotions were correctly perceived and judged by autistic adults, no motor resonance was observed. In contrast to neurotypicals, who exhibited *Zygomaticus major* activity reflecting perceived emotion and *Corrugator supercilii* activity linked to explicit emotional prosody processing, autistic adults showed no motor activity associated with either motor resonance or emotional judgment. More specifically, their motor response differed, with a marked relaxation of both muscles during sentence processing, suggesting an atypical integration of vocal emotional information beyond mere perception.

Distinct patterns of muscle activity were measured during the listening and processing of emotional prosody in the two groups. In autistic adults, during the listening of social stimuli, facial muscles initially tensed significantly and then relaxed, possibly due to the unpleasant nature of the auditory experience. This finding may align with behavioral results indicating that autistic adults tend to assign unpleasantness ratings to vocal stimuli<sup>34</sup>. This is also consistent with previous findings showing a decreased engagement of the somatosensory system during emotion processing in autism<sup>35</sup>. These results also suggest that cross-modal motor resonance (from hearing to visual expression) does not compensate for the lack of facial mimicry previously extensively described in the visual modality in autism<sup>2,20,22</sup>.

The current study indicates a difference in automatic and controlled processes in response to emotional prosody in autism. Indeed, controlled processes, such as choosing the most smiling sound from a pair of similar sounds, appeared to be preserved in autistic adults. This type of process relies on low-level acoustic processing, involving a purely frequency analysis that remains intact (or even enhanced) in autism<sup>36-38</sup>. The preserved perceptual representations and accurate recognition of vocal emotions in adults with autism are consistent with a compensatory strategy for controllable behaviors<sup>39</sup>. The use of acoustic cues similar to those of neurotypical

adults may have been acquired through experience and repeated exposure to such information.

In contrast, the automatic facial response to vocal emotion was absent in autistic adults.

An effect of gender on emotional prosody typicality was also observed with autistic women displaying more typical vocal smile models than men, which could explain their better performance in emotional prosody recognition. This aligns with studies by Chevallier et al.<sup>40</sup> and Globerson et al.<sup>41</sup>, which reported superior emotion recognition in autistic women, possibly due to more neurotypical-like acoustic processing.

Interestingly, the similar mental representation of vocal smiles and the equivalent performance in emotional prosody recognition between autistic and neurotypical adults raise questions about impaired recognition reported in several studies<sup>10,15</sup>. In the visual modality, it has been suggested<sup>42</sup> that while the representation of emotions may be similar between autistics and neurotypicals (possibly even more precise for autistics), differences in recognition and in subsequent response to emotions, could be attributed to autistic people being less guided by emotion representation as *a priori* within the Bayesian framework<sup>43</sup>.

Furthermore, the 'double empathy' problem posits that autistic and neurotypical individuals may have mutual misunderstandings rather than a deficit on one side<sup>44</sup>, suggesting a disconnect in reciprocal emotional understanding between autistic and neurotypical individuals. Consistent with this theory, the present study showed that autistic adults exhibited a lack of motor resonance while accurately recognizing emotional prosody, suggesting that their integration of emotions is different rather than impaired, and that difficulties in reciprocity are not a result of atypical emotional representations. This absence of motor resonance aligns with previous findings indicating challenges in the production and response to facial emotions in individuals with autism<sup>1</sup>. The absence of congruent motor activity in response to emotional prosody, along with the lack of correlation between accuracy and muscular activity suggests that emotion embodiment in autism may fundamentally differ from that in neurotypicals. Embodied emotion

theory states that emotional experiences are deeply integrated with physical responses, such as facial expressions and bodily sensations. These processes play a crucial role in how emotions are experienced and understood<sup>45</sup>. Consistent with this theory, it has been proposed that individuals with autism may experience a disconnect or alteration in how emotions are physically felt, expressed and processed<sup>46</sup>. Thereby, studies indicate that individuals with autism often show atypical facial expressions and reduced facial muscle activity in response to emotional stimuli, suggesting differences in embodied emotional processing<sup>1,47</sup>. This suggests that embodied automatic responses to emotional stimuli, such as smiling or frowning, may be less prevalent or different in individuals with autism. Nevertheless, despite differences in motor resonance, autistic adults were able to recognize emotional prosody accurately, indicating that autistic individuals might use alternative, more cognitive strategies for emotion recognition<sup>48</sup>. The atypicalities in motor resonance observed in autistic adults may stem from difficulties experienced during childhood. In infancy, emotional resonance is crucial for developing accurate perception and response to emotions even before the emergence of empathetic abilities<sup>49,50</sup>. Such early deficits in the embodiment of emotion, would lead to atypical emotional responses and, subsequently, challenges in social interaction and reciprocity. As previously indicated, emotional resonance is an automatic process; therefore, early difficulties in resonance to vocal emotions are likely to persist and impact emotional reactivity in adulthood.

The small sample size limits the generalizability of our findings, particularly given the heterogeneity of autism profiles in emotional prosody representations and responses. Future research should recruit larger and more diverse samples to capture the variability in vocal smile representations across different autistic profiles. Additionally, incorporating explicit emotional imitation tasks could help determine whether autistic individuals can voluntarily engage in motor mimicry and how this capacity relates to spontaneous emotional expression<sup>51</sup>. Furthermore, including a spontaneous vocal smile production task in autistic adults would allow

researchers to analyze the acoustic characteristics of smile production and compare these with perceptual representations and associated resonance capacities.

This study provides valuable clinical insights, especially the finding that autistic and neurotypical adults share similar representations of vocal smiles. This highlights the ability of autistic individuals to use relevant acoustic indices to classify vocal emotions effectively. Supporting autistic individuals in applying these robust representations in everyday life could be achieved through targeted interventions, such as speech and language therapy or psychomotor rehabilitation. These approaches could focus on enhancing the sensory and motor embodiment of emotions, aiming to improve the adaptation of emotional responses to social contexts. By leveraging pre-existing efficient representations and reinforcing their application, such therapies could help autistic individuals navigate emotional interactions more effectively in real-life situations.

The study of embodied emotion in autism reveals important differences in how emotions are integrated and expressed in this population. While autistic individuals may not exhibit typical motor resonance, their ability to understand and recognize emotions suggests that they experience emotions in a way that is different but not necessarily deficient. These findings challenge traditional deficit-based perspectives on emotion processing in autism and advocate for a more nuanced understanding of these processes. Furthermore, the present study offers new insights into the lack of synchrony in emotional prosody processing in autistic adults.

#### **4. Methods**

##### **4.1. Population**

Twenty autistic (ASD) and 21 neurotypical (NT) adults participated in this study. The autistic participants were recruited through the Autism Resource Center of Centre Val de Loire and diagnosed by a trained clinical team following DSM-5 criteria, with confirmation using ADOS-2 and ADI-R<sup>8,52</sup>. Verbal and non-verbal abilities in NT participants were assessed using

selected subtests from the WAIS-IV<sup>53</sup>, while ASD participants completed the full scales. All participants completed the Autism Quotient 50-questions (AQ) and the Empathy Quotient 40-questions (EQ)<sup>54,55</sup>. The NT group reported no developmental difficulties in language or sensorimotor acquisition. Both groups reported no central nervous system diseases, infectious or metabolic diseases, epilepsy, or abnormal hearing. Hearing level was checked for both ears using a brief subjective audiometry test. A detailed description of the two groups can be found in Table 1.

	NT	ASD	Comparison
<b>Sex</b>	7 ♀ 14 ♂	7 ♀ 13 ♂	$\chi^2 = 0; p > .05$
<b>Age</b>	26.1 ± 6	31.4 ± 10	t(30) = 2.2; p < .05
<b>vIQ</b>	125.5 ± 10 (19)	117.6 ± 20 (18)	t(24) = -1.6; p > .05
<b>nvIQ</b>	106.3 ± 12 (19)	102.7 ± 18 (18)	t(29) = -.5; p > .05
<b>AQ</b>	15.8 ± 7 (18)	33.9 ± 7 (18)	t(34) = 7.3; p < .001
<b>EQ</b>	44.2 ± 9 (18)	24.9 ± 15 (18)	t(30) = -4.8; p < .001
<b>ADOS</b>	/	10.4 ± 6 (18)	/

Table 1. Characteristics of Neurotypical (NT) and Autistic (ASD) group (mean ± standard deviation). vIQ : verbal Intelligence Quotient calculated from the WAIS-IV in ASD, and composite score calculated from vocabulary and similarities subtests in NT. nvIQ : non-verbal Intelligence Quotient calculated from the WAIS-IV in ASD, and composite score calculated from matrix reasoning and cubes subtests in NT. ADOS: score calculated as the sum of the subdimensions communication and social interaction behaviors. The values in brackets correspond to the number of values included in the calculation of the mean and standard deviation due to missing data.

#### 4.2. Ethics statement

The study was approved by *Comité de Protection des Personnes* CPP Est I (PROSCEA2017/23; ID RCB: 2017-A00756-47). All methods in this study were carried out in accordance with the relevant guidelines and regulations, and all data in this study were obtained with informed consent from all subjects and/or their legal guardian(s).

#### 4.3. Procedure

### **4.3.1. Preliminary validation of vocal smile model – Reverse correlation**

#### **4.3.1.1. Sound Modulations**

The method and stimuli from Ponsot et al. <sup>28</sup> was used to evaluate mental representations of vocal smiles in both groups. A French male speaker's /a/ phoneme with constant pitch was recorded, and a 500 ms stationary segment with constant spectral energy was selected. Spectral variants were generated by modulating the original sound using a random frequency equalizer. This equalizer, based on 25 log-spaced frequency points (100–10,000 Hz), interpolated linearly between gain values (in dB) drawn from Gaussian distributions (SD = 5 dB, clipped at  $\pm 2.5$  SD) using the Cleese Python toolbox <sup>56</sup>.

#### **4.3.1.2. Procedure**

The reverse correlation task consisted of paired comparisons, with each trial requiring participants to identify the "most smiling" voice. Autistic participants completed 150 trials (~15 min), while neurotypical participants completed 300 trials (~20 min), analyzed for convergence using the first 150 trials. A preliminary study with neurotypical adults confirmed that 150 trials were sufficient to generate reliable vocal smile representations <sup>31</sup>, showing no significant difference from the norm established with 500 trials <sup>28</sup>. The order between the reverse correlation and the main task was counterbalanced between participants.

#### **4.3.2. Main task**

The same task as in Merchie et al. <sup>31</sup> was presented using the same sentences corpus as in Arias et al.'s <sup>29,30</sup>. Semantically neutral sentences were artificially modified to a Smiling and an Unsmiling prosody according to the model of vocal smile demonstrated by Ponsot et al. <sup>28</sup>. Two 300 ms tones, each at a distinct frequency, were generated using Praat <sup>57</sup>, with energy normalization and fade-in and fade-out applied. The first tone (440 Hz) signaled the start of a new trial. After a delay of 1000 ms, the sentence was presented, followed by a 2000 ms period of silence. The second tone (275 Hz) then marked the beginning of the rating period and a return

to baseline, lasting 5000 ms. The timeline for each trial is illustrated in Figure 5A. The instructions given to participants were: "You will hear sentences, and you will need to rate how joyful, or not, the sentence was pronounced. Focus on the pronunciation rather than the content of the sentence." Ratings were recorded using a 4-key response pad. The image displayed on the screen throughout the task was a rating scale (Figure 5B): the two right keys on the pad represented two levels of Smile intensity, and the two left keys represented two levels of Unsmile intensity.

Note that the timing and luminance conditions were adapted to pupil dilation measurements that was a secondary objective of the study. Unfortunately, many autistic participants did not stare sufficiently at the screen during the task, resulting in too much missing data, which prevented a quality pupillometric analysis.

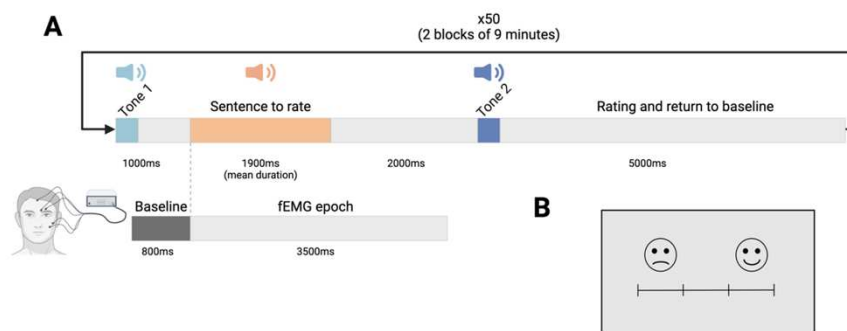


Figure 5: A: Timeline of a trial in the experimental sequence. Tone 1 indicated the beginning of a new trial, tone 2 indicated the time when the subject should rate the sentence. B: Rating scale displayed on screen throughout the duration of the task; fEMG: facial ElectroMyoGraphy. Adapted from Merchie et al. <sup>31</sup>.

Participants were seated comfortably 70 cm from the screen (1980x1080 px) and speakers, without the use of a chinrest. The sound intensity was controlled and maintained between 58- and 62-dB SPL.

#### 4.4. Measurements

Facial EMG (fEMG) was recorded using 4 mm Ag-AgCl electrodes (EL254 and EL254S) with

the BIOPAC MP36 system and AcqKnowledge® 4.1 software (BIOPAC® Systems Inc., Goleta, CA). To measure the activity of two facial muscles, the *Zygomaticus major* (ZM) and the *Corrugator supercilii* (CS), electrodes were placed on the participant's left face. The skin was scrubbed with an abrasive paste to remove any traces of moisturizer, makeup, or sweat, following the standard guidelines by Fridlund and Cacioppo<sup>58</sup>. Data were recorded at a sampling rate of 1024 Hz. For each muscle, two electrodes, separated by approximately 1 cm, were used to measure the potential difference between anode and cathode, and were placed in the same position on every participant's face. Additionally, a reference electrode was placed on the middle of the forehead at the hairline.

### **4.5. Data analysis**

#### **4.5.1. Reverse correlation model analysis**

The mental representation of vocal smiles was modeled using the classification image technique<sup>59</sup>. For each participant, the mean pitch contour of “unsmiling” voices was subtracted from that of “smiling” voices, and the result was normalized by dividing it by the sum of its absolute values. Individual representations were then averaged within each group, and comparisons of gain at each of the 25-modulated frequency were conducted between NT and ASD participants. To evaluate if the recognition of applied filters was related with the representation of vocal smile. The mean Euclidian distance for each frequency between the combined model of ASD and NT and a participant's model was computed as the Kernel Distance.

#### **4.5.2. Accuracy of the rating in the main task**

In the following parts, the acoustical modifications of the sentences are referred to as Filter (Smiling vs. Unsmiling) and the ratings of the smiliness of the sentences are referred to as Choice (Smile vs. Unsmile).

Accuracy was calculated as the number of Correct responses divided by the total number of Smiling and Unsmiling sentences (80 for each participant) to reflect the accurate rating of both filters (Equation 1). Accuracy was compared to the chance level (50%) to evaluate the

recognition of emotional prosodic modulation.

$$Accuracy = \frac{nSmiling_{smile} + nUnsmiling_{unsmile}}{nSmiling + nUnsmiling}$$

*Equation 1: Accuracy formula;  $nSmiling_{smile}$ : number of trials for smiling sentences rated as smile;  $nUnsmiling_{unsmile}$ : number of trials for Unsmiling sentences rated as unsmile;  $nSmiling$ : total of trials for the Smiling Filter;  $nUnsmiling$ : total of trials for the Unsmiling Filter.*

#### 4.5.3. fEMG processing

Analysis and pre-processing of fEMG data were conducted using the MNE toolbox <sup>60</sup> in Python 3.9.4. The data were filtered with a 50 Hz IIR high-pass filter and a 250 Hz low-pass filter. Subsequently, any significant muscular activity identified as artifacts was manually removed after visually inspecting all recordings. This was followed by segmenting the data into epochs with an 800 ms pre-stimulus baseline and a duration of 3500 ms, starting at the beginning of each sentence. The absolute value of the muscular activity was smoothed and averaged with a 300 ms sliding window, then normalized using z-scores based on the baseline of the current trial. On average, in the autistic group  $22 \pm 17\%$  (mean  $\pm$  SD) and  $20 \pm 11\%$  in neurotypical adults of trials were removed in each participant due to excessive movement artifacts.

#### 4.5.4. Statistical analysis

Statistical analyses were performed using Rstudio 4.2.2 <sup>61,62</sup> with the ggplot2 <sup>63</sup>, ez <sup>64</sup>, tidyverse <sup>65</sup>, and dplyr <sup>66</sup> packages.

Pearson's Correlations between the Kernel Distance and the Accuracy and between the Accuracy and the AQ/EQ scores were computed. Correlations between the muscular activity and the AQ/EQ scores were also computed in both groups. In case of multiple comparisons, a Bonferroni's correction was applied on p-value ( $p_{corr}$ ).

To evaluate the effect of listening sentences on muscular tonus in both groups the mean muscular activity was computed during Listening (0-1900 ms) and during Processing (1900-4000 ms) of sentences. A repeated-measures ANOVA on mean muscular activity with Timing (Listening vs Processing) and Muscle (ZM vs CS) as within factor, and Group (ASD vs NT) as between subject factors. For each muscle activity, the effects of Filter and Choice were

estimated through the mean amplitude within selected time windows, both between and within groups (ASD and NT). To specify the significant effect in the selected time window, a repeated-measures ANOVA (Filter x Choice) was performed on mean muscular activity.

In case of significant main effect or interaction, post-hoc Student's tests were performed with a Bonferroni's correction for multiple comparisons.

The effect of the gender on the Kernel Distance was estimated in each group with Wilcoxon test as the ratio between male and female is not balanced.

## 5. References

1. Trevisan, D. A., Hoskyn, M. & Birmingham, E. Facial Expression Production in Autism: A Meta-Analysis. *Autism Res.* **11**, 1586–1601 (2018).
2. McIntosh, D. N., Reichmann-Decker, A., Winkielman, P. & Wilbarger, J. L. When the social mirror breaks: deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Dev. Sci.* **9**, 295–302 (2006).
3. Yirmiya, N., Sigman, M. D., Kasari, C. & Mundy, P. Empathy and Cognition in High-Functioning Children with Autism. *Child Dev.* **63**, 150–160 (1992).
4. Faso, D. J., Sasson, N. J. & Pinkham, A. E. Evaluating posed and evoked facial expressions of emotion from adults with autism spectrum disorder. *J. Autism Dev. Disord.* **45**, 75–89 (2015).
5. Grossman, R. B. Judgments of social awkwardness from brief exposure to children with and without high-functioning autism. *Autism* **19**, 580–587 (2015).
6. Paul, R., Augustyn, A., Klin, A. & Volkmar, F. R. Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *J. Autism Dev. Disord.* **35**, 205–220 (2005).
7. Shriberg, L. D. *et al.* Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *J. Speech Lang. Hear. Res. JSLHR* **44**, 1097–1115 (2001).
8. Lord, C. *et al.* The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *J. Autism Dev. Disord.* **30**, 205–223 (2000).
9. Wang, J.-E. & Tsao, F.-M. Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Res. Dev. Disabil.* **37**, 162–170 (2015).
10. Hubbard, D. J., Faso, D. J., Assmann, P. F. & Sasson, N. J. Production and perception of emotional prosody by adults with autism spectrum disorder: Affective prosody in ASD. *Autism Res.* **10**, 1991–2001 (2017).
11. Rosenblau, G., Kliemann, D., Dziobek, I. & Heekeren, H. R. Emotional prosody processing in autism spectrum disorder. *Soc. Cogn. Affect. Neurosci.* **12**, 224–239 (2016).
12. Zhang, M. *et al.* Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. *Autism* **26**, 798–813 (2022).
13. Golan, O., Baron-Cohen, S., Hill, J. J. & Rutherford, M. D. The ‘Reading the Mind in the Voice’ Test-Revised: A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *J. Autism Dev. Disord.* **37**, 1096–1106 (2007).

14. Peppé, S., McCann, J., Gibbon, F., O'Hare, A. & Rutherford, M. Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *J. Speech Lang. Hear. Res.* **50**, 1015–1028 (2007).
15. Leung, F. Y. N. *et al.* Emotion recognition across visual and auditory modalities in autism spectrum disorder: A systematic review and meta-analysis. *Dev. Rev.* **63**, 101000 (2022).
16. Dimberg, U., Thunberg, M. & Elmehed, K. Unconscious Facial Reactions to Emotional Facial Expressions. *Psychol. Sci.* **11**, 86–89 (2000).
17. Helt, M. S., Eigsti, I.-M., Snyder, P. J. & Fein, D. A. Contagious Yawning in Autistic and Typical Development: Contagious Yawning. *Child Dev.* **81**, 1620–1631 (2010).
18. Helt, M. S., Sorensen, T. M., Scheub, R. J., Nakhle, M. B. & Luddy, A. C. Patterns of Contagious Yawning and Itching Differ Amongst Adults With Autistic Traits vs. Psychopathic Traits. *Front. Psychol.* **12**, 645310 (2021).
19. Senju, A. *et al.* Absence of contagious yawning in children with autism spectrum disorder. *Biol. Lett.* **3**, 706–708 (2007).
20. Beall, P. M., Moody, E. J., McIntosh, D. N., Hepburn, S. L. & Reed, C. L. Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. *J. Exp. Child Psychol.* **101**, 206–223 (2008).
21. Magnée, M. J. C. M., De Gelder, B., Van Engeland, H. & Kemner, C. Facial electromyographic responses to emotional information from faces and voices in individuals with pervasive developmental disorder. *J. Child Psychol. Psychiatry* **48**, 1122–1130 (2007).
22. Oberman, L. M., Winkielman, P. & Ramachandran, V. S. Slow echo: facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Dev. Sci.* **12**, 510–520 (2009).
23. Norscia, I., Zanolli, A., Gamba, M. & Palagi, E. Auditory Contagious Yawning Is Highest Between Friends and Family Members: Support to the Emotional Bias Hypothesis. *Front. Psychol.* **11**, (2020).
24. Giganti, F. & Esposito Ziello, M. Contagious and spontaneous yawning in autistic and typically developing children. *Curr. Psychol. Lett. Behav. Brain Cogn.* (2009).
25. Rychlowska, M. *et al.* Functional Smiles: Tools for Love, Sympathy, and War. *Psychol. Sci.* **28**, 1259–1270 (2017).
26. Dawson, G., Hill, D., Spencer, A., Galpert, L. & Watson, L. Affective exchanges between young autistic children and their mothers. *J. Abnorm. Child Psychol.* **18**, 335–345 (1990).
27. Merchie, A. *et al.* Intact Representation of Vocal Smile in Autism: A reverse correlation approach. Preprint at <https://doi.org/10.31234/osf.io/q4n7z> (2024).

28. Ponsot, E., Arias, P. & Aucouturier, J.-J. Uncovering mental representations of smiled speech using reverse correlation. *J. Acoust. Soc. Am.* **143**, EL19–EL24 (2018).
29. Arias, P., Belin, P. & Aucouturier, J.-J. Auditory smiles trigger unconscious facial imitation. *Curr. Biol.* **28**, R782–R783 (2018).
30. Arias, P., Bellmann, C. & Aucouturier, J.-J. Facial mimicry in the congenitally blind. *Curr. Biol.* **31**, R1112–R1114 (2021).
31. Merchie, A. *et al.* Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance. *Sci. Rep.* **14**, 25043 (2024).
32. Brewer, R. *et al.* Can Neurotypical Individuals Read Autistic Facial Expressions? Atypical Production of Emotional Facial Expressions in Autism Spectrum Disorders. *Autism Res. Off. J. Int. Soc. Autism Res.* **9**, 262–271 (2016).
33. Lindström, R. *et al.* Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders. *Clin. Neurophysiol.* **129**, 2411–2420 (2018).
34. Michel, L., Ricou, C., Bonnet-Brilhault, F., Houy-Durand, E. & Latinus, M. Sounds Pleasantness Ratings in Autism: Interaction Between Social Information and Acoustical Noise Level. *J. Autism Dev. Disord.* (2023) doi:10.1007/s10803-023-05989-6.
35. Fanghella, M., Gaigg, S. B., Candidi, M., Forster, B. & Calvo-Merino, B. Somatosensory evoked potentials reveal reduced embodiment of emotions in autism. *J. Neurosci.* (2022) doi:10.1523/JNEUROSCI.0706-21.2022.
36. Bonnel, A. *et al.* Enhanced pitch sensitivity in individuals with autism: a signal detection analysis. *J. Cogn. Neurosci.* **15**, 226–235 (2003).
37. Bonnel, A. *et al.* Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia* **48**, 2465–2475 (2010).
38. Samson, F. *et al.* Atypical processing of auditory temporal complexity in autistics. *Neuropsychologia* **49**, 546–555 (2011).
39. Livingston, L. A., Shah, P. & Happé, F. Compensatory strategies below the behavioural surface in autism: a qualitative study. *Lancet Psychiatry* **6**, 766–777 (2019).
40. Chevallier, C., Noveck, I., Happé, F. & Wilson, D. What’s in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia* **49**, 507–517 (2011).
41. Globerson, E., Amir, N., Kishon-Rabin, L. & Golan, O. Prosody recognition in adults with high-functioning autism spectrum disorders: from psychoacoustics to cognition. *Autism Res. Off. J. Int. Soc. Autism Res.* **8**, 153–163 (2015).

42. Keating, C. T., Ichijo, E. & Cook, J. L. Autistic adults exhibit highly precise representations of others' emotions but a reduced influence of emotion representations on emotion recognition accuracy. *Sci. Rep.* **13**, 11875 (2023).
43. Lawson, R. P., Rees, G. & Friston, K. J. An aberrant precision account of autism. *Front. Hum. Neurosci.* **8**, (2014).
44. Milton, D. E. M. On the ontological status of autism: the 'double empathy problem'. *Disabil. Soc.* **27**, 883–887 (2012).
45. Niedenthal, P. M. Embodying Emotion. *Science* **316**, 1002–1005 (2007).
46. Eigsti, I.-M. A Review of Embodiment in Autism Spectrum Disorders. *Front. Psychol.* **4**, (2013).
47. Sato, W., Fujimura, T., Kochiyama, T. & Suzuki, N. Relationships among Facial Mimicry, Emotional Experience, and Emotion Recognition. *PLOS ONE* **8**, e57889 (2013).
48. Jones, C. R. G. *et al.* A multimodal approach to emotion recognition ability in autism spectrum disorders. *J. Child Psychol. Psychiatry* **52**, 275–285 (2011).
49. Meltzoff, A. N. & Decety, J. What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **358**, 491–500 (2003).
50. Waal, F. B. & Ferrari, P. F. Towards a bottom-up perspective on animal and human cognition. *Trends Cogn. Sci.* **14**, 201–207 (2010).
51. Heaton, P., Williams, K., Cummins, O. & Happé, F. G. E. Beyond perception: musical representation and on-line processing in autism. *J. Autism Dev. Disord.* **37**, 1355–1360 (2007).
52. Lord, C., Rutter, M. & Le Couteur, A. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* **24**, 659–685 (1994).
53. Wechsler, D. Wechsler Adult Intelligence Scale--Fourth Edition. <https://doi.org/10.1037/t15169-000> (2012).
54. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).
55. Baron-Cohen, S. & Wheelwright, S. The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *J. Autism Dev. Disord.* **34**, 163–175 (2004).

56. Burred, J. J., Ponsot, E., Goupil, L., Liuni, M. & Aucouturier, J.-J. CLEESE: An open-source audio-transformation toolbox for data-driven experiments in speech and music cognition. *PLOS ONE* **14**, e0205943 (2019).
57. Boersma, P. Praat, a system for doing phonetics by computer. *Glott Int.* **5**, 341–345 (2002).
58. Fridlund, A. J. & Cacioppo, J. T. Guidelines for human electromyographic research. *Psychophysiology* **23**, 567–589 (1986).
59. Brinkman, L., Todorov, A. & Dotsch, R. Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *Eur. Rev. Soc. Psychol.* **28**, 333–361 (2017).
60. Gramfort, A. *et al.* MNE software for processing MEG and EEG data. *NeuroImage* **86**, 446–460 (2014).
61. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2022).
62. Team, R. C. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2021).
63. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).
64. Lawrence, M. A. *Ez: Easy Analysis and Visualization of Factorial Experiments*. (2016).
65. Wickham, H. *et al.* Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
66. Wickham, H., François, R., Henry, L. & Müller, K. *Dplyr: A Grammar of Data Manipulation*. (2021).

### **Acknowledgements**

We thank all the volunteers for their time and participation in this study. We also would like to thank Shasha Morel-Kohlmeyer and Marine Bessé for helping in the recruitment of participants and Marta Manenti for her English proofreading.

### **Statements**

#### **Ethics approval and consent to participate**

The protocol received approval from Ethics Committee (PROSCEA2017/23; ID RCB: 2017-A00756-47). Each participant signed an informed consent form.

#### **Availability of data and materials**

Please contact the corresponding author for data and experimental stimuli requests.

#### **Conflict of interest disclosure**

All authors declare that they have no conflicts of interest.

#### **Funding**

Annabelle Merchie was funded by the FEANS (Fondation Européenne pour l'Avancement des Neurosciences) and the ANR (ANR-19-CE37-0022-01). The funding sources were not involved in the study design, data collection and analysis, or the writing of the report.

### 3.4. Synthèse

Comme présenté dans l'étude n°3 les adultes autistes ont une représentation perceptive du sourire vocal similaire à celle des adultes neurotypiques, ce qui ne permet pas d'expliquer les difficultés de réponses automatiques aux émotions. Quand les représentations sont similaires, la comparaison d'une émotion perçue à ces représentations devrait générer une réponse motrice appropriée, ce qui contribuerait à une contagion émotionnelle, comme observé chez les sujets neurotypiques (étude n°4). Cependant, chez les adultes autistes, aucune modulation de l'activité musculaire en réponse à des émotions vocales perçues (résonance motrice implicite) ou jugées (résonance motrice explicite) n'a été mise en évidence. La différence dans la réponse (action), et non dans les représentations suppose que l'absence de réponse aux émotions dans l'autisme pourrait être associée spécifiquement à cette dernière étape, possiblement en relation avec des difficultés d'incarnation des émotions.

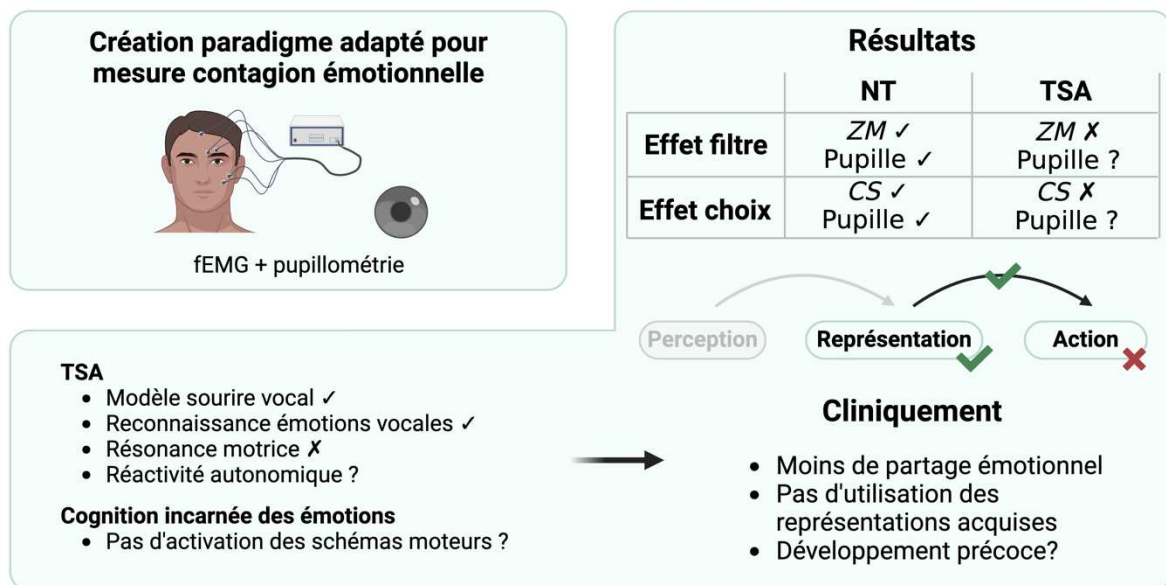


Figure 21: Schéma de synthèse des études n°4 et 5 à propos de la résonance motrice et autonome chez l'adulte neurotypique et autiste



## **Discussion générale**

---



## 1. Discussion

### 1.1. Synthèse des résultats

Dans ce travail, l'objectif était d'étudier le traitement de la prosodie émotionnelle dans l'autisme dans le cadre de la boucle *Perception – Représentation – Action*. Pour cela plusieurs techniques d'exploration (EEG, fEMG, pupillométrie, corrélation inverse) ont été utilisées conduisant à la réalisation de 5 études différentes. Une synthèse des principaux résultats de chacune des études, dans le cadre de la boucle *Perception – Représentation – Action*, est présentée dans la figure 22.

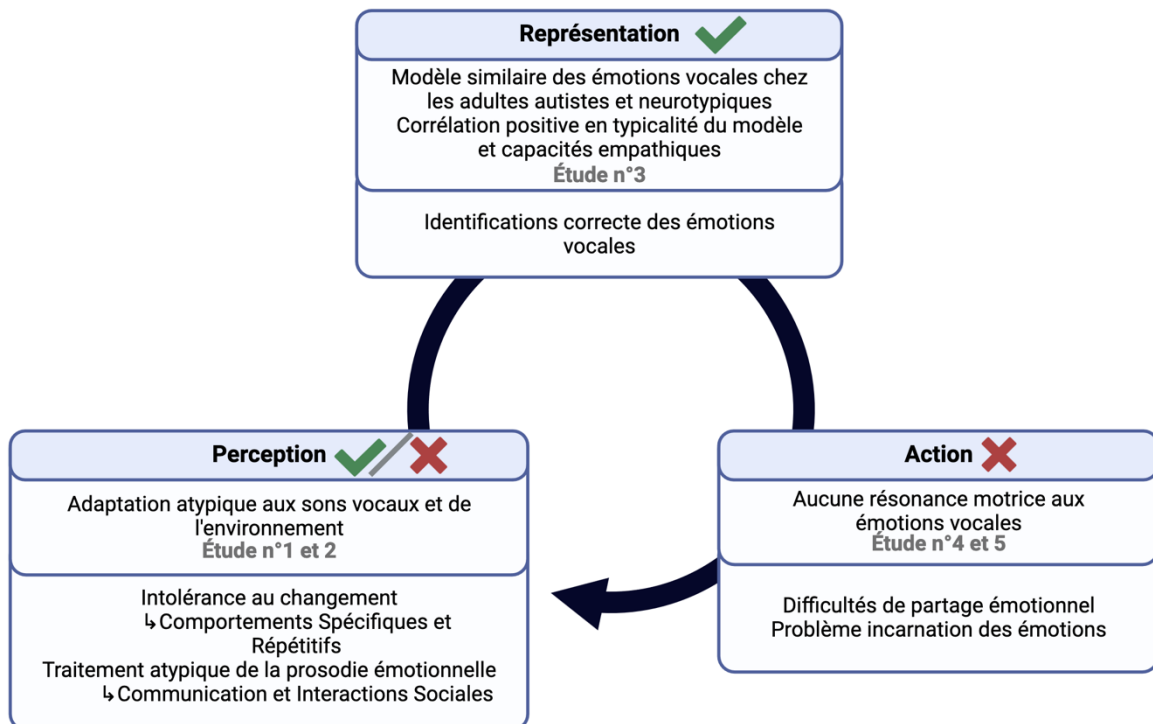
Ces différentes études ont révélé des atypies dans le traitement et la réponse à la prosodie émotionnelle dans l'autisme. Les études n°1 et 2 ont montré que chez les adultes neurotypiques, c'est la complexité acoustique d'un son qui joue un rôle prépondérant dans l'adaptation neurale. En effet, l'adaptation est plus rapide pour des sons non-vocaux que pour des sons vocaux, mais sans distinction en fonction du contenu émotionnel. En revanche, chez les personnes autistes, l'adaptation est plus rapide pour les sons vocaux que pour les sons non-vocaux, pour lesquels aucune modulation de l'amplitude de la réponse neurale n'est observée. Ainsi, bien que les RP soient similaires dans les deux groupes, l'adaptation à la voix est légèrement plus rapide chez les adultes autistes que chez les adultes neurotypiques. Cependant, le contenu émotionnel des sons vocaux ralentit ce processus d'adaptation chez les TSA. Dès l'étape de perception et de traitement sensoriel de la prosodie émotionnelle des atypies sont observées dans l'autisme, pouvant avoir un impact sur les étapes suivantes, et par conséquent sur la réponse aux émotions vocales.

Après avoir observé des atypies dans l'adaptation neurale (études n°1 et 2), l'étude n°3 s'est intéressée à un autre aspect du traitement de la prosodie émotionnelle : les représentations perceptives, cette fois-ci en focalisant sur la perception du sourire vocal chez les adultes autistes. Cette étude a montré que dans l'ensemble, les adultes autistes partagent les mêmes représentations que les adultes neurotypiques. Seul un tiers de l'échantillon d'adultes autistes a une représentation différente ou trop variable pour être interprétée. Cette absence de différence pour la plupart des adultes autistes reflète l'utilisation d'indices acoustiques pertinents dans la classification des émotions vocales. Un lien entre la typicalité de la représentation du sourire vocal et les capacités empathiques a aussi été mis en lumière chez

## Discussion

les adultes autistes et neurotypiques. De meilleures capacités empathiques améliorent les représentations mentales des émotions vocales et donc, potentiellement, leur reconnaissance.

En complément des études sur l'adaptation neurale et les représentations perceptives, les deux dernières études (n°4 et 5) se sont concentrées sur la résonance motrice et la contagion émotionnelle au sourire vocal, offrant un aperçu des réponses automatiques aux émotions dans l'autisme. Une contagion émotionnelle en réponse à des émotions vocales a été mise en évidence chez l'adulte neurotypique (étude n°4) grâce à la mesure conjointe des activités musculaire faciale et pupillaire. Cependant, chez les adultes autistes, malgré des performances de reconnaissance préservées, aucune modulation de l'activité musculaire n'a été observée (étude n°5) en réponse à un sourire vocal. Aucun lien n'a été montré entre la représentation perceptive et la résonance motrice chez les adultes neurotypiques ou autistes. L'absence de résonance motrice n'apparaît donc pas comme une conséquence d'une différence de représentation perceptive.



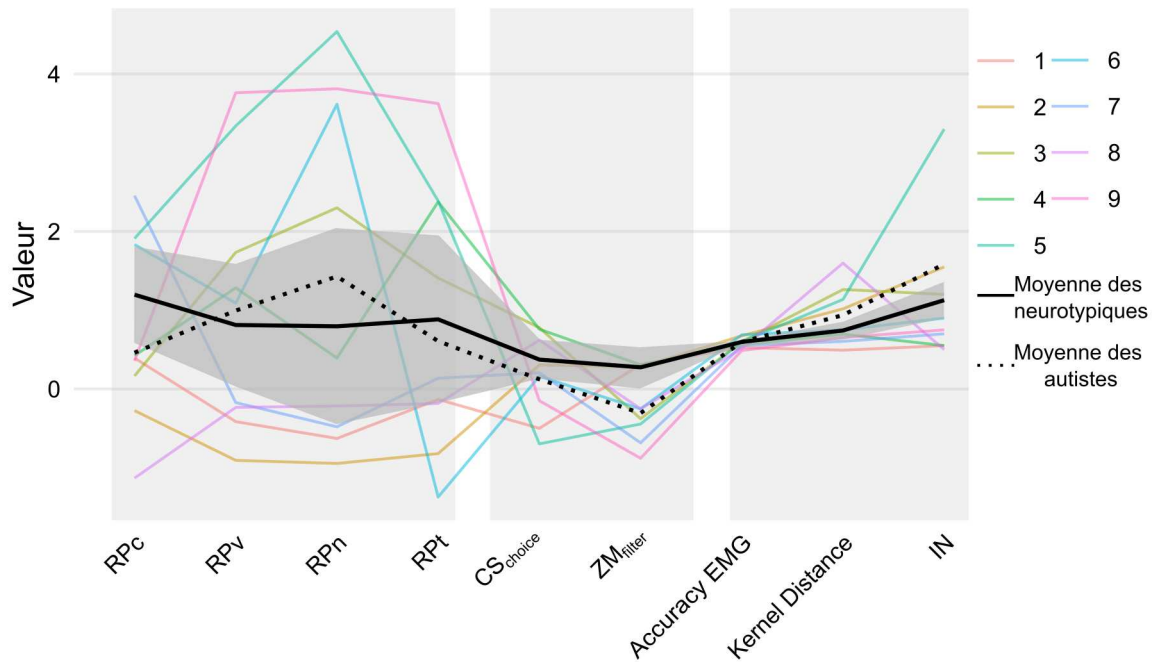
**Figure 22 : Récapitulatif des résultats dans l'autisme des différentes études et conséquences prévisibles sur le plan clinique**

Les différentes études menées au cours de la thèse ont permis de répondre aux questions de recherche pour chacune des étapes indépendamment. Néanmoins, l'étude des différentes

## Discussion

étapes chez une même personne autiste est importante pour une compréhension intégrative du traitement de la prosodie émotionnelle. Ainsi une description qualitative des résultats chez les 9 participants autistes qui ont participé à toutes les études a été réalisée. La figure 23 reprend une partie des résultats de ces adultes autistes ayant participé à l'ensemble des études. Pour l'adaptation, seules les amplitudes moyennes des RP (RPc, RPv, RPn, RPt) ont été prises en compte. Pour la résonance motrice, seule l'activité musculaire en fonction du choix et de l'émotion a été choisie ( $CS_{\text{choice}}$ ,  $ZM_{\text{filter}}$ ). Enfin, des données comportementales ont été prises en compte : la performance à la tâche de reconnaissance d'EMG (Accuracy EMG) et les variables issues de la tâche de corrélation inverse (kernel distance : distance par rapport au modèle de Ponsot et al., 2018, IN : bruit interne estimé). La moyenne des résultats de l'ensemble du groupe d'adultes neurotypiques et autistes est aussi représentée. D'après cette analyse qualitative, aucun profil ne ressort clairement comme étant partagé entre plusieurs individus autistes. Néanmoins, on remarque que les sujets 5 et 2 ont un bruit interne plus élevé que les autres patients et que la moyenne des neurotypiques. Leur RP sont également les plus éloignées de la moyenne des neurotypiques (plus ample pour le patient n°5 et négative pour le patient n°2). On peut aussi observer que le patient n°4 a une RP plus ample en réponse aux sons tristes par rapport aux autres sons et que sa résonance motrice est similaire à celle des neurotypiques ( $CS_{\text{choice}}$  et  $ZM_{\text{filter}} > 0$ ), ce qui pourrait indiquer un profil similaire à celui de la moyenne des neurotypiques.

Cette représentation graphique permet également de rendre compte de performances homogènes chez les patients autistes et les adultes neurotypiques pour la reconnaissance des émotions vocales alors que les autres indices sont plus hétérogènes. L'inclusion de plus de patients dans l'ensemble des paradigmes permettrait de tester l'existence de profils de patients à partir des données expérimentales. Enfin, la comparaison avec les profils cliniques et/ou sensoriels serait intéressante pour avoir une meilleure prise en charge des patients.



**Figure 23 : Profils des participants autistes ayant participé à l'ensemble des études**

Chaque couleur représente un patient, la ligne pleine épaisse noire la moyenne des résultats les adultes neurotypiques ( $\pm$ intervalle de confiance à 95%) et la ligne noire en pointillés les résultats chez les adultes autistes. RPs: amplitude moyenne de la Repetition Positivity en dans la fenêtre 90-200ms sur l'électrode Fz en réponse aux sons complexes, vocaux, neutres et tristes. CS<sub>choice</sub>: différence de l'amplitude moyenne du Corrugator supercilii dans la fenêtre 2900-3400ms en fonction du choix (Unsmile – Smile). ZM<sub>filter</sub>: différence de l'amplitude moyenne du Zygomaticus major dans la fenêtre 2900-3400ms en fonction du filtre (Smiling – Unsmiling). Accuracy EMG: taux de reconnaissance des émotions vocales (Smiling et Unsmiling) dans le protocole d'EMG et de pupillométrie. Kernel Distance: distance du modèle individuel au modèle du sourire vocal mis en évidence par Ponsot et al., (2018). IN: Internal Noise, bruit interne à l'issue de la tâche de corrélation inverse.

## 1.2. Réponse contrôlée versus réponse automatique : stratégies de compensation ?

Les résultats de ce travail de thèse révèlent des différences en fonction de la nature contrôlée ou automatique des tâches effectuées. En effet, dans les tâches de corrélation inverse et de reconnaissance des émotions vocales, les participants autistes ont des performances similaires à celles des neurotypiques. En revanche, les processus automatisés, tels que l'adaptation neurale et la résonance motrice, se révèlent atypiques chez les adultes autistes. Cela suggère qu'il existe des différences dans le traitement et la réponse à la prosodie émotionnelle selon que le processus sous-jacent soit contrôlé ou automatique.

Dans l'autisme il a été montré que le traitement acoustique de bas niveau était préservé, voire amélioré, pour des sons purs (Bonnel et al., 2003, 2010; Samson et al., 2006). Dans l'étude en

## Discussion

corrélation inverse, nous avons montré que le traitement fréquentiel de bas niveau semble également intact chez les adultes autistes dans le cadre des émotions vocales. Les adultes autistes ont pu reconnaître et identifier correctement les émotions vocales en se basant sur des représentations et des indices acoustiques similaires à ceux des neurotypiques.

De plus, le fait que la reconnaissance des filtres, créés et appliqués à partir des représentations d'adultes neurotypiques soit supérieure au hasard chez les autistes suggère que ces représentations perceptives puissent être généralisées à l'ensemble de notre population. L'utilisation de ces représentations pour modifier des phrases neutres a entraîné une résonance motrice congruente chez les adultes neurotypiques. Toutefois, aucune modulation de l'activité musculaire chez les autistes n'a été observée, malgré une reconnaissance correcte. Comme mentionné précédemment, les adultes autistes se fient principalement à un traitement de bas niveau (traitement des informations purement acoustiques) pour classer et identifier les émotions vocales.

Les processus contrôlés, qui nécessitent un traitement conscient, comme le choix d'un son plus souriant parmi une paire de sons similaires et la constance dans ce choix, semblent préservés chez les adultes autistes. Ce type de processus repose, comme indiqué précédemment, sur un traitement acoustique, impliquant une analyse purement fréquentielle qui reste intacte dans l'autisme (Bonnell et al., 2003, 2010; Samson et al., 2011). Les performances similaires dans ces tâches d'analyse acoustique, contrastant avec les difficultés de traitement des émotions d'autrui au quotidien, pourraient s'inscrire dans la théorie de la « faible cohérence centrale » dans l'autisme (Frith, 2003). La cohérence centrale désigne la capacité du cerveau humain à intégrer et structurer les informations de manière à en former un ensemble cohérent, ce qui implique de percevoir les informations dans un contexte plus large, d'identifier les relations entre elles et d'en extraire une signification globale, plutôt que de se concentrer uniquement sur des détails isolés. Dans l'autisme, on parle de faible cohérence centrale, ce qui suppose une attention accrue aux détails au détriment d'une perception globale (Happé & Frith, 2006; Shah & Frith, 1993). Les résultats comportementaux de cette thèse (corrélation inverse et reconnaissance des filtres) soutiennent cette théorie : bien que le traitement acoustique soit fin et précis chez les adultes autistes, la généralisation des représentations, hors du laboratoire ou dans des conditions plus écologiques, reste difficile.

## Discussion

Les représentations perceptives préservées et la bonne reconnaissance des émotions vocales chez les adultes autistes concordent avec l'idée de la mise en place de stratégies de compensation pour les comportements contrôlables ou maîtrisables (Livingston et al., 2019). En effet, l'utilisation d'indices acoustiques similaires à ceux des adultes neurotypiques pourrait avoir été acquise grâce à l'expérience et à l'exposition répétée à ces informations. Cette acquisition pourrait être facilitée chez les femmes autistes. En effet, une différence de genre a été observée dans l'étude n°5, avec de représentations perceptives des émotions vocales chez les femmes autistes par rapport aux hommes. Cette différence dans la reconnaissance des émotions faciales et vocales a déjà été relevée à plusieurs reprises dans l'autisme (Harms et al., 2010; Lacroix et al., 2024; Lai et al., 2015). Toutefois, dans ce travail, elle semble également présente au niveau de la typicalité des représentations perceptives des émotions vocales chez les femmes autistes. Dans l'autisme, l'hypothèse d'un camouflage, notamment social, plus fréquent chez les femmes que chez les hommes a été proposée (Lai et al., 2019; Schuck et al., 2019). Cependant, malgré une bonne reconnaissance des émotions et des représentations perceptuelles typiques, les difficultés de résonances motrices sont présentes, quel que soit le genre.

Ainsi la capacité à se « synchroniser » avec les émotions perçues chez autrui pourrait être moins fluide chez les personnes autistes, ce qui affecterait la perception immédiate des émotions. Néanmoins, pour les processus moins automatisés, plus délibérés et conscients, comme la reconnaissance des émotions ou l'identification du sourire vocal, les individus autistes peuvent recourir à des stratégies compensatoires ou à un traitement plus analytique qui semble préservé.

### **1.3. Traitement de la prosodie émotionnelle dans l'autisme : cas du cerveau Bayésien ?**

Dans l'autisme, les théories bayésiennes et de codage prédictif (*predictive coding*) ont été proposées pour expliquer les difficultés perceptuelles observées (Brock, 2012; Lawson et al., 2014; Pellicano & Burr, 2012; Van de Cruys et al., 2013). Ces théories s'appuient sur le théorème de Bayes, adapté à la perception. Selon ce modèle, la perception consciente (*a posteriori*) d'un stimulus repose à la fois sur les informations sensorielles traitées (*vraisemblance*) et sur les représentations internes (*a priori*) de ce stimulus :

$$a \text{ posteriori} \propto \text{vraisemblance} \cdot a \text{ priori}$$

## Discussion

Dans le cadre de l'autisme, un déséquilibre dans la précision des trois éléments nécessaires à la perception a été proposé (Brock, 2012; Friston, 2016; Lawson et al., 2014; Pellicano & Burr, 2012; Sinha et al., 2014; Van de Cruys et al., 2013). Différentes théories suggèrent soit une faible influence des a priori (Pellicano & Burr, 2012), soit une hypersensibilité sensorielle (Brock, 2012), ce qui, dans les deux cas, conduit à une perception moins biaisée par les attentes préalables mais fortement liée à l'expérience sensorielle. Une autre hypothèse suggère que les différences comportementales et perceptuelles dans l'autisme pourraient résulter d'un système qui attribue un poids trop important tant aux a priori qu'aux perceptions sensorielles (Van de Cruys et al., 2013, 2014). Ces études se sont d'abord concentrées sur l'explication des premières étapes de la boucle *Perception – Représentation – Action* en examinant l'adaptation neurale, notamment à travers des recherches sur la détection du changement et l'erreur de prédiction dans le cadre des théories bayésiennes (Friston, 2003).

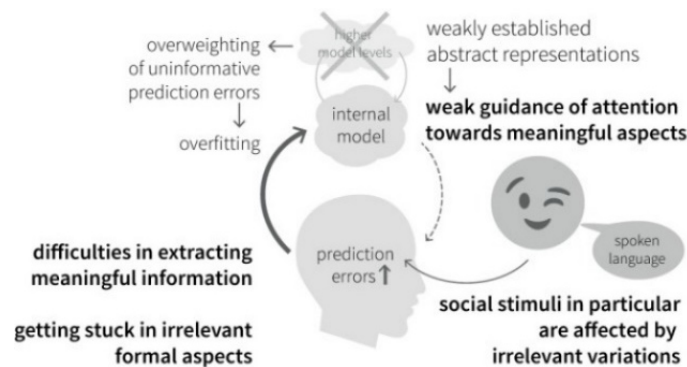
Dans ce cadre théorique, nos résultats neurophysiologiques peuvent être interprétés de la manière suivante. L'adaptation neurale contribuerait à la formation et à l'affinage des représentations internes, où la RP reflète la mise en place progressive de ces représentations. Les représentations internes (*a priori*) semblent intactes dans l'autisme, mais leur utilisation présente des difficultés.

L'étude de Ruiz-Martinez et al. (2020) montre qu'il n'y a pas de différence dans l'adaptation neurale (ou un équivalent d'adaptation neurale, étant donné que leur paradigme évaluait la détection du changement plutôt que l'encodage de la régularité) en fonction de la nature du son chez les enfants autistes, contrairement aux enfants neurotypiques. Nos résultats restent en accord avec cette étude, avec une adaptation neurale atypique aux sons vocaux et non vocaux pouvant expliquer la recherche d'immuabilité et l'intolérance au changement observées sur le plan comportemental. De plus, dans l'autisme, des erreurs de prédiction atypique ont été mises en évidence à travers de la MMN avec une latence (Charpentier, Kovarski, Houy-Durand, et al., 2018; Kujala et al., 2007; Lepistö et al., 2005, 2006) ou une amplitude (Charpentier, Kovarski, Houy-Durand, et al., 2018; Jansson-Verkasalo et al., 2003; Kasai et al., 2005; Lepistö et al., 2006) atypique. Comme présenté dans la revue de la littérature dans l'introduction, ces difficultés dans la détection du changement pourraient être conséquentes à une adaptation atypique comme montré dans l'étude n°2. En outre, une réduction de l'adaptation neurale a été associée aux difficultés sociales observées dans l'autisme (Green et al., 2018) et à des erreurs

## Discussion

de prédiction atypique, ces difficultés pouvant entraîner des différences dans la perception sensorielle avec un poids trop important accordé au percept par rapport à l'*a priori* (Pellicano & Burr, 2012).

Dans le cadre du cerveau bayésien, les représentations perceptives évaluées par la tâche de corrélation inverse pourraient refléter l'*a priori* auquel le percept est comparé (Haker et al., 2016) (figure 24). Les différences observées dans la réponse motrice et l'adaptation neurale pourraient alors résulter d'un poids excessif accordé à cet *a priori* (Van de Cruys et al., 2013). Cela entraînerait une comparaison imparfaite avec le percept, conduisant à une erreur de prédiction dans la réponse.



**Figure 24 : Cerveau Bayésien et prédiction dans l'autisme (Haker et al., 2016)**

*Des erreurs de prédiction de bas niveau surpondérées, dues à une précision sensorielle trop élevée, entraînent un surajustement du modèle interne et des difficultés à extraire des informations significatives empêchant la formation de représentations abstraites de haut niveau.*

Ainsi, les résultats de cette thèse sont en cohérence avec les théories du codage prédictif et du cerveau bayésien dans l'autisme. Les difficultés à généraliser les modèles en dehors des contextes expérimentaux et les atypies d'adaptation neurale aux sons de l'environnement pourraient constituer des points clés à explorer pour améliorer la qualité de vie et les interactions sociales des personnes autistes.

### 1.4. L'autisme : une cognition incarnée atypique ?

Dans l'autisme, des différences dans la cognition incarnée ont été proposées (Eigsti, 2013; Moseley & Pulvermüller, 2018), et nos résultats peuvent également être interprétés dans ce cadre théorique. En effet, la similarité des représentations perceptuelles entre adultes autistes et neurotypiques, combinée aux différences de réponses neuronales et motrices aux émotions

## Discussion

vocales, semble aller dans ce sens. Comme évoqué précédemment, les processus les plus contrôlés semblent préservés dans l'autisme, tandis que les processus automatisés sont différents. La cognition incarnée est caractérisée par l'interaction entre le corps, le cerveau et l'environnement, où les processus cognitifs dépendent des expériences sensori-motrices et sont influencés par le contexte et les actions physiques. La réponse aux émotions, et en particulier dans le cas du sourire, peut ainsi être traitées dans le cadre de la cognition incarnée, (Niedenthal et al., 2010). Par exemple, dans des tâches de lecture de mots émotionnels (selon leur contenu sémantique) en IRMf, des chercheurs ont observé une absence d'activation des zones motrices et limbiques chez des adolescents (Moseley, Ypma, et al., 2015) et des adultes autistes (Moseley, Shtyrov, et al., 2015), contrairement à l'activité observée dans un groupe contrôle.

Les difficultés liées à la cognition incarnée chez les personnes autistes ne semblent pas se limiter aux réponses émotionnelles ; des différences dans l'activité des schémas moteurs pour des actions sans contenu émotionnel ont également été mises en évidence (Moseley & Pulvermüller, 2018). Ainsi pour des contenus sémantiques non émotionnels, les adultes autistes étaient « plus lents mais pas moins efficaces » que le groupe contrôle dans le traitement des verbes d'action (Moseley & Pulvermüller, 2018). Dans sa revue de la littérature sur la cognition incarnée dans l'autisme, Eigsti (2013) propose trois hypothèses pour expliquer les différences observées. Les résultats des études menées dans ce travail de thèse peuvent apporter des éléments de réponse à ces hypothèses :

- ***L'encodage original des informations du stimulus pourrait être plus imprécis.***

Nos résultats en électrophysiologie convergent dans le sens de cette hypothèse, en montrant un traitement et un encodage neural de la prosodie émotionnelle et de la voix sensiblement différent des adultes neurotypiques chez les personnes autistes. De plus, ces différences électrophysiologiques pourraient influencer la réponse motrice : si l'encodage est moins efficace, la réponse motrice qui en découle ne pourra pas l'être non plus.

- ***L'incarnation est, fondamentalement, un problème de réponse et de planification motrice.***

## Discussion

Dans l'étude n°5, des différences d'activité musculaire ont été observées entre les adultes autistes et neurotypiques. Chez les adultes autistes, le tonus musculaire reste constant pendant l'écoute de la phrase, puis il est suivi d'un relâchement. Cette première phase tonique exclue une hypotonie supposée des muscles de la face en réponse aux émotions vocales dans l'autisme (Al Abdulmohsen & Kruger, 2011; Serdarevic et al., 2017). Au contraire, chez les neurotypiques, la réponse musculaire reste soutenue après l'écoute et se trouve modulée par le contenu et le jugement de la phrase. Il est possible que les adultes autistes n'aient pas encodé correctement les réponses motrices associées aux émotions entendues, ce qui conduirait à une incarnation atypique de l'émotion.

Néanmoins, nos données ne permettent pas de tirer des conclusions à propos de la planification motrice.

- ***Ce sont les différences de priorité attentionnelle qui régissent le choix des expériences stockées en vue d'une réactivation ultérieure***

Une fois de plus, nos résultats électrophysiologiques pourraient être interprétés dans ce sens : une attention réduite aux informations sociales et émotionnelles pourrait expliquer l'adaptation neurale atypique observée pour ces sons.

Les différences observées concernent les processus automatisés, pour lesquels la mise en place de stratégies de compensation est plus complexe. Par conséquent, les stimuli auxquels une personne autiste est confrontée peuvent être moins liés aux conditions sensorielles et motrices associées aux premières rencontres avec le stimulus (Eigsti, 2013).

Ces éléments peuvent également être envisagés sous un angle développemental. En effet, la résonance motrice se manifeste dès les premiers mois de vie, bien avant l'émergence des capacités empathiques chez les enfants (Meltzoff & Decety, 2003; Waal & Ferrari, 2010). Si, dans l'autisme, une diminution de l'attention pour les stimuli et les interactions sociales est observée dès les premières années de vie, cela pourrait entraver la mise en place de la résonance motrice. De plus, si l'encodage et le traitement automatique des émotions sont atypiques, la seule façon pour les personnes autistes de répondre de manière appropriée aux émotions vocales pourrait être de passer par un apprentissage conscient et contrôlé.

## 2. Limitations

Comme toute étude expérimentale, ce travail de recherche présente des limites, notamment méthodologiques, qui sont abordées dans cette section. Des recherches supplémentaires sont nécessaires pour tester la reproductibilité des résultats et pour pallier ces limitations.

### 2.1. Limites méthodologiques relatives aux protocoles

Tout d'abord, dans l'étude n°2, pour examiner l'effet du contenu émotionnel de la voix sur le processus d'adaptation neurale, nous avons choisi d'utiliser des voix tristes plutôt que des voix souriantes. Ce choix était basé sur les résultats d'une étude préliminaire (voir annexe 2) qui montrait une adaptation trop rapide et qui ne se maintenait pas à la voix souriante chez des adultes neurotypiques. Cependant, il aurait été intéressant d'utiliser des voix souriantes, conformément au modèle du sourire vocal mis en évidence par la tâche de corrélation inverse (Ponsot, Arias, et al., 2018). Dans de futures recherches, il serait intéressant de comparer l'adaptation neurale à différentes émotions vocales pour déterminer si la valence ou l'éveil (*arousal*) influence les réponses chez les personnes autistes ou neurotypiques. De plus, il serait pertinent d'examiner la relation entre la distance des représentations perceptuelles et l'amplitude de la *Repetition Positivity* (RP) pour obtenir une vue d'ensemble plus complète de la perception de la prosodie émotionnelle dans l'autisme. En effet, la RP peut être considérée, comme spécifié dans le cadre des théories bayésiennes, comme un marqueur de traitement abouti de la représentation interne. Elle a également été considérée à l'origine comme reflétant simplement une baisse de la réactivité neurale à une stimulation répétée. La comparaison entre l'amplitude de cette RP et la représentation perceptive du sourire vocal serait intéressante afin de pouvoir conclure sur l'interprétation de la RP dans ce contexte.

Il semble important de souligner que les paradigmes mis en place dans le cadre de cette thèse apparaissent comme adaptés et fiables pour mettre en évidence les processus d'intérêt. En effet, les résultats chez les adultes neurotypiques sont robustes et reproductibles dans chacune des études. Pour l'adaptation neurale, dans les études n°1 et 2 et dans l'étude préliminaire (annexe 2) nous avons mis en évidence le même résultat principal avec une adaptation plus rapide pour les sons non-vocaux que vocaux. Les résultats de l'étude n°4 sont aussi globalement similaires à ceux de l'étude préliminaire (annexe 3) avec une activité du

## Limitations

Zygomatique associée au filtre et une activité du Corrugateur liée au jugement émis. Ces résultats sont aussi en accord avec une autre étude qui cherchait à caractériser la résonance motrice chez des adultes neurotypiques (Arias et al., 2018). Enfin, nos résultats de corrélation inverse sont similaires dans les études n°3 et 4 pour les adultes neurotypiques et sont concordant avec ceux de l'étude de Ponsot et collaborateurs (2018). Les paradigmes créés et mis en place dans cette thèse sont donc robustes et permettent d'étudier efficacement les différentes étapes de traitement de la prosodie émotionnelle dans l'autisme.

Dans l'étude n°5, nous n'avons pas pu inclure les données de pupillométrie acquises chez les autistes en raison d'une quantité trop importante de données manquantes. Cette situation s'explique par le fait que les adultes neurotypiques ont regardé l'écran de manière plus soutenue, permettant ainsi un suivi du regard plus efficace. En revanche, les adultes autistes ont moins focalisé leur regard sur l'écran, certains regardant leurs mains au moment du jugement ou fermant les yeux pour mieux se concentrer. En conséquence, l'absence de données de pupillométrie analysables nous empêche de tirer des conclusions sur la réponse autonome aux émotions vocales chez les personnes autistes et, par extension, sur la contagion émotionnelle. Pour pallier ce problème dans des recherches futures, il serait utile d'envisager l'utilisation d'une mentonnière pour limiter les mouvements du participant (bien que cela puisse induire un inconfort), ou de dissimuler les mains et le clavier de réponse avec un tissu ou une table afin d'éviter que les participants ne regardent leurs mains. L'ajout de données sur l'activité autonome dans l'autisme pourrait, en outre, contribuer au débat théorique sur les émotions (Cannon, 1927; Damasio & Blanc, 1995; James, 1884; Ledoux, 2005). En effet, si une activité autonome est présente sans réponse motrice, cela pourrait indiquer l'existence d'une modification de l'état émotionnel dans l'autisme, indépendamment de l'activation des effecteurs somatiques.

### **2.2. Taille de l'échantillon et reflet de l'hétérogénéité clinique**

Dans ce travail, malgré nos efforts pour inclure des patients présentant des profils variés, nous ne pouvons pas pleinement rendre compte de l'hétérogénéité clinique de l'autisme. En effet, les tailles d'échantillon sont relativement petites : 20 adultes autistes pour les études n°2 et 5, et 30 pour l'étude n°3. Bien que ces effectifs soient comparables à ceux d'autres études utilisant des méthodes expérimentales similaires, ils ne permettent pas de tirer des

## Limitations

conclusions généralisables à une échelle plus large. De plus, bien que nous ayons tenté d'utiliser des protocoles adaptés à la population clinique, deux des trois protocoles étaient des tâches actives nécessitant une réponse du participant, ce qui a rendu plus difficile l'inclusion de patients avec des habilités cognitives moins développées. Les résultats des deux dernières étapes de la boucle pourraient ainsi refléter les stratégies de compensation mises en place par les participants ayant de meilleures compétences. Cependant, nous avons veillé à ne pas mettre les personnes en difficulté en leur proposant les paradigmes les plus adaptés, selon notre jugement et celui de l'équipe clinique. En conséquence, tous les participants inclus ont réalisé au moins une tâche. Seul un patient autiste n'a pas été en mesure de compléter la tâche de corrélation inverse, malgré notre tentative de l'inclure. De plus, comme indiqué dans l'annexe 1, des enfants autistes et neurotypiques, âgés de 8 à 14 ans, ont depuis été inclus dans des études utilisant les mêmes protocoles et ont été capables de réaliser les tâches proposées. Cela suggère que les paradigmes mis en place et utilisés dans cette thèse sont globalement adaptés à la population clinique cible, et permettent aussi de tester des participants avec des habilités variables.

Enfin, l'inclusion de participants différents dans chacune des études a limité notre capacité à établir des profils individuels pour un grand nombre de personnes autistes. Cela est lié à notre méthode de recrutement. La plupart des patients autistes (17 sur 38) ont été recrutés grâce à la collaboration entre le Centre Ressources Autisme Centre-Val-de-Loire et l'équipe de recherche au sein du centre d'excellence Autisme et Troubles du Neurodéveloppement de Tours (EXAC-T). Ces patients participaient à la recherche pendant leur bilan diagnostique, mais le temps alloué à cette participation était limité à 1h30, ce qui ne permettait pas de compléter les trois paradigmes expérimentaux qui nécessitaient une durée totale de 2 heures.

De plus, une partie de notre échantillon a été recrutée pour cette recherche, mais aussi pour d'autres études menées au laboratoire. Dans ces cas-là, il était impossible de réaliser l'intégralité du protocole en raison du manque de temps, et pour éviter une trop grande fatigue des participants. En conséquence, seuls 9 patients ont pu participer aux trois protocoles, ce qui a limité notre capacité à analyser les profils expérimentaux pour ces individus.

Bien que certaines limitations méthodologiques aient été identifiées, elles offrent également des pistes intéressantes pour les futures recherches. En effet, en surmontant ces obstacles,

nous pourrions améliorer notre compréhension du traitement de la prosodie émotionnelle dans l'autisme et développer des approches thérapeutiques encore plus ciblées.

### **3. Perspectives**

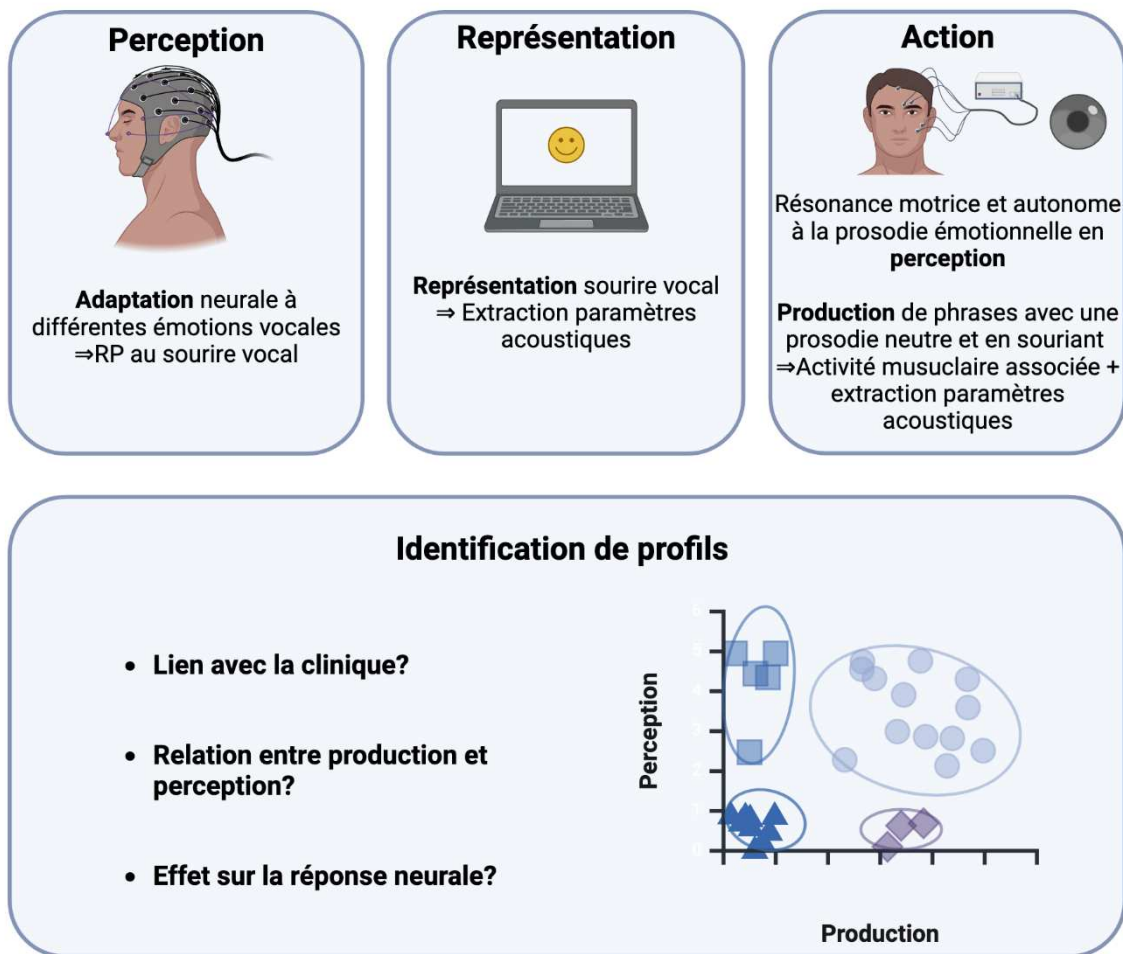
#### **3.1. Perspectives de recherche**

Dans cette thèse, nous avons concentré notre analyse sur la perception de la prosodie émotionnelle dans l'autisme en utilisant diverses méthodes d'exploration physiologiques et comportementales, mais nous n'avons pas mesuré la production de la prosodie. Il a été suggéré que la perception de la prosodie était préservée dans l'autisme, comme l'indiquent les résultats de la tâche de corrélation inverse. En revanche, des déficits de production prosodique sont rapportés et pourraient être liés à d'autres aspects du langage, comme la pragmatique et/ou la syntaxe (Paul, Augustyn, et al., 2005). Pour enrichir les résultats de cette étude, il serait intéressant d'ajouter des tâches de production langagière et de prosodie émotionnelle afin de comparer la perception et la production chez les mêmes patients. Ainsi, les patients dont les représentations perceptuelles des émotions vocales diffèrent de celles d'un groupe témoin, produisent-ils des indices acoustiques différents pour exprimer cette émotion ? Il serait également pertinent de mesurer les variations acoustiques prosodiques des productions des patients lors de tâches de production langagière sans prosodie émotionnelle, telles que les répétitions de phrases (Manenti et al., 2023; Tuller et al., 2021) ou de non-mots (Ferré et al., 2022) développées au sein de l'équipe (LITMUS-SR-FR-Adults et LITMUS-QU-NWR-FR-Adults). Ces tâches sont standardisées et utilisent une prosodie naturelle qui facilite la répétition pour les participants. L'intégration de tâches de production, en plus des tâches de perception et de traitement, permettrait d'extraire et de comparer certains paramètres acoustiques avec les résultats perceptuels. À partir de ces tâches, une étude a permis de classifier des participants autistes et neurotypiques selon leur profil clinique en utilisant les caractéristiques acoustiques de la voix lors de la répétition de non-mots (Briend et al., 2023). Un projet de ce type est en cours de développement au sein de l'équipe de recherche dans le cadre d'une thèse intitulée « Prosodie et langage formel dans le trouble du spectre de l'autisme », dirigée par Sandrine Ferré et Marie Gomot, qui débutera en octobre 2024.

## Perspectives

Il serait également intéressant de comparer l'expression émotionnelle spontanée ou dirigée avec la résonance motrice et la représentation perceptive des émotions vocales, dans une optique plus intégrative. Les contractions musculaires mesurées dans les études n°4 et 5 sont des micro-contractions musculaires, non perceptibles à l'œil nu. Il serait intéressant de comparer les amplitudes de ces réponses musculaires avec l'amplitude d'une contraction induite et forcée du *Zygomaticus major* et du *Corrugator supercilii* chez des adultes autistes. En effet, cette comparaison d'amplitude permettrait de rendre compte de l'intensité de la réponse motrice aux émotions vocales, au regard des capacités de contraction musculaire. De plus, il serait pertinent de demander aux participants de prononcer des phrases de manière neutre et émotionnelle, en utilisant les mêmes phrases que celles employées dans les études n°4 et 5. Cela permettrait de comparer les indices acoustiques (comme les modulations fréquentielles et d'intensité) de la prosodie émotionnelle produite par les participants, avec les filtres appliqués. Cette approche permettrait d'examiner s'il existe un lien entre la distance des représentations perceptives et les modulations prosodiques en production. La figure 25 reprend ces idées et des résultats simulés à partir d'une étude de ce type. À partir d'un tel protocole, l'hypothèse de l'existence de différents profils selon la perception et la production de la prosodie peut être posée, permettant une meilleure compréhension de ce volet de la symptomatologie de l'autisme.

La compréhension de ces processus dans une approche intégrative similaire dans une autre modalité et dans une approche multimodale serait aussi intéressante. Les différences et les similarités entre personnes autistes et neurotypiques mises en lumière dans ce travail sont-elles retrouvées dans la modalité visuelle en réponse aux émotions faciales ? Est-ce que dans des tâches impliquant des émotions vocales et faciales simultanées, l'ajout d'indice dans une autre modalité faciliterait l'intégration et la réponse aux émotions ? Il a été démontré que l'intégration d'informations audiovisuelles facilite la réponse aux émotions par rapport à l'analyse séparée des informations visuelles ou auditives (Collignon et al., 2008; Kreifelts et al., 2007; Paulmann & Pell, 2011). Ce sont des questions auxquelles il serait possible de répondre dans l'autisme avec un protocole similaire dans une autre modalité sensorielle.



*Figure 25 : Protocole et résultats envisagés pour tester conjointement la perception et la production de la prosodie émotionnelle*

Enfin, une autre perspective de recherche consiste à explorer la perception de la prosodie émotionnelle au cours du développement typique ainsi que dans l'autisme. Une telle étude permettrait de déterminer si les processus associés à la perception émotionnelle se mettent en place rapidement chez les enfants, ou s'ils nécessitent une expérience préalable. Plus spécifiquement, une étude comparative entre le développement typique et l'autisme pourrait révéler des fenêtres critiques dans le développement de ces processus, offrant ainsi des opportunités importantes pour des interventions de remédiation dans l'autisme. De plus, identifier ces processus dans le développement pourrait mettre en lumière des différences avec les résultats observés chez les adultes. Cela pourrait révéler des stratégies de compensation ou d'apprentissage qui se développent au fil du temps, et qui sont plus intégrées à l'âge adulte.

### 3.2. Applications thérapeutiques

Au-delà des observations théoriques, ces résultats soulèvent des questions cruciales pour la prise en charge clinique. En effet, les atypies de traitement des émotions vocales que nous avons mises en évidence pourraient influencer directement les interventions visant à améliorer les compétences émotionnelles des personnes autistes. Nos résultats suggèrent des atypies dans la perception et la réponse automatique aux émotions vocales chez les personnes autistes, même si leurs représentations de ces émotions ne semblent pas fondamentalement différentes de celles des personnes neurotypiques. Dans un contexte clinique, ces résultats pourraient guider une prise en charge plus ciblée et individualisée des patients. En effet, deux scénarios peuvent être envisagés pour améliorer la réponse et le traitement des émotions vocales chez les adultes autistes, afin de favoriser leur intégration sociale. Actuellement, il existe peu de thérapies spécifiquement conçues pour améliorer l'intégration d'indices sociaux, notamment ceux liés à la voix (Berggren et al., 2018). La plupart des thérapies se concentrent de manière indirecte sur ces aspects, dans le cadre de rééducation des compétences pragmatiques ou sociales. Ces interventions adoptent une approche globale, visant à aider les patients à mieux identifier les émotions en utilisant divers indices, tels que la prosodie, le contexte et les expressions faciales.

Il semble que les difficultés rencontrées proviennent d'un défaut de généralisation des émotions vocales, ce qui complique l'utilisation efficace d'outils basés sur ces apprentissages, comme le programme *Mind Reading* (Baron-Cohen, 2004), un guide informatisé pour la reconnaissance des émotions. Pour les patients autistes qui possèdent des représentations préservées des émotions vocales, il serait pertinent de proposer un programme de remédiation axé sur l'amélioration de la perception et de la réponse à ces émotions, en s'appuyant sur les compétences déjà présentes. Une approche intéressante serait de leur enseigner comment utiliser ces représentations afin d'y répondre de manière appropriée. Cela pourrait se faire par le biais d'une rééducation orthophonique ou psychomotrice, en mettant l'accent sur l'incarnation sensorielle et motrice des émotions. L'objectif serait d'améliorer leur capacité à adapter leurs réponses émotionnelles en fonction des contextes sociaux, en utilisant les représentations déjà intégrées et en renforçant leur compréhension et leur utilisation de ces émotions dans des situations réelles.

## Perspectives

Pour les patients autistes ayant des représentations mentales atypiques des émotions vocales, une approche potentiellement bénéfique pourrait être l'intégration d'une pratique musicale ou d'une exposition accrue à la musique. La thérapie musicale est déjà utilisée avec succès chez les enfants autistes pour développer les compétences en communication sociale, en tirant parti de leur intérêt souvent marqué pour la musique (Geretsegger et al., 2022; LaGasse, 2014). Cette thérapie peut inclure diverses activités telles que le chant, la pratique instrumentale, et l'écoute de morceaux, pratiquées en groupe ou individuellement. Les recherches ont montré que cette approche améliore la communication (Boso et al., 2007; Geretsegger et al., 2014), la motivation (Kim et al., 2008) et l'engagement social (Ghasemtabar et al., 2015; LaGasse, 2014) des personnes autistes. Les progrès sont souvent facilités par des mécanismes tels que l'imitation, la répétition et les interactions non verbales durant les séances d'improvisation musicale (Edgerton, 1994; Kim et al., 2008; Mössler et al., 2019). Chez les adultes autistes, ces méthodes pourraient également être exploitées pour affiner les représentations émotionnelles. L'expertise musicale a été associée à une meilleure perception de la prosodie émotionnelle (Lima & Castro, 2011; Pinheiro et al., 2015; Pralus et al., 2019) ce qui pourrait améliorer le traitement et la réponse aux émotions dans l'autisme. En utilisant la musique pour développer des compétences perceptuelles et émotionnelles, il est possible de renforcer la capacité des patients à reconnaître et à répondre aux émotions vocales de manière plus adaptée et nuancée.

Il est essentiel de noter que l'efficacité des outils thérapeutiques peut être limitée si ces outils sont uniquement informatisés. Un contact direct entre le thérapeute et le patient reste fondamental. En effet, des études récentes en imagerie fonctionnelle ont montré que le traitement du discours et des émotions nécessite des processus de mentalisation plus élaborés lors d'interactions directes, comparé à des situations préenregistrées moins naturelles (Rice & Redcay, 2016). Les personnes avec des traits autistiques marqués peuvent présenter des altérations dans les mécanismes neuraux associés à ces interactions directes (Rice & Redcay, 2016). Ainsi, une approche en deux étapes pour les thérapies comportementales pourrait être plus efficace : débiter avec un entraînement virtuel (par exemple, via des programmes informatiques) et passer ensuite à des thérapies de groupe. En effet le feedback des pairs dans les thérapies de groupe améliore le comportement prosocial chez les adolescents autistes (Van Hoorn et al., 2017), et pourrait aider à transférer les compétences acquises aux situations sociales réelles.

## Conclusion

Enfin, il est important de souligner que nos résultats proviennent principalement d'adultes autistes ayant un niveau fonctionnel relativement élevé. Par conséquent, les outils thérapeutiques suggérés ne peuvent pas être généralisés à l'ensemble des personnes sur le spectre de l'autisme. Il est crucial de mieux comprendre ces processus chez les patients présentant des difficultés sociales et langagières plus prononcées afin d'adapter les interventions et optimiser l'accompagnement pour améliorer leur réponse sociale et émotionnelle.

## 4. Conclusion

Ce travail de thèse a permis de caractériser les différentes étapes de la boucle *Perception – Représentation – Action* chez des adultes autistes. Nous avons mis en évidence que les réponses cérébrales liées à l'adaptation neurale sont influencées par le contenu social et prosodique tant chez les adultes autistes que neurotypiques. L'adaptation neurale à la voix est généralement plus lente que celle aux sons de l'environnement chez les adultes neurotypiques, tandis que chez les adultes autistes, cette adaptation est plus marquée en réponse aux émotions vocales, révélant des atypies dans ce domaine.

Nous avons également mis en lumière que, pour une majorité d'adultes autistes, les représentations perceptives des émotions vocales ne diffèrent pas significativement de celles des neurotypiques. Cependant, cette similitude perceptuelle ne suffit pas à induire une résonance motrice appropriée. En effet, alors que les émotions vocales déclenchent une résonance motrice et une contagion émotionnelle chez les adultes neurotypiques, ce phénomène est absent chez les adultes autistes, ce qui en fait une étape primordiale dans les difficultés d'adaptation sociale rencontrées par les personnes autistes.

Ce travail a ainsi permis de mieux comprendre le traitement et la réponse aux émotions vocales dans l'autisme, éclairant sur les difficultés d'interaction sociale rencontrées par ces individus. Des recherches complémentaires sont nécessaires pour explorer ces phénomènes en réponse à d'autres émotions ou dans d'autres modalités sensorielles. Enfin, l'analyse de ces processus au cours du développement typique et dans l'autisme pourrait aider à identifier une fenêtre critique pour une intervention efficace et adaptée concernant la prosodie émotionnelle, permettant ainsi d'optimiser les périodes d'accompagnement les plus appropriées.



## Références

---



## Références

- Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2008). Right-Hemisphere Auditory Cortex Is Dominant for Coding Syllable Patterns in Speech. *Journal of Neuroscience*, *28*(15), 3958-3965. <https://doi.org/10.1523/JNEUROSCI.0187-08.2008>
- Adl Zarrabi, A., Jeulin, M., Bardet, P., Commère, P., Naccache, L., Aucouturier, J.-J., Ponsot, E., & Villain, M. (2024). A simple psychophysical procedure separates representational and noise components in impairments of speech prosody perception after right-hemisphere stroke. *Scientific Reports*, *14*(1), 15194. <https://doi.org/10.1038/s41598-024-64295-y>
- Aguert, M., Laval, V., Lacroix, A., Gil, S., & Le Bigot, L. (2013). Inferring Emotions from Speech Prosody: Not So Easy at Age Five. *PLoS ONE*, *8*(12), e83657. <https://doi.org/10.1371/journal.pone.0083657>
- Aguillon-Hernandez, N., Mofid, Y., Latinus, M., Roché, L., Bufo, M. R., Lemaire, M., Malvy, J., Martineau, J., Wardak, C., & Bonnet-Brilhault, F. (2020). The pupil : A window on social automatic processing in autism spectrum disorder children. *Journal of Child Psychology and Psychiatry*, *61*(7), 768-778. <https://doi.org/10.1111/jcpp.13170>
- Al Abdulmohsen, T., & Kruger, T. H. C. (2011). The contribution of muscular and auditory pathologies to the symptomatology of autism. *Medical Hypotheses*, *77*(6), 1038-1047. <https://doi.org/10.1016/j.mehy.2011.08.044>
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890425596>
- American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5-TR). American Psychiatric Association Publishing. <https://doi.org/10.1176/appi.books.9780890425787>
- Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., & Haynes, J.-D. (2011). Flow of affective information between communicating brains. *NeuroImage*, *54*(1), 439-446. <https://doi.org/10.1016/j.neuroimage.2010.07.004>
- Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology*, *84*(5), 1054-1068. <https://doi.org/10.1037/0022-3514.84.5.1054>
- Andronikof, A., & Fontan, P. (2016). Grounia Efimovna Soukhareva : La première description du syndrome dit d'Asperger. *Neuropsychiatrie de l'Enfance et de l'Adolescence*, *64*(1), 58-70. <https://doi.org/10.1016/j.neurenf.2015.07.007>

## Références

- Arias, P., Belin, P., & Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. *Current Biology*, 28(14), R782-R783. <https://doi.org/10.1016/j.cub.2018.05.084>
- Arias, P., Bellmann, C., & Aucouturier, J.-J. (2021). Facial mimicry in the congenitally blind. *Current Biology*, 31(19), R1112-R1114. <https://doi.org/10.1016/j.cub.2021.08.059>
- Arias, P., Rachman, L., Liuni, M., & Aucouturier, J.-J. (2021). Beyond Correlation : Acoustic Transformation Methods for the Experimental Study of Emotional Voice and Speech. *Emotion Review*, 13(1), 12-24. <https://doi.org/10.1177/1754073920934544>
- Asperger, H. (1944). Die „Autistischen Psychopathen“ im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten*, 117(1), 76-136. <https://doi.org/10.1007/BF01837709>
- Aston-Jones, G., & Cohen, J. D. (2005). AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*, 28(1), 403-450. <https://doi.org/10.1146/annurev.neuro.28.061604.135709>
- Aubergé, V., & Cathiard, M. (2003). Can we hear the prosody of smile? *Speech Communication*, 40(1-2), 87-97. [https://doi.org/10.1016/S0167-6393\(02\)00077-8](https://doi.org/10.1016/S0167-6393(02)00077-8)
- Babiker, A., Faye, I., Prehn, K., & Malik, A. (2015). Machine Learning to Differentiate Between Positive and Negative Emotions Using Pupil Diameter. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01921>
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614-636. <https://doi.org/10.1037/0022-3514.70.3.614>
- Baron-Cohen, S. (2004). *Mind Reading: The Interactive Guide to Emotions*. Jessica Kingsley Publishers.
- Barthel, H., & Quené, H. (2015). Acoustic-phonetic properties of smiling revised—Measurements on a natural video corpus. In M. Wolters, J. Livingstone, B. Beattie, R. Smith, M. MacMahon, J. Stuart-Smith, & J. M. Scobbie (Éds.), *18th International Congress of Phonetic Sciences, ICPHS 2015, Glasgow, UK, August 10-14, 2015*. University of Glasgow. <https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0337.pdf>
- Basso, F., & Oullier, O. (2010). “Smile down the phone” : Extending the effects of smiles to vocal social interactions. *Behavioral and Brain Sciences*, 33(6), 435-436. <https://doi.org/10.1017/S0140525X10001469>

## Références

- Beall, P. M., Moody, E. J., McIntosh, D. N., Hepburn, S. L., & Reed, C. L. (2008). Rapid facial reactions to emotional facial expressions in typically developing children and children with autism spectrum disorder. *Journal of Experimental Child Psychology*, *101*(3), 206-223. <https://doi.org/10.1016/j.jecp.2008.04.004>
- Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In *Handbook of psychophysiology, 2nd ed.* (p. 142-162). Cambridge University Press.
- Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding Voice Perception. *British Journal of Psychology*, *102*(4), 711-725. <https://doi.org/10.1111/j.2044-8295.2011.02041.x>
- Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice : Neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129-135. <https://doi.org/10.1016/j.tics.2004.01.008>
- Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices : A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531-539. <https://doi.org/10.3758/BRM.40.2.531>
- Belin, P., Zatorre, R. J., & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Research. Cognitive Brain Research*, *13*(1), 17-26. [https://doi.org/10.1016/s0926-6410\(01\)00084-2](https://doi.org/10.1016/s0926-6410(01)00084-2)
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309-312. <https://doi.org/10.1038/35002078>
- Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, *87*(1), 527-570. <https://doi.org/10.1007/BF01797193>
- Berggren, S., Fletcher-Watson, S., Milenkovic, N., Marschik, P. B., Bölte, S., & Jonsson, U. (2018). Emotion recognition training in autism spectrum disorder : A systematic review of challenges related to generalizability. *Developmental Neurorehabilitation*, *21*(3), 141-154. <https://doi.org/10.1080/17518423.2017.1305004>
- Bidet-Caulet, A., Latinus, M., Roux, S., Malvy, J., Bonnet-Brilhault, F., & Bruneau, N. (2017). Atypical sound discrimination in children with ASD as indicated by cortical ERPs. *Journal of Neurodevelopmental Disorders*, *9*(1), 13. <https://doi.org/10.1186/s11689-017-9194-9>
- Blasi, A., Lloyd-Fox, S., Sethna, V., Brammer, M. J., Mercure, E., Murray, L., Williams, S. C. R., Simmons, A., Murphy, D. G. M., & Johnson, M. H. (2015). Atypical processing of voice sounds in infants at risk for autism spectrum disorder. *Cortex; a Journal Devoted to the*

## Références

- Study of the Nervous System and Behavior*, 71, 122-133.  
<https://doi.org/10.1016/j.cortex.2015.06.015>
- Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341-345.
- Bolinger, D. (1985). *Intonation and Its Parts : Melody in Spoken English*. Stanford University Press.
- Bonaiuto, J., & Arbib, M. A. (2010). Extending the mirror neuron system model, II : What did I just do? A new role for mirror neurons. *Biological Cybernetics*, 102(4), 341-359.  
<https://doi.org/10.1007/s00422-010-0371-0>
- Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., Burack, J. A., & Mottron, L. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia*, 48(9), 2465-2475.  
<https://doi.org/10.1016/j.neuropsychologia.2010.04.020>
- Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A.-M. (2003). Enhanced pitch sensitivity in individuals with autism : A signal detection analysis. *Journal of Cognitive Neuroscience*, 15(2), 226-235. <https://doi.org/10.1162/089892903321208169>
- Boso, M., Emanuele, E., Minazzi, V., Abbamonte, M., & Politi, P. (2007). Effect of long-term interactive music therapy on behavior profile and musical skills in young adults with severe autism. *Journal of Alternative and Complementary Medicine (New York, N.Y.)*, 13(7), 709-712. <https://doi.org/10.1089/acm.2006.6334>
- Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody : Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology*, 41(2), 259-268. <https://doi.org/10.1111/j.1469-8986.2003.00142.x>
- Boxtel, A. (2010). Facial EMG as a tool for inferring affective states. In *Proceedings of measuring behavior*.
- Bradley, M., & Lang, P. (2007). *IADS-2*. NIMH Center for the Study of Emotion and Attention.
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 45(4), 602-607.  
<https://doi.org/10.1111/j.1469-8986.2008.00654.x>
- Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, 5(4), 1567-1573.  
<https://doi.org/10.1016/j.rasd.2011.03.002>

## Références

- Brewer, R., Biotti, F., Catmur, C., Press, C., Happé, F., Cook, R., & Bird, G. (2016). Can Neurotypical Individuals Read Autistic Facial Expressions? Atypical Production of Emotional Facial Expressions in Autism Spectrum Disorders. *Autism Research: Official Journal of the International Society for Autism Research*, 9(2), 262-271. <https://doi.org/10.1002/aur.1508>
- Briend, F., David, C., Silleresi, S., Malvy, J., Ferré, S., & Latinus, M. (2023). Voice acoustics allow classifying autism spectrum disorder with high accuracy. *Translational Psychiatry*, 13(1), 1-8. <https://doi.org/10.1038/s41398-023-02554-8>
- Brinkman, L., Goffin, S., van de Schoot, R., van Haren, N. E. M., Dotsch, R., & Aarts, H. (2019). Quantifying the informational value of classification images. *Behavior Research Methods*, 51(5), 2059-2073. <https://doi.org/10.3758/s13428-019-01232-2>
- Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations : A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, 28(1), 333-361. <https://doi.org/10.1080/10463283.2017.1381469>
- Brock, J. (2012). Alternative Bayesian accounts of autistic perception : Comment on Pellicano and Burr. *Trends Cogn. Sci*, 16, 573-574.
- Brooks, P. J., & Ploog, B. O. (2013). Attention to emotional tone of voice in speech perception in children with autism. *Research in Autism Spectrum Disorders*, 7(7), 845-857. <https://doi.org/10.1016/j.rasd.2013.03.003>
- Brown-Iannuzzi, J. L., Dotsch, R., Cooley, E., & Payne, B. K. (2017). The relationship between mental representations of welfare recipients and attitudes toward welfare. *Psychological Science*, 28(1), 92-103. <https://doi.org/10.1177/0956797616674999>
- Bryant, G., & Barrett, H. C. (2008). Vocal Emotion Recognition Across Disparate Cultures. *Journal of Cognition and Culture*, 8(1-2), 135-148. <https://doi.org/10.1163/156770908X289242>
- Bufo, M. R., Guidotti, M., De Faria, C., Mofid, Y., Bonnet-Brilhault, F., Wardak, C., & Aguillon-Hernandez, N. (2022). Autonomic tone in children and adults : Pupillary, electrodermal and cardiac activity at rest. *International Journal of Psychophysiology*, 180, 68-78. <https://doi.org/10.1016/j.ijpsycho.2022.07.009>
- Bullack, A., Büdenbender, N., Roden, I., & Kreutz, G. (2018). Psychophysiological Responses to “Happy” and “Sad” Music A Replication Study. *Music Perception*, 35(4), 502-517. <https://doi.org/10.1525/mp.2018.35.4.502>

## Références

- Burley, D. T., & Daughters, K. (2020). The effect of oxytocin on pupil response to naturalistic dynamic facial expressions. *Hormones and Behavior*, *125*, 104837. <https://doi.org/10.1016/j.yhbeh.2020.104837>
- Burred, J. J., Ponsot, E., Goupil, L., Liuni, M., & Aucouturier, J.-J. (2019). CLEESE : An open-source audio-transformation toolbox for data-driven experiments in speech and music cognition. *PLOS ONE*, *14*(4), e0205943. <https://doi.org/10.1371/journal.pone.0205943>
- Burrows, A. M., Waller, B. M., Parr, L. A., & Bonar, C. J. (2006). Muscles of facial expression in the chimpanzee (*Pan troglodytes*): Descriptive, comparative and phylogenetic contexts. *Journal of Anatomy*, *208*(2), 153-167. <https://doi.org/10.1111/j.1469-7580.2006.00523.x>
- Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). Psychophysiological science. *Handbook of psychophysiology*, *2*, 3-23.
- Cannon, W. B. (1927). The James-Lange Theory of Emotions : A Critical Examination and an Alternative Theory. *The American Journal of Psychology*, *39*(1/4), 106-124. <https://doi.org/10.2307/1415404>
- Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior*, *23*(1), 91-100. <https://doi.org/10.1023/A:1021339410262>
- Chabin, T., Gabriel, D., Comte, A., Haffen, E., Moulin, T., & Pazart, L. (2022). Interbrain emotional connection during music performances is driven by physical proximity and individual traits. *Annals of the New York Academy of Sciences*, *1508*(1), 178-195. <https://doi.org/10.1111/nyas.14711>
- Chan, M. M. Y., & Han, Y. M. Y. (2020). Differential mirror neuron system (MNS) activation during action observation with and without social-emotional components in autism : A meta-analysis of neuroimaging studies. *Molecular Autism*, *11*(1), 72. <https://doi.org/10.1186/s13229-020-00374-x>
- Charest, I., Pernet, C. R., Rousselet, G. A., Quiñones, I., Latinus, M., Fillion-Bilodeau, S., Chartrand, J.-P., & Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience*, *10*(1), 127. <https://doi.org/10.1186/1471-2202-10-127>
- Charpentier, J., Kovarski, K., Houy-Durand, E., Malvy, J., Saby, A., Bonnet-Brilhault, F., Latinus, M., & Gomot, M. (2018). Emotional prosodic change detection in autism Spectrum disorder : An electrophysiological investigation in children and adults. *Journal of Neurodevelopmental Disorders*, *10*(1), 28. <https://doi.org/10.1186/s11689-018-9246-9>

## Références

- Charpentier, J., Kovarski, K., Roux, S., Houy-Durand, E., Saby, A., Bonnet-Brilhault, F., Latinus, M., & Gomot, M. (2018). Brain mechanisms involved in angry prosody change detection in school-age children and adults, revealed by electrophysiology. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(4), 748-763. <https://doi.org/10.3758/s13415-018-0602-8>
- Charpentier, J., Latinus, M., Andersson, F., Saby, A., Cottier, J.-P., Bonnet-Brilhault, F., Houy-Durand, E., & Gomot, M. (2020). Brain correlates of emotional prosodic change detection in autism spectrum disorder. *NeuroImage: Clinical*, *28*, 102512. <https://doi.org/10.1016/j.nicl.2020.102512>
- Cherng, Y.-G., Baird, T., Chen, J.-T., & Wang, C.-A. (2020). Background luminance effects on pupil size associated with emotion and saccade preparation. *Scientific Reports*, *10*(1), Article 1. <https://doi.org/10.1038/s41598-020-72954-z>
- Chevallier, C., Noveck, I., Happé, F., & Wilson, D. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, *49*(3), 507-517. <https://doi.org/10.1016/j.neuropsychologia.2010.11.042>
- Chronaki, G., Broyd, S., Garner, M., Hadwin, J. A., Thompson, M. J. J., & Sonuga-Barke, E. J. S. (2012). Isolating N400 as neural marker of vocal anger processing in 6-11-year old children. *Dev. Cogn. Neurosci*, *2*, 268-276.
- Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, *8*(1), 8659. <https://doi.org/10.1038/s41598-018-26889-1>
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126-135. <https://doi.org/10.1016/j.brainres.2008.04.023>
- Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia*, *48*(1), 220-225. <https://doi.org/10.1016/j.neuropsychologia.2009.09.007>
- Cosme, G., Rosa, P. J., Lima, C. F., Tavares, V., Scott, S., Chen, S., Wilcockson, T. D. W., Crawford, T. J., & Prata, D. (2021). Pupil dilation reflects the authenticity of received nonverbal vocalizations. *Scientific Reports*, *11*(1), 1-14. <https://doi.org/10.1038/s41598-021-83070-x>
- Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental*

## Références

- Psychology: Learning, Memory, and Cognition*, 19(4), 909-921.  
<https://doi.org/10.1037/0278-7393.19.4.909>
- Crespo, C., Santos, S., Canavarro, M. C., Kielikowski, M., Pryor, J., & Féres-Carneiro, T. (2013). Family routines and rituals in the context of chronic conditions : A review. *International Journal of Psychology*, 48(5), 729-746.  
<https://doi.org/10.1080/00207594.2013.806811>
- Croft, P. G. (1961). Photomotor Reflex as an Indicator of Consciousness during Curarization. *Nature*, 190(4777), 726-726. <https://doi.org/10.1038/190726a0>
- Cummins, F. (2002). *Speech rhythm and rhythmic taxonomy*. 121-124.  
<https://doi.org/10.21437/SpeechProsody.2002-17>
- Damasio, A. R., & Blanc, M. (1995). *L'erreur de Descartes : La raison des émotions*. O. Jacob.
- Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others : Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9(1), 28-30.  
<https://doi.org/10.1038/nn1611>
- Darwin, C. (2009). *The Expression of the Emotions in Man and Animals, Anniversary Edition* (Anniversary ed édition). Oxford University Press Inc.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44(1), 113-126. <https://doi.org/10.1037/0022-3514.44.1.113>
- Davis, P. A. (1939). Effects of acoustic stimuli on the waking human brain. *Journal of Neurophysiology*, 2, 494-499.
- Dawson, G. D. (1954). A summation technique for the detection of small evoked potentials. *Electroencephalography and Clinical Neurophysiology*, 6, 65-84.  
[https://doi.org/10.1016/0013-4694\(54\)90007-3](https://doi.org/10.1016/0013-4694(54)90007-3)
- Dawson, G., Hill, D., Spencer, A., Galpert, L., & Watson, L. (1990). Affective exchanges between young autistic children and their mothers. *Journal of Abnormal Child Psychology*, 18(3), 335-345. <https://doi.org/10.1007/BF00916569>
- De Lucia, M., Clarke, S., & Murray, M. M. (2010). A Temporal Hierarchy for Conspecific Vocalization Discrimination in Humans. *Journal of Neuroscience*, 30(33), 11210-11221.  
<https://doi.org/10.1523/JNEUROSCI.2239-10.2010>
- Delobel-Ayoub, M., Saemundsen, E., Gissler, M., Ego, A., Moilanen, I., Ebeling, H., Rafnsson, V., Klapouszczak, D., Thorsteinsson, E., Arnaldsdóttir, K. M., Roge, B., Arnaud, C., &

## Références

- Schendel, D. (2020). Prevalence of Autism Spectrum Disorder in 7–9-Year-Old Children in Denmark, Finland, France and Iceland : A Population-Based Registries Approach Within the ASDEU Project. *Journal of Autism and Developmental Disorders*, 50(3), 949-959. <https://doi.org/10.1007/s10803-019-04328-y>
- Demos, K. E., Kelley, W. M., Ryan, S. L., Davis, F. C., & Whalen, P. J. (2008). Human Amygdala Sensitivity to the Pupil Size of Others. *Cerebral Cortex*, 18(12), 2729-2734. <https://doi.org/10.1093/cercor/bhn034>
- Dietrich, S., Ackermann, H., Szameitat, D. P., & Alter, K. (2006). Psychoacoustic studies on the processing of vocal interjections : How to disentangle lexical and prosodic information? *Progress in Brain Research*, 156, 295-302. [https://doi.org/10.1016/S0079-6123\(06\)56016-9](https://doi.org/10.1016/S0079-6123(06)56016-9)
- Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, 39(1), 39-45. <https://doi.org/10.1111/1467-9450.00054>
- Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science*, 11(1), 86-89. <https://doi.org/10.1111/1467-9280.00221>
- Dondi, M., Agnoli, S., & Franchin, L. (2010). A new look at the very early origins of smiling. In *Origins as a paradigm in the sciences and in the humanities* (p. 131-146).
- Dotsch, R., Wigboldus, D. H. J., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, 19(10), 978-980. <https://doi.org/10.1111/j.1467-9280.2008.02186.x>
- Dotsch, R., Wigboldus, D. H. J., & van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Personality and Social Psychology*, 100(6), 999-1014. <https://doi.org/10.1037/a0023026>
- Drahota, A., Costall, A., & Reddy, V. (2008). The vocal communication of different kinds of smile. *Speech Communication*, 50(4), 278-287. <https://doi.org/10.1016/j.specom.2007.10.001>
- Duchenne (de Boulogne), G.-B. (1876). *Mécanisme de la physionomie humaine. Deuxième édition* ((2)). Librairie J.-B. Bailliere et Fils. <http://vlp.mpiwg-berlin.mpg.de/references?id=lit38953>
- Edgerton, C. L. (1994). The Effect of Improvisational Music Therapy on the Communicative Behaviors of Autistic Children. *J. Music Ther*, 31, 31-62.

## Références

- Eibl-Eibesfeldt, I. (2017). *Human ethology*. Routledge.
- Eigsti, I.-M. (2013). A Review of Embodiment in Autism Spectrum Disorders. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00224>
- Eigsti, I.-M., & Fein, D. A. (2013). More is less : Pitch discrimination and language delays in children with optimal outcomes from autism. *Autism research : official journal of the International Society for Autism Research*, 6(6), 10.1002/aur.1324. <https://doi.org/10.1002/aur.1324>
- Eigsti, I.-M., Schuh, J., Mencl, E., Schultz, R. T., & Paul, R. (2012). The neural underpinnings of prosody in autism. *Child neuropsychology: a journal on normal and abnormal development in childhood and adolescence*, 18(6), 600-617. <https://doi.org/10.1080/09297049.2011.639757>
- Ekman, P., Davidson, R. J., & Friesen, W. V. (1990). The Duchenne smile : Emotional expression and brain physiology: II. *Journal of Personality and Social Psychology*, 58(2), 342-353. <https://doi.org/10.1037/0022-3514.58.2.342>
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face* Englewood cliffs. *Spectrum-Prentice Hall, New Jersey*.
- Ekman, P., & Keltner, D. (1997). Universal facial expressions of emotion : An old controversy and new findings. In *Nonverbal communication : Where nature meets culture* (p. 27-46). Lawrence Erlbaum Associates, Inc.
- Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science (New York, N.Y.)*, 164(3875), 86-88. <https://doi.org/10.1126/science.164.3875.86>
- El Haddad, K., Torre, I., Gilmartin, E., Çakmak, H., Dupont, S., Dutoit, T., & Campbell, N. (2017). Introducing AmuS : The Amused Speech Database. *Statistical Language and Speech Processing: 5th International Conference, SLSP 2017, Le Mans, France, October 23-25, 2017, Proceedings*, 229-240. [https://doi.org/10.1007/978-3-319-68456-7\\_19](https://doi.org/10.1007/978-3-319-68456-7_19)
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition : A meta-analysis. *Psychological Bulletin*, 128(2), 203-235. <https://doi.org/10.1037/0033-2909.128.2.203>
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., Reiterer, S., Grodd, W., & Wildgruber, D. (2006). Impact of voice on emotional judgment of faces : An event-related fMRI study. *Human Brain Mapping*, 27(9), 707-714. <https://doi.org/10.1002/hbm.20212>

## Références

- Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4), Article 4. <https://doi.org/10.1038/nrn2258>
- Fan, Y.-T., & Cheng, Y. (2014). Atypical Mismatch Negativity in Response to Emotional Voices in People with Autism Spectrum Conditions. *PLoS ONE*, 9(7), e102471. <https://doi.org/10.1371/journal.pone.0102471>
- Fawcett, C., Arslan, M., Falck-Ytter, T., Roeyers, H., & Gredebäck, G. (2017). Human eyes with dilated pupils induce pupillary contagion in infants. *Scientific Reports*, 7(1), 9601. <https://doi.org/10.1038/s41598-017-08223-3>
- Fawcett, C., Wesevich, V., & Gredebäck, G. (2016). Pupillary Contagion in Infancy : Evidence for Spontaneous Transfer of Arousal. *Psychological Science*, 27(7), 997-1003. <https://doi.org/10.1177/0956797616643924>
- Fecteau, S., Armony, J. L., Joanette, Y., & Belin, P. (2005). Sensitivity to voice in human prefrontal cortex. *Journal of Neurophysiology*, 94(3), 2251-2254. <https://doi.org/10.1152/jn.00329.2005>
- Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage*, 36(2), 480-487. <https://doi.org/10.1016/j.neuroimage.2007.02.043>
- Ferré, S., Mahé, O., Manenti, M., Pimenta, H., & Prévost, P. (2022). *Détecter un déficit phonologique chez les adultes—Un module adulte pour LITMUS-QU-NWR-FR*. 280-288. <https://doi.org/10.21437/JEP.2022-30>
- Filipe, M. G., Branco, P., Frota, S., Castro, S. L., & Vicente, S. G. (2015). Affective prosody in European Portuguese : Perceptual and acoustic characterization of one-word utterances. *Speech Communication*, 67, 58-64. <https://doi.org/10.1016/j.specom.2014.09.007>
- Filippi, P., Ocklenburg, S., Bowling, D. L., Heege, L., Güntürkün, O., Newen, A., & de Boer, B. (2017). More than words (and faces) : Evidence for a Stroop effect of prosody in emotion word processing. *Cognition and Emotion*, 31(5), 879-891. <https://doi.org/10.1080/02699931.2016.1177489>
- Fischer, A., & LaFrance, M. (2015). What Drives the Smile and the Tear : Why Women Are More Emotionally Expressive Than Men. *Emotion Review*, 7(1), 22-29. <https://doi.org/10.1177/1754073914544406>
- Fombonne, E. (2018). Editorial : The rising prevalence of autism. *Journal of Child Psychology and Psychiatry*, 59(7), 717-720. <https://doi.org/10.1111/jcpp.12941>

## Références

- Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, 23(5), 567-589. <https://doi.org/10.1111/j.14698986.1986.tb00676.x>
- Frijda, N. H., & Mesquita, B. (1994). The social roles and functions of emotions. In *Emotion and culture : Empirical studies of mutual influence* (p. 51-87). American Psychological Association. <https://doi.org/10.1037/10152-002>
- Friston, K. (2003). Learning and inference in the brain. *Neural Netw. Off. J. Int. Neural Netw. Soc*, 16, 1325-1352.
- Friston, K. (2016). The Bayesian Savant. *Biological Psychiatry*, 80(2), 87-89. <https://doi.org/10.1016/j.biopsych.2016.05.006>
- Frith, U. (2003). *Autism : Explaining the enigma*. Blackwell publishing.
- Frühholz, S., Hofstetter, C., Cristinzio, C., Saj, A., Seeck, M., Vuilleumier, P., & Grandjean, D. (2015). Asymmetrical effects of unilateral right or left amygdala damage on auditory cortical processing of vocal emotions. *Proceedings of the National Academy of Sciences*, 112(5), 1583-1588. <https://doi.org/10.1073/pnas.1411315112>
- Fry, D. B. (1955). Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal of the Acoustical Society of America*, 27(4), 765-768. <https://doi.org/10.1121/1.1908022>
- Gabriels, R. L., Cuccaro, M. L., Hill, D. E., Ivers, B. J., & Goldson, E. (2005). Repetitive behaviors in autism : Relationships with associated clinical features. *Research in Developmental Disabilities*, 26(2), 169-181. <https://doi.org/10.1016/j.ridd.2004.05.003>
- Galazka, M. A., Åsberg Johnels, J., Zürcher, N. R., Hippolyte, L., Lemonnier, E., Billstedt, E., Gillberg, C., & Hadjikhani, N. (2019). Pupillary Contagion in Autism. *Psychological Science*, 30(2), 309-315. <https://doi.org/10.1177/0956797618809382>
- Gallese, V. (2005). Embodied simulation : From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences*, 4(1), 23-48. <https://doi.org/10.1007/s11097-005-4737-z>
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2(12), 493-501. [https://doi.org/10.1016/s1364-6613\(98\)01262-5](https://doi.org/10.1016/s1364-6613(98)01262-5)
- Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant Pupil Diameter Changes in Response to Others' Positive and Negative Emotions. *PLoS ONE*, 6(11), e27132. <https://doi.org/10.1371/journal.pone.0027132>

## Références

- Geelhand, P., Papastamou, F., Belenger, M., Clin, E., Hickman, L., Keating, C. T., & Sowden, S. (2023). Autism-Related Language Preferences of French-Speaking Autistic Adults : An Online Survey. *Autism in Adulthood*. <https://doi.org/10.1089/aut.2022.0056>
- Geretsegger, M., Elefant, C., Mössler, K. A., & Gold, C. (2014). Music therapy for people with autism spectrum disorder. *The Cochrane Database of Systematic Reviews*, 2014(6), CD004381. <https://doi.org/10.1002/14651858.CD004381.pub3>
- Geretsegger, M., Fusar-Poli, L., Elefant, C., Mössler, K. A., Vitale, G., & Gold, C. (2022). Music therapy for autistic people. *Cochrane Database of Systematic Reviews*, 5. <https://doi.org/10.1002/14651858.CD004381.pub4>
- Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthélémy, C., Brunelle, F., Samson, Y., & Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience*, 7(8), 801-802. <https://doi.org/10.1038/nn1291>
- Geschwind, D. H. (2011). Genetics of Autism Spectrum Disorders. *Trends in cognitive sciences*, 15(9), 409-416. <https://doi.org/10.1016/j.tics.2011.07.003>
- Ghasemtabar, S. N., Hosseini, M., Fayyaz, I., Arab, S., Naghashian, H., & Poudineh, Z. (2015). Music therapy : An effective approach in improving social skills of children with autism. *Advanced Biomedical Research*, 4, 157. <https://doi.org/10.4103/2277-9175.161584>
- Gibson, J. J. (1986). *The Ecological Approach to Visual Perception* (1st edition). Routledge.
- Giganti, F., & Esposito Ziello, M. (2009). Contagious and spontaneous yawning in autistic and typically developing children. *Current Psychology Letters. Behaviour, Brain & Cognition*, Vol. 25, Issue 1, 2009, Article Vol. 25, Issue 1, 2009. <https://journals.openedition.org/cpl/4810>
- Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'Reading the Mind in the Voice' Test-Revised : A Study of Complex Emotion Recognition in Adults with and Without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, 37(6), 1096-1106. <https://doi.org/10.1007/s10803-006-0252-5>
- Gomot, M., & Wicker, B. (2012). A challenging, unpredictable world for people with autism spectrum disorder. *Int. J. Psychophysiol. Off. J. Int. Organ. Psychophysiol*, 83, 240-247.
- Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-020-20649-4>

## Références

- Grandjean, D., Sander, D., Lucas, N., Scherer, K. R., & Vuilleumier, P. (2008). Effects of emotional prosody on auditory extinction for voices in patients with spatial neglect. *Neuropsychologia*, 46(2), 487-496. <https://doi.org/10.1016/j.neuropsychologia.2007.08.025>
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, 8(2), 145-146. <https://doi.org/10.1038/nn1392>
- Green, B., Jääskeläinen, I. P., Sams, M., & Rauschecker, J. P. (2018). Distinct brain areas process novel and repeating tone sequences. *Brain and Language*, 187, 104-114. <https://doi.org/10.1016/j.bandl.2018.09.006>
- Grossman, R. B. (2015). Judgments of social awkwardness from brief exposure to children with and without high-functioning autism. *Autism*, 19(5), 580-587. <https://doi.org/10.1177/1362361314536937>
- Grossman, R. B., Bemis, R. H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *J. Speech Lang. Hear. Res. JSLHR*, 53, 778-793.
- Grossman, R. B., Edelson, L. R., & Tager-Flusberg, H. (2013). Emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. *Journal of Speech, Language, and Hearing Research: JSLHR*, 56(3), 1035-1044. [https://doi.org/10.1044/1092-4388\(2012/12-0067\)](https://doi.org/10.1044/1092-4388(2012/12-0067))
- Grossmann, T., Striano, T., & Friederici, A. D. (2005). Infants' electric brain responses to emotional prosody. *NeuroReport: For Rapid Communication of Neuroscience Research*, 16(16), 1825-1828. <https://doi.org/10.1097/01.wnr.0000185964.34336.b1>
- Haag, C. (2019). *La Contagion émotionnelle*. Albin Michel. <https://hal.science/hal-02298302>
- Haddad, K. E. (2015). *Towards a Speech Synthesis System with Controllable Amusement Levels*. [https://www.academia.edu/66322194/Towards\\_a\\_Speech\\_Synthesis\\_System\\_with\\_Controllable\\_Amusement\\_Levels](https://www.academia.edu/66322194/Towards_a_Speech_Synthesis_System_with_Controllable_Amusement_Levels)
- Haker, H., Schneebeli, M., & Stephan, K. E. (2016). Can Bayesian Theories of Autism Spectrum Disorder Help Improve Clinical Practice? *Frontiers in Psychiatry*, 7. <https://www.frontiersin.org/articles/10.3389/fpsy.2016.00107>

## Références

- Hamilton, A. F. de C. (2008). Emulation and Mimicry for Social Interaction : A Theoretical Approach to Imitation in Autism. *Quarterly Journal of Experimental Psychology*, 61(1), 101-115. <https://doi.org/10.1080/17470210701508798>
- Hamilton, A. F. de C. (2013). Reflecting on the mirror neuron system in autism : A systematic review of current theories. *Developmental Cognitive Neuroscience*, 3, 91-105. <https://doi.org/10.1016/j.dcn.2012.09.008>
- Hammerschmidt, K., & Jürgens, U. (2007). Acoustical Correlates of Affective Prosody. *Journal of Voice*, 21(5), 531-540. <https://doi.org/10.1016/j.jvoice.2006.03.002>
- Happé, F., & Frith, U. (2006). The Weak Coherence Account : Detail-focused Cognitive Style in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 36(1), 5-25. <https://doi.org/10.1007/s10803-005-0039-0>
- Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial Emotion Recognition in Autism Spectrum Disorders : A Review of Behavioral and Neuroimaging Studies. *Neuropsychology Review*, 20(3), 290-322. <https://doi.org/10.1007/s11065-010-9138-6>
- Harrison, N. A., Gray, M. A., & Critchley, H. D. (2009). Dynamic pupillary exchange engages brain regions encoding social salience. *Social Neuroscience*, 4(3), 233-243. <https://doi.org/10.1080/17470910802553508>
- Harrison, N. A., Singer, T., Rotshtein, P., Dolan, R. J., & Critchley, H. D. (2006). Pupillary contagion : Central mechanisms engaged in sadness processing. *Social Cognitive and Affective Neuroscience*, 1(1), 5-17. <https://doi.org/10.1093/scan/nsl006>
- Harrison, N. A., Wilson, C. E., & Critchley, H. D. (2007). Processing of observed pupil size modulates perception of sadness and predicts empathy. *Emotion*, 7(4), 724-729. <https://doi.org/10.1037/1528-3542.7.4.724>
- Hatfield, E., Bensman, L., Thornton, P. D., & Rapson, R. L. (2014). New Perspectives on Emotional Contagion : A Review of Classic and Recent Research on Facial Mimicry and Contagion. *Interpersona: An International Journal on Personal Relationships*, 8(2), Article 2. <https://doi.org/10.5964/ijpr.v8i2.162>
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1992). Primitive emotional contagion. In *Emotion and social behavior* (p. 151-177). Sage Publications, Inc.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). *Emotional Contagion*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139174138>

## Références

- Heaton, P., Davis, R. E., & Happé, F. G. E. (2008). Research note : Exceptional absolute pitch perception for spoken words in an able adult with autism. *Neuropsychologia*, *46*(7), 2095-2098. <https://doi.org/10.1016/j.neuropsychologia.2008.02.006>
- Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior discrimination of speech pitch and its relationship to verbal ability in autism spectrum disorders. *Cognitive Neuropsychology*, *25*(6), 771-782. <https://doi.org/10.1080/02643290802336277>
- Helt, M. S., de Marchena, A. B., Schineller, M. E., Kirk, A. I., Scheub, R. J., & Sorensen, T. M. (2021). Contagious itching is heightened in children with autism spectrum disorders. *Developmental Science*, *24*(2), e13024. <https://doi.org/10.1111/desc.13024>
- Helt, M. S., Eigsti, I.-M., Snyder, P. J., & Fein, D. A. (2010). Contagious Yawning in Autistic and Typical Development : Contagious Yawning. *Child Development*, *81*(5), 1620-1631. <https://doi.org/10.1111/j.1467-8624.2010.01495.x>
- Helt, M. S., Fein, D. A., & Vargas, J. E. (2020). Emotional contagion in children with autism spectrum disorder varies with stimulus familiarity and task instructions. *Development and Psychopathology*, *32*(1), 383-393. <https://doi.org/10.1017/S0954579419000154>
- Helt, M. S., Sorensen, T. M., Scheub, R. J., Nakhle, M. B., & Luddy, A. C. (2021). Patterns of Contagious Yawning and Itching Differ Amongst Adults With Autistic Traits vs. Psychopathic Traits. *Frontiers in Psychology*, *12*, 645310. <https://doi.org/10.3389/fpsyg.2021.645310>
- Hesling, I., Dilharreguy, B., Peppé, S., Amirault, M., Bouvard, M., & Allard, M. (2010). The Integration of Prosodic Speech in High Functioning Autism : A Preliminary fMRI Study. *PLoS ONE*, *5*(7), e11571. <https://doi.org/10.1371/journal.pone.0011571>
- Hess, E. H. (1965). ATTITUDE AND PUPIL SIZE. *Scientific American*, *212*(4), 46-55. JSTOR.
- Hess, E. H., & Polt, J. M. (1964). Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science*, *143*(3611), 1190-1192. <https://doi.org/10.1126/science.143.3611.1190>
- Hess, U., Beaupré, M. G., & Cheung, N. (2002). Who to whom and why – cultural differences and similarities in the function of smiles. In *An Empirical Reflection on the Smile* (p. 187-216).
- Hess, U., & Bourgeois, P. (2010). You smile–I smile : Emotion expression in social interaction. *Biological Psychology*, *84*(3), 514-520. <https://doi.org/10.1016/j.biopsycho.2009.11.001>

## Références

- Hess, U., & Fischer, A. (2014). Emotional Mimicry : Why and When We Mimic Emotions. *Social and Personality Psychology Compass*, 8(2), 45-57. <https://doi.org/10.1111/spc3.12083>
- Hess, U., Houde, S., & Fischer, A. (2014). Do we mimic what we see or what we know? In C. von Scheve & M. Salmela (Éds.), *Collective Emotions* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199659180.003.0007>
- Hietanen, J. K., Surakka, V., & Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology*, 35(5), 530-536. <https://doi.org/10.1017/S0048577298970445>
- Hobson, R. P. (2019). *Autism and the Development of Mind*. Routledge. <https://doi.org/10.4324/9780203765272>
- Hoffmann, E., Brück, C., Kreifelts, B., Ethofer, T., & Wildgruber, D. (2016). Reduced functional connectivity to the frontal cortex during processing of social cues in autism spectrum disorder. *Journal of Neural Transmission (Vienna, Austria: 1996)*, 123(8), 937-947. <https://doi.org/10.1007/s00702-016-1544-3>
- Holochwost, S. J., & Izard, C. E. (2008). Evidence from young children regarding emotional responses to music. *Behavioral and Brain Sciences*, 31(5), 581-582. <https://doi.org/10.1017/S0140525X08005360>
- Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). The window of my eyes : Task disengagement and mental fatigue covary with pupil dynamics. *Biological Psychology*, 110, 100-106. <https://doi.org/10.1016/j.biopsycho.2015.06.013>
- Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder : Affective prosody in ASD. *Autism Research*, 10(12), 1991-2001. <https://doi.org/10.1002/aur.1847>
- Hull, L., Petrides, K. V., Allison, C., Smith, P., Baron-Cohen, S., Lai, M.-C., & Mandy, W. (2017). « Putting on My Best Normal » : Social Camouflaging in Adults with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, 47(8), 2519-2534. <https://doi.org/10.1007/s10803-017-3166-5>
- Hyman, S. L., Levy, S. E., Myers, S. M., & COUNCIL ON CHILDREN WITH DISABILITIES, SECTION ON DEVELOPMENTAL AND BEHAVIORAL PEDIATRICS. (2020). Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. *Pediatrics*, 145(1), e20193447. <https://doi.org/10.1542/peds.2019-3447>

## Références

- Iacoboni, M. (2009). Imitation, Empathy, and Mirror Neurons. *Annual Review of Psychology*, 60(Volume 60, 2009), 653-670. <https://doi.org/10.1146/annurev.psych.60.110707.163604>
- Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, 7(12), 942-951. <https://doi.org/10.1038/nrn2024>
- Imhoff, R., Dotsch, R., Bianchi, M., Banse, R., & Wigboldus, D. H. J. (2011). Facing Europe : Visualizing spontaneous in-group projection. *Psychological Science*, 22(12), 1583-1590. <https://doi.org/10.1177/0956797611419675>
- Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face ! Visual encoding of stereotype content. *Frontiers in Psychology*, 4. <https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00386>
- Izard, C. E. (1971). *The face of emotion* (p. xii, 468). Appleton-Century-Crofts.
- Izard, C. E. (1977). *Human Emotions* (1977e édition). Kluwer Academic/Plenum Publishers.
- Jaaskelainen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levanen, S., Lin, F.-H., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H., & Belliveau, J. W. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences*, 101(17), 6809-6814. <https://doi.org/10.1073/pnas.0303760101>
- James, W. (1884). What is an Emotion? *Mind*, 9(34), 188-205.
- Jansson-Verkasalo, E., Ceponiene, R., Kielinen, M., Suominen, K., Jäntti, V., Linna, S.-L., Moilanen, I., & Näätänen, R. (2003). Deficient auditory processing in children with Asperger Syndrome, as indexed by event-related potentials. *Neuroscience Letters*, 338(3), 197-200. [https://doi.org/10.1016/S0304-3940\(02\)01405-2](https://doi.org/10.1016/S0304-3940(02)01405-2)
- Jin, A. B., Steding, L. H., & Webb, A. K. (2015). Reduced emotional and cardiovascular reactivity to emotionally evocative stimuli in major depressive disorder. *International Journal of Psychophysiology*, 97(1), 66-74. <https://doi.org/10.1016/j.ijpsycho.2015.04.014>
- Jones, E. J. H., Gliga, T., Bedford, R., Charman, T., & Johnson, M. H. (2014). Developmental pathways to autism : A review of prospective studies of infants at risk. *Neuroscience & Biobehavioral Reviews*, 39, 1-33. <https://doi.org/10.1016/j.neubiorev.2013.12.001>
- Jürgens, R., Fischer, J., & Schacht, A. (2018). Hot Speech and Exploding Bombs : Autonomic Arousal During Emotion Classification of Prosodic Utterances and Affective Sounds. *Frontiers in Psychology*, 9. <https://doi.org/10.3389/fpsyg.2018.00228>

## Références

- Juslin, P. N. (2013). From everyday emotions to aesthetic emotions : Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235-266. <https://doi.org/10.1016/j.plrev.2013.05.008>
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance : Different channels, same code? *Psychological Bulletin*, 129(5), 770-814. <https://doi.org/10.1037/0033-2909.129.5.770>
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music : The need to consider underlying mechanisms. *The Behavioral and Brain Sciences*, 31(5), 559-575; discussion 575-621. <https://doi.org/10.1017/S0140525X08005293>
- Kaczmarek, L. D., Behnke, M., Kashdan, T. B., Kusiak, A., Marzec, K., Mistrzak, M., & Włodarczyk, M. (2018). Smile intensity in social networking profile photographs is related to greater scientific achievements. *The Journal of Positive Psychology*, 13(5), 435-439. <https://doi.org/10.1080/17439760.2017.1326519>
- Kamiloğlu, R. G., Fischer, A. H., & Sauter, D. A. (2020). Good vibrations : A review of vocal expressions of positive emotions. *Psychonomic Bulletin & Review*, 27(2), 237-265. <https://doi.org/10.3758/s13423-019-01701-x>
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217-250.
- Kanski, J. J., & Bowling, B. (2015). *Kanski's Clinical Ophthalmology E-Book : A Systematic Approach*. Elsevier Health Sciences.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area : A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, 17(11), 4302. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- Kargas, N., López, B., Morris, P., & Reddy, V. (2016). Relations Among Detection of Syllable Stress, Speech Abnormalities, and Communicative Ability in Adults With Autism Spectrum Disorders. *Journal of Speech, Language, and Hearing Research: JSLHR*, 59(2), 206-215. [https://doi.org/10.1044/2015\\_JSLHR-S-14-0237](https://doi.org/10.1044/2015_JSLHR-S-14-0237)
- Kasai, K., Hashimoto, O., Kawakubo, Y., Yumoto, M., Kamio, S., & Itoh, K. (2005). Delayed automatic detection of change in speech sounds in adults with autism : A magnetoencephalographic study. *Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol*, 116, 1655-1664.
- Keysers, C., & Gazzola, V. (2010). Social Neuroscience : Mirror Neurons Recorded in Humans. *Current Biology*, 20(8), R353-R354. <https://doi.org/10.1016/j.cub.2010.03.013>

## Références

- Kim, J., Wigram, T., & Gold, C. (2008). The effects of improvisational music therapy on joint attention behaviors in autistic children: A randomized controlled study. *Journal of Autism and Developmental Disorders*, 38(9), 1758-1766. <https://doi.org/10.1007/s10803-008-0566-6>
- Korb, S., & Delplanque, S. (2013). Chapitre 8. Électromyographie (EMG): In *Electrophysiologie de la cognition* (p. 179-194). Dunod. <https://doi.org/10.3917/dunod.hot.2013.01.0179>
- Korpilahti, P., Jansson-Verkasalo, E., Mattila, M.-L., Kuusikko, S., Suominen, K., Rytty, S., Pauls, D. L., & Moilanen, I. (2006). Processing of Affective Speech Prosody is Impaired in Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 37(8), 1539-1549. <https://doi.org/10.1007/s10803-006-0271-2>
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: An event-related fMRI study. *NeuroImage*, 37(4), 1445-1456. <https://doi.org/10.1016/j.neuroimage.2007.06.020>
- Kret, M. E., & De Dreu, C. K. W. (2017). Pupil-mimicry conditions trust in partners: Moderation by oxytocin and group membership. *Proceedings of the Royal Society B: Biological Sciences*, 284(1850), 20162554. <https://doi.org/10.1098/rspb.2016.2554>
- Kret, M. E., Fischer, A. H., & De Dreu, C. K. W. (2015). Pupil Mimicry Correlates With Trust in In-Group Partners With Dilating Pupils. *Psychological Science*, 26(9), 1401-1410. <https://doi.org/10.1177/0956797615588306>
- Kret, M. E., Tomonaga, M., & Matsuzawa, T. (2014). Chimpanzees and Humans Mimic Pupil-Size of Conspecifics. *PLOS ONE*, 9(8), e104886. <https://doi.org/10.1371/journal.pone.0104886>
- Krolak-Salmon, P., Hénaff, M. A., Bertrand, O., Vighetto, A., & Mauguière, F. (2006). Les visages et leurs émotions: Partie II: La reconnaissance des expressions faciales. *Revue Neurologique*, 162(11), 1047-1058. [https://doi.org/10.1016/S0035-3787\(06\)75117-7](https://doi.org/10.1016/S0035-3787(06)75117-7)
- Kuchinke, L., Schneider, D., Kotz, S. A., & Jacobs, A. M. (2011). Spontaneous but not explicit processing of positive sentences impaired in Asperger's syndrome: Pupillometric evidence. *Neuropsychologia*, 49(3), 331-338. <https://doi.org/10.1016/j.neuropsychologia.2010.12.026>
- Kujala, T., Aho, E., Lepistö, T., Jansson-Verkasalo, E., Nieminen-von Wendt, T., von Wendt, L., & Näätänen, R. (2007). Atypical pattern of discriminating sound features in adults with Asperger syndrome as reflected by the mismatch negativity. *Biological Psychology*, 75(1), 109-114. <https://doi.org/10.1016/j.biopsycho.2006.12.007>

## Références

- Lacroix, A., Harquel, S., Mermillod, M., Garrido, M., Barbosa, L., Vercueil, L., Aleysson, D., Dutheil, F., Kovarski, K., & Gomot, M. (2024). Sex modulation of faces prediction error in the autistic brain. *Communications Biology*, 7(1), 1-12. <https://doi.org/10.1038/s42003-024-05807-4>
- LaFrance, M. (2011). *Lip Service : Smiles in Life, Death, Trust, Lies, Work, Memory, Sex, and Politics*. W. W. Norton & Company.
- LaGasse, A. B. (2014). Effects of a music therapy group intervention on enhancing social skills in children with autism. *Journal of Music Therapy*, 51(3), 250-275. <https://doi.org/10.1093/jmt/thu012>
- Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism : Setting the Scene for Future Research. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(1), 11-24. <https://doi.org/10.1016/j.jaac.2014.10.003>
- Lai, M.-C., Lombardo, M. V., Chakrabarti, B., Ruigrok, A. N., Bullmore, E. T., Suckling, J., Auyeung, B., Happé, F., Szatmari, P., Baron-Cohen, S., Bailey, A. J., Bolton, P. F., Carrington, S., Catani, M., Craig, M. C., Daly, E. M., Deoni, S. C., Ecker, C., Henty, J., ... Williams, S. C. (2019). Neural self-representation in autistic women and association with 'compensatory camouflaging'. *Autism*, 23(5), 1210-1223. <https://doi.org/10.1177/1362361318807159>
- Lai, M.-C., Lombardo, M. V., Ruigrok, A. N., Chakrabarti, B., Auyeung, B., Szatmari, P., Happé, F., & Baron-Cohen, S. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism*, 21(6), 690-702. <https://doi.org/10.1177/1362361316671012>
- Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion*, 12(3), 529-539. <https://doi.org/10.1037/a0026827>
- Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition : Impact of sensory modality and emotional category. *Cognition and Emotion*, 28(3), 452-469. <https://doi.org/10.1080/02699931.2013.837378>
- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures : Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3), 261-273. <https://doi.org/10.1111/j.1469-8986.1993.tb03352.x>
- Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, 40(5), 776-785. <https://doi.org/10.1111/1469-8986.00078>

## Références

- Lasarczyk, E., & Trouvain, J. (2008). Spread Lips + Raised Larynx + Higher F<sub>0</sub> = Smiled Speech? – An Articulatory Synthesis Approach. *Proceedings of ISSP 2008 - 8th International Seminar on Speech Production*.
- Latinus, M., & Belin, P. (2011). Human voice perception. *Curr Biol*, *21*(4), 143-145. <https://doi.org/10.1016/j.cub.2010.12.033>.
- Latinus, M., Mofid, Y., Kovarski, K., Charpentier, J., Batty, M., & Bonnet-Brilhault, F. (2019). Atypical Sound Perception in ASD Explained by Inter-Trial (In)consistency in EEG. *Frontiers in Psychology*, *10*. <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01177>
- Laukka, P. (2005). Categorical perception of vocal emotion expressions. *Emotion*, *5*(3), 277-295. <https://doi.org/10.1037/1528-3542.5.3.277>
- Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, *8*. <https://doi.org/10.3389/fnhum.2014.00302>
- Ledoux, J. (2005). *Le Cerveau des émotions*. Odile Jacob.
- Leekam, S. R., Prior, M. R., & Uljarevic, M. (2011). Restricted and repetitive behaviors in autism spectrum disorders : A review of research in the last decade. *Psychological Bulletin*, *137*(4), 562-593. <https://doi.org/10.1037/a0023341>
- Legris, E., Galvin, J., Mofid, Y., Aguillon-Hernandez, N., Roux, S., Aoustin, J.-M., Gomot, M., & Bakhos, D. (2022). Relationship between Behavioral and Objective Measures of Sound Intensity in Normal-Hearing Listeners and Hearing-Aid Users : A Pilot Study. *Brain Sciences*, *12*(3), Article 3. <https://doi.org/10.3390/brainsci12030392>
- Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, *1066*(1-2), 147-157. <https://doi.org/10.1016/j.brainres.2005.10.052>
- Lepistö, T., Nieminen-von Wendt, T., von Wendt, L., Näätänen, R., & Kujala, T. (2007). Auditory cortical change detection in adults with Asperger syndrome. *Neuroscience Letters*, *414*(2), 136-140. <https://doi.org/10.1016/j.neulet.2006.12.009>
- Lepistö, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Näätänen, R., & Kujala, T. (2006). Auditory perception and attention as reflected by the brain event-related potentials in children with Asperger syndrome. *Clinical Neurophysiology*, *117*(10), 2161-2171. <https://doi.org/10.1016/j.clinph.2006.06.709>

## Références

- Leung, F. Y. N., Sin, J., Dawson, C., Ong, J. H., Zhao, C., Veić, A., & Liu, F. (2022). Emotion recognition across visual and auditory modalities in autism spectrum disorder : A systematic review and meta-analysis. *Developmental Review*, 63, 101000. <https://doi.org/10.1016/j.dr.2021.101000>
- Likowski, K. U., Muehlberger, A., Gerdes, A. B., Wieser, M. J., Pauli, P., & Weyers, P. (2012). Facial mimicry and the mirror neuron system : Simultaneous acquisition of facial electromyography and functional magnetic resonance imaging. *Frontiers in Human Neuroscience*, 6. <https://doi.org/10.3389/fnhum.2012.00214>
- Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear : Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, 11, 1021-1031. <https://doi.org/10.1037/a0024521>
- Livingston, L. A., & Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders : Reflections from autism spectrum disorder. *Neuroscience & Biobehavioral Reviews*, 80, 729-742. <https://doi.org/10.1016/j.neubiorev.2017.06.005>
- Livingston, L. A., Shah, P., & Happé, F. (2019). Compensatory strategies below the behavioural surface in autism : A qualitative study. *The Lancet. Psychiatry*, 6(9), 766-777. [https://doi.org/10.1016/S2215-0366\(19\)30224-X](https://doi.org/10.1016/S2215-0366(19)30224-X)
- Lockwood Estrin, G., Milner, V., Spain, D., Happé, F., & Colvert, E. (2021). Barriers to Autism Spectrum Disorder Diagnosis for Young Women and Girls : A Systematic Review. *Review Journal of Autism and Developmental Disorders*, 8(4), 454-470. <https://doi.org/10.1007/s40489-020-00225-8>
- Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *Lancet (London, England)*, 392(10146), 508-520. [https://doi.org/10.1016/S0140-6736\(18\)31129-2](https://doi.org/10.1016/S0140-6736(18)31129-2)
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic : A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, 30(3), 205-223. <https://doi.org/10.1023/A:1005592401947>
- Maenner, M. J. (2021). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. *MMWR. Surveillance Summaries*, 70. <https://doi.org/10.15585/mmwr.ss7011a1>
- Maenner, M. J. (2023). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11

## Références

- Sites, United States, 2020. *MMWR. Surveillance Summaries*, 72. <https://doi.org/10.15585/mmwr.ss7202a1>
- Magnée, M. J. C. M., Stekelenburg, J. J., Kemner, C., & de Gelder, B. (2007). Similar facial electromyographic responses to faces, voices, and body expressions: *NeuroReport*, 18(4), 369-372. <https://doi.org/10.1097/WNR.0b013e32801776e6>
- Manenti, M., Tuller, L., Houy-Durand, E., Bonnet-Brilhault, F., & Prévost, P. (2023). Assessing structural language skills of autistic adults : Focus on sentence repetition. *Lingua*, 294, 103598. <https://doi.org/10.1016/j.lingua.2023.103598>
- Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit : Estimating the information employed for face classifications. *Cognitive Science*, 28(2), 209-226. [https://doi.org/10.1207/s15516709cog2802\\_4](https://doi.org/10.1207/s15516709cog2802_4)
- Martinet, A. (1982). Linguistique, paralinguistique, extralinguistique. *International Review of Applied Linguistics in Language Teaching*, 20(1-4), 83-84.
- Mason, D., Capp, S. J., Stewart, G. R., Kempton, M. J., Glaser, K., Howlin, P., & Happé, F. (2021). A Meta-analysis of Outcome Studies of Autistic Adults : Quantifying Effect Size, Quality, and Meta-regression. *Journal of Autism and Developmental Disorders*, 51(9), 3165-3179. <https://doi.org/10.1007/s10803-020-04763-2>
- Mathôt, S. (2018). Pupillometry : Psychology, Physiology, and Function. *Journal of Cognition*, 1(1). <https://doi.org/10.5334/joc.18>
- Mathôt, S., & Naber, M. (2018). There is no evidence that pupil mimicry is a social phenomenon. *Proceedings of the National Academy of Sciences*, 115(50), E11565-E11565. <https://doi.org/10.1073/pnas.1814429115>
- Mauersberger, H., & Hess, U. (2019). When smiling back helps and scowling back hurts : Individual differences in emotional mimicry are associated with self-reported interaction quality during conflict interactions. *Motivation and Emotion*, 43(3), 471-482. <https://doi.org/10.1007/s11031-018-9743-x>
- Mauersberger, H., Kastendieck, T., & Hess, U. (2022). I looked at you, you looked at me, I smiled at you, you smiled at me—The impact of eye contact on emotional mimicry. *Frontiers in Psychology*, 13. <https://www.frontiersin.org/articles/10.3389/fpsyg.2022.970954>
- McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders : A critical review. *International Journal of Language & Communication Disorders*, 38(4), 325-350. <https://doi.org/10.1080/1368282031000154204>

## Références

- McIntosh, D. N., Reichmann-Decker, A., Winkielman, P., & Wilbarger, J. L. (2006). When the social mirror breaks : Deficits in automatic, but not voluntary, mimicry of emotional facial expressions in autism. *Developmental Science*, 9(3), 295-302. <https://doi.org/10.1111/j.1467-7687.2006.00492.x>
- Mehrabian, A. (1981). *Silent messages : Implicit communication of emotions and attitudes* (2d ed). Wadsworth Pub. Co.
- Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition : A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 491-500. <https://doi.org/10.1098/rstb.2002.1261>
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of Facial and Manual Gestures by Human Neonates. *Science*, 198(4312), 75-78.
- Milton, D. E. M. (2012). On the ontological status of autism : The 'double empathy problem'. *Disability & Society*, 27(6), 883-887. <https://doi.org/10.1080/09687599.2012.710008>
- Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons : From motion to 'e'motion. *Social Cognitive and Affective Neuroscience*, 1(3), 235-241. <https://doi.org/10.1093/scan/nsl029>
- Moore, V., & Goodson, S. (2003). How Well Does Early Diagnosis of Autism Stand the Test of Time? : Follow-Up Study of Children Assessed for Autism at Age 2 and Development of an Early Diagnostic Service. *Autism*, 7(1), 47-63. <https://doi.org/10.1177/1362361303007001005>
- Morningstar, M., Venticinque, J., & Nelson, E. E. (2019). Differences in adult and adolescent listeners' ratings of valence and arousal in emotional prosody. *Cognition and Emotion*. <https://www.tandfonline.com/doi/abs/10.1080/02699931.2018.1561422>
- Moseley, R. L., & Pulvermüller, F. (2018). What can autism teach us about the role of sensorimotor systems in higher cognition? New clues from studies on language, action semantics, and abstract emotional concept processing. *Cortex*, 100, 149-190. <https://doi.org/10.1016/j.cortex.2017.11.019>
- Moseley, R. L., Shtyrov, Y., Mohr, B., Lombardo, M. V., Baron-Cohen, S., & Pulvermüller, F. (2015). Lost for emotion words : What motor and limbic brain activity reveals about autism and semantic theory. *NeuroImage*, 104, 413-422. <https://doi.org/10.1016/j.neuroimage.2014.09.046>
- Moseley, R. L., Ypma, R. J. F., Holt, R. J., Floris, D., Chura, L. R., Spencer, M. D., Baron-Cohen, S., Suckling, J., Bullmore, E., & Rubinov, M. (2015). Whole-brain functional

## Références

- hypoconnectivity as an endophenotype of autism in adolescents. *NeuroImage: Clinical*, 9, 140-152. <https://doi.org/10.1016/j.nicl.2015.07.015>
- Mössler, K., Gold, C., Aßmus, J., Schumacher, K., Calvet, C., Reimer, S., Iversen, G., & Schmid, W. (2019). The Therapeutic Relationship as Predictor of Change in Music Therapy with Young Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49(7), 2795-2809. <https://doi.org/10.1007/s10803-017-3306-y>
- Mundy, P., & Sigman, M. (1989). The Theoretical Implications of Joint-attention Deficits in Autism. *Development and Psychopathology*, 1, 173-183. <https://doi.org/10.1017/S0954579400000365>
- Murphy, G. H., Beadle-Brown, J., Wing, L., Gould, J., Shah, A., & Holmes, N. (2005). Chronicity of Challenging Behaviours in People with Severe Intellectual Disabilities and/or Autism : A Total Population Sample. *Journal of Autism and Developmental Disorders*, 35(4), 405-418. <https://doi.org/10.1007/s10803-005-5030-2>
- Näätänen, R. (2000). Mismatch negativity (MMN) : Perspectives for application. *International Journal of Psychophysiology*, 37(1), 3-10. [https://doi.org/10.1016/S0167-8760\(00\)00091-X](https://doi.org/10.1016/S0167-8760(00)00091-X)
- Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42(4), 313-329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- Nadel, J., & Potier, C. (2002). Imiter, imitez, il en restera toujours quelque chose : Le statut développemental de l'imitation dans le cas d'autisme. *Enfance*, 54(1), 76-85. <https://doi.org/10.3917/enf.541.0076>
- Neal, D., & Chartrand, T. (2011). Embodied Emotion Perception Amplifying and Dampening Facial Feedback Modulates Emotion Perception Accuracy. *Social Psychological and Personality Science*, 2, 673-678. <https://doi.org/10.1177/1948550611406138>
- Neri, P. (2010). How inherently noisy is human sensory processing? *Psychonomic Bulletin & Review*, 17(6), 802-808. <https://doi.org/10.3758/PBR.17.6.802>
- Nestor, A., & Tarr, M. J. (2008). Gender Recognition of Human Faces Using Color. *Psychological Science*, 19(12), 1242-1246. <https://doi.org/10.1111/j.1467-9280.2008.02232.x>
- Niedenthal, P. M., Korb, S., Wood, A., & Rychlowska, M. (2016). Revisiting the Simulation of Smiles model : The what, when, and why of mimicking smiles. In *Emotional mimicry in social context* (p. 44-71). Cambridge University Press. <https://doi.org/10.1017/CBO9781107587595.004>

## Références

- Niedenthal, P. M., Mermillod, M., Maringer, M., & Hess, U. (2010). The Simulation of Smiles (SIMS) model: Embodied simulation and the meaning of facial expression. *The Behavioral and Brain Sciences*, 33(6), 417-480. <https://doi.org/10.1017/S0140525X10000865>
- Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face : Blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, 2(3-4), 167-178. <https://doi.org/10.1080/17470910701391943>
- Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2009). Slow echo : Facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Developmental Science*, 12(4), 510-520. <https://doi.org/10.1111/j.1467-7687.2008.00796.x>
- Olszanowski, M., & Wróbel, M. (2022). *Why We Mimic Emotions Even When No One is Watching: Limited Visual Contact and Emotional Mimicry*. PsyArXiv. <https://doi.org/10.31234/osf.io/qdzjw>
- Olszanowski, M., Wróbel, M., & Hess, U. (2020). Mimicking and sharing emotions : A re-examination of the link between facial mimicry and emotional contagion. *Cognition and Emotion*, 34(2), 367-376. <https://doi.org/10.1080/02699931.2019.1611543>
- Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 112(4), 713-719. [https://doi.org/10.1016/s1388-2457\(00\)00527-7](https://doi.org/10.1016/s1388-2457(00)00527-7)
- Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-Based Assessment of Autism Spectrum Disorders in Children and Adolescents. *Journal of Clinical Child & Adolescent Psychology*, 34(3), 523-540. [https://doi.org/10.1207/s15374424jccp3403\\_8](https://doi.org/10.1207/s15374424jccp3403_8)
- Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L. J., Constantino, J. N., Dobkins, K., Hutman, T., Iverson, J. M., Landa, R., Rogers, S. J., Sigman, M., & Stone, W. L. (2011). Recurrence Risk for Autism Spectrum Disorders : A Baby Siblings Research Consortium Study. *Pediatrics*, 128(3), e488-e495. <https://doi.org/10.1542/peds.2010-2825>
- Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation : A computationally guided review. *Neural Networks*, 19(3), 254-271. <https://doi.org/10.1016/j.neunet.2006.02.002>

## Références

- Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, 59(1-2), 185-198. [https://doi.org/10.1016/S1071-5819\(03\)00017-X](https://doi.org/10.1016/S1071-5819(03)00017-X)
- Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and Production of Prosody by Speakers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 35(2), 205-220. <https://doi.org/10.1007/s10803-004-1999-1>
- Paul, R., Shriberg, L. D., McSweeney, J., Cicchetti, D., Klin, A., & Volkmar, F. (2005). Brief report : Relations between prosodic performance and communication and socialization ratings in high functioning speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 35(6), 861-869. <https://doi.org/10.1007/s10803-005-0031-8>
- Paulmann, S., & Kotz, S. A. (2008a). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, 105(1), 59-69. <https://doi.org/10.1016/j.bandl.2007.11.005>
- Paulmann, S., & Kotz, S. A. (2008b). Early emotional prosody perception based on different speaker voices. *NeuroReport*, 19(2), 209-213. <https://doi.org/10.1097/WNR.0b013e3282f454db>
- Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion*, 35(2), 192-201. <https://doi.org/10.1007/s11031-011-9206-0>
- Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language*, 104(3), 262-269. <https://doi.org/10.1016/j.bandl.2007.03.002>
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech : Evidence from 7-week-old infants. *Infant Behavior and Development*, 15(3), 325-345. [https://doi.org/10.1016/0163-6383\(92\)80003-D](https://doi.org/10.1016/0163-6383(92)80003-D)
- Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing Emotions in a Foreign Language. *Journal of Nonverbal Behavior*, 33(2), 107-120. <https://doi.org/10.1007/s10919-008-0065-7>
- Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological Psychology*, 111, 14-25. <https://doi.org/10.1016/j.biopsycho.2015.08.008>

## Références

- Pellicano, E., & Burr, D. (2012). When the world becomes 'too real' : A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504-510. <https://doi.org/10.1016/j.tics.2012.08.009>
- Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and Expressive Prosodic Ability in Children With High-Functioning Autism. *Journal of Speech, Language, and Hearing Research*, 50(4), 1015-1028. [https://doi.org/10.1044/1092-4388\(2007/071\)](https://doi.org/10.1044/1092-4388(2007/071))
- Pernet, C. R., McAleer, P., Latinus, M., Gorgolewski, K. J., Charest, I., Bestelmeyer, P. E. G., Watson, R. H., Fleming, D., Crabbe, F., Valdes-Sosa, M., & Belin, P. (2015). The human voice areas : Spatial organization and inter-individual variability in temporal and extra-temporal cortices. *NeuroImage*, 119, 164-174. <https://doi.org/10.1016/j.neuroimage.2015.06.050>
- Pinheiro, A. P., Barros, C., Vasconcelos, M., Obermeier, C., & Kotz, S. A. (2017). Is laughter a better vocal change detector than a growl? *Cortex*, 92, 233-248. <https://doi.org/10.1016/j.cortex.2017.03.018>
- Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music of language : An ERP investigation of the effects of musical training on emotional prosody processing. *Brain and Language*, 140, 24-34. <https://doi.org/10.1016/j.bandl.2014.10.009>
- Podesva, R. J., Callier, P., Voigt, R., & Jurafsky, D. (2015). The Connection Between Smiling and GOAT Fronting : Embodied Affect in Sociophonetic Variation. *Proceedings of the 18th International Congress of Phonetic Sciences*.
- Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition*, 92(1-2), 1-12. <https://doi.org/10.1016/j.cognition.2003.11.001>
- Ponsot, E., Arias, P., & Aucouturier, J.-J. (2018). Uncovering mental representations of smiled speech using reverse correlation. *The Journal of the Acoustical Society of America*, 143(1), EL19-EL24. <https://doi.org/10.1121/1.5020989>
- Ponsot, E., Burred, J. J., Belin, P., & Aucouturier, J.-J. (2018). Cracking the social code of speech prosody using reverse correlation. *Proceedings of the National Academy of Sciences*, 115(15), 3972-3977. <https://doi.org/10.1073/pnas.1716090115>
- Posserud, M.-B., Skretting Solberg, B., Engeland, A., Haavik, J., & Klungsøyr, K. (2021). Male to female ratios in autism spectrum disorders by age, intellectual disability and attention-deficit/hyperactivity disorder. *Acta Psychiatrica Scandinavica*, 144(6), 635-646. <https://doi.org/10.1111/acps.13368>

## Références

- Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2019). Emotional prosody in congenital amusia: Impaired and spared processes. *Neuropsychologia*, *134*, 107234. <https://doi.org/10.1016/j.neuropsychologia.2019.107234>
- Press, C., Richardson, D., & Bird, G. (2010). Intact imitation of emotional facial actions in autism spectrum conditions. *Neuropsychologia*, *48*(11), 3291-3297. <https://doi.org/10.1016/j.neuropsychologia.2010.07.012>
- Preston, S. D., & Waal, F. B. M. de. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1-20. <https://doi.org/10.1017/S0140525X02000018>
- Prochazkova, E., & Kret, M. E. (2017). Connecting minds and sharing emotions through mimicry: A neurocognitive model of emotional contagion. *Neuroscience and Biobehavioral Reviews*, *80*, 99-114. <https://doi.org/10.1016/j.neubiorev.2017.05.013>
- Pruvost-Robieux, E. (2023). *Les patients en état de conscience altérée sont-ils sensibles à la prosodie?* Sorbonne Université.
- Quené, H., Semin, G. R., & Foroni, F. (2012). Audible smiles and frowns affect speech comprehension. *Speech Communication*, *54*(7), 917-922. <https://doi.org/10.1016/j.specom.2012.03.004>
- Quintin, E.-M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2011). Emotion Perception in Music in High-Functioning Adolescents With Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *41*(9), 1240-1255. <https://doi.org/10.1007/s10803-010-1146-0>
- Reis, H. T., Wilson, I. M., Monestere, C., Bernstein, S., Clark, K., Seidl, E., Franco, M., Gioioso, E., Freeman, L., & Radoane, K. (1990). What is smiling is beautiful and good. *European Journal of Social Psychology*, *20*(3), 259-267. <https://doi.org/10.1002/ejsp.2420200307>
- Rice, K., & Redcay, E. (2016). Interaction matters: A perceived social partner alters the neural processing of human speech. *NeuroImage*, *129*, 480-488. <https://doi.org/10.1016/j.neuroimage.2015.11.041>
- Rimé, B. (2009). Emotion Elicits the Social Sharing of Emotion: Theory and Empirical Review. *Emotion Review*, *1*(1), 60-85. <https://doi.org/10.1177/1754073908097189>
- Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin*, *95*(1), 52-77. <https://doi.org/10.1037/0033-2909.95.1.52>

## Références

- Rizzolatti, G., & Craighero, L. (2004). THE MIRROR-NEURON SYSTEM. *Annual Review of Neuroscience*, 27(Volume 27, 2004), 169-192. <https://doi.org/10.1146/annurev.neuro.27.070203.144230>
- Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Research. Cognitive Brain Research*, 3(2), 131-141. [https://doi.org/10.1016/0926-6410\(95\)00038-0](https://doi.org/10.1016/0926-6410(95)00038-0)
- Rodger, S., & Umaibalan, V. (2011). The Routines and Rituals of Families of Typically Developing Children Compared with Families of Children with Autism Spectrum Disorder : An Exploratory Study. *British Journal of Occupational Therapy*, 74(1), 20-26. <https://doi.org/10.4276/030802211X12947686093567>
- Rogier, O., Roux, S., Belin, P., Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4- to 5-year-old children. *International Journal of Psychophysiology*, 75(1), 44-47. <https://doi.org/10.1016/j.ijpsycho.2009.10.013>
- Rosenblau, G., Kliemann, D., Dziobek, I., & Heekeren, H. R. (2016). Emotional prosody processing in autism spectrum disorder. *Social Cognitive and Affective Neuroscience*, 12(2), 224-239. <https://doi.org/10.1093/scan/nsw118>
- Ruiz-Martínez, F. J., Rodríguez-Martínez, E. I., Wilson, C. E., Yau, S., Saldaña, D., & Gómez, C. M. (2020). Impaired P1 Habituation and Mismatch Negativity in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 50(2), 603-616. <https://doi.org/10.1007/s10803-019-04299-0>
- Russ, J. B., Gur, R. C., & Bilker, W. B. (2008). Validation of affective and neutral sentence content for prosodic testing. *Behavior Research Methods*, 40(4), 935-939. <https://doi.org/10.3758/BRM.40.4.935>
- Rychlowska, M., Jack, R. E., Garrod, O. G. B., Schyns, P. G., Martin, J. D., & Niedenthal, P. M. (2017). Functional Smiles : Tools for Love, Sympathy, and War. *Psychological Science*, 28(9), 1259-1270. <https://doi.org/10.1177/0956797617706082>
- Safran, A. B., & Roth, A. (1982). [Photomotor pupillary reflex in the functional examination of the optic nerve. Recent data]. *Bulletin De La Societe Belge D'ophtalmologie*, 199-200, 217-233.
- Sams, M., Aulanko, R., Hämäläinen, M., Hari, R., Lounasmaa, O. V., Lu, S. T., & Simola, J. (1991). Seeing speech : Visual information from lip movements modifies activity in the human auditory cortex. *Neuroscience Letters*, 127(1), 141-145. [https://doi.org/10.1016/0304-3940\(91\)90914-f](https://doi.org/10.1016/0304-3940(91)90914-f)

## Références

- Samson, F., Hyde, K. L., Bertone, A., Soulières, I., Mendrek, A., Ahad, P., Mottron, L., & Zeffiro, T. A. (2011). Atypical processing of auditory temporal complexity in autistics. *Neuropsychologia*, *49*(3), 546-555. <https://doi.org/10.1016/j.neuropsychologia.2010.12.033>
- Samson, F., Mottron, L., Jemel, B., Belin, P., & Ciocca, V. (2006). Can Spectro-Temporal Complexity Explain the Autistic Pattern of Performance on Auditory Tasks? *Journal of Autism and Developmental Disorders*, *36*(1), 65-76. <https://doi.org/10.1007/s10803-005-0043-4>
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition : Brain regions involved in processing anger prosody. *NeuroImage*, *28*(4), 848-858. <https://doi.org/10.1016/j.neuroimage.2005.06.023>
- Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(6), 2408-2412. <https://doi.org/10.1073/pnas.0908239106>
- Sauter, D. A., Panattoni, C., & Happé, F. (2013). Children's recognition of emotions from vocal cues. *British Journal of Developmental Psychology*, *31*(1), 97-113. <https://doi.org/10.1111/j.2044-835X.2012.02081.x>
- Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379-399. <https://doi.org/10.1037/h0046234>
- Schall, S., Kiebel, S. J., Maess, B., & von Kriegstein, K. (2015). Voice identity recognition : Functional division of the right STS and its behavioral relevance. *Journal of Cognitive Neuroscience*, *27*(2), 280-291. [https://doi.org/10.1162/jocn\\_a\\_00707](https://doi.org/10.1162/jocn_a_00707)
- Schelinski, S., Borowiak, K., & von Kriegstein, K. (2016). Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition. *Social Cognitive and Affective Neuroscience*, *11*(11), 1812-1822. <https://doi.org/10.1093/scan/nsw089>
- Scherer, K. (2003). Vocal communication of emotion : A review of research paradigms. *Speech Communication*, *40*(1-2), 227-256. [https://doi.org/10.1016/S0167-6393\(02\)00084-5](https://doi.org/10.1016/S0167-6393(02)00084-5)
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere : Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24-30. <https://doi.org/10.1016/j.tics.2005.11.009>

## Références

- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Research. Cognitive Brain Research*, 14(2), 228-233. [https://doi.org/10.1016/s0926-6410\(02\)00108-8](https://doi.org/10.1016/s0926-6410(02)00108-8)
- Schuck, R. K., Flores, R. E., & Fung, L. K. (2019). Brief Report : Sex/Gender Differences in Symptomology and Camouflaging in Adults with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 49(6), 2597-2604. <https://doi.org/10.1007/s10803-019-03998-y>
- Senju, A., Maeda, M., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2007). Absence of contagious yawning in children with autism spectrum disorder. *Biology Letters*, 3(6), 706-708. <https://doi.org/10.1098/rsbl.2007.0337>
- Serdarevic, F., Ghassabian, A., van Batenburg-Eddes, T., White, T., Blanken, L. M. E., Jaddoe, V. W. V., Verhulst, F. C., & Tiemeier, H. (2017). Infant muscle tone and childhood autistic traits : A longitudinal study in the general population. *Autism Research*, 10(5), 757-768. <https://doi.org/10.1002/aur.1739>
- Seropian, L., Ferschneider, M., Cholvy, F., Micheyl, C., Bidet-Caulet, A., & Moulin, A. (2022). Comparing methods of analysis in pupillometry : Application to the assessment of listening effort in hearing-impaired patients. *Heliyon*, 8(6), e09631. <https://doi.org/10.1016/j.heliyon.2022.e09631>
- Shah, A., & Frith, U. (1993). Why Do Autistic Individuals Show Superior Performance on the Block Design Task? *Journal of Child Psychology and Psychiatry*, 34(8), 1351-1364. <https://doi.org/10.1111/j.1469-7610.1993.tb02095.x>
- Sheppard, E., Pillai, D., Wong, G. T.-L., Ropar, D., & Mitchell, P. (2016). How Easy is it to Read the Minds of People with Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, 46(4), 1247-1254. <https://doi.org/10.1007/s10803-015-2662-8>
- Shore, D. M., & Heerey, E. A. (2011). The value of genuine and polite smiles. *Emotion*, 11, 169-174. <https://doi.org/10.1037/a0022601>
- Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A. M., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research: JSLHR*, 44(5), 1097-1115. [https://doi.org/10.1044/1092-4388\(2001/087\)](https://doi.org/10.1044/1092-4388(2001/087))
- Simpson, E. A., Murray, L., Paukner, A., & Ferrari, P. F. (2014). The mirror neuron system as revealed through neonatal imitation: Presence from birth, predictive power and evidence of plasticity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 369(1644), 20130289. <https://doi.org/10.1098/rstb.2013.0289>

## Références

- Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), 15220-15225. <https://doi.org/10.1073/pnas.1416797111>
- Soussignan, R. (2002). Duchenne smile, emotional experience, and autonomic reactivity : A test of the facial feedback hypothesis. *Emotion (Washington, D.C.)*, *2*(1), 52-74. <https://doi.org/10.1037/1528-3542.2.1.52>
- Ssucharewa, Dr., G. E. (1926). Die schizoiden Psychopathien im Kindesalter. (Part 1 of 2). *Monatsschrift für Psychiatrie und Neurologie*, *60*(3-4), 235-247. <https://doi.org/10.1159/000190478>
- Stanutz, S., Wapnick, J., & Burack, J. A. (2014). Pitch discrimination and melodic memory in children with autism spectrum disorders. *Autism*, *18*(2), 137-147. <https://doi.org/10.1177/1362361312462905>
- Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile : A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*(5), 768-777. <https://doi.org/10.1037/0022-3514.54.5.768>
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and Communication in Autism. In *Handbook of autism and pervasive developmental disorders : Diagnosis, development, neurobiology, and behavior, Vol. 1, 3rd ed* (p. 335-364). John Wiley & Sons, Inc.
- Tartter, V. C. (1980). Happy talk : Perceptual and acoustic effects of smiling on speech. *Perception & Psychophysics*, *27*(1), 24-27. <https://doi.org/10.3758/BF03199901>
- Tartter, V. C., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *The Journal of the Acoustical Society of America*, *96*(4), 2101-2107. <https://doi.org/10.1121/1.410151>
- Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7*(1), 53-71. [https://doi.org/10.1207/s15327078in0701\\_5](https://doi.org/10.1207/s15327078in0701_5)
- Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijdsdijk, F. (2016). Heritability of autism spectrum disorders : A meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *57*(5), 585-595. <https://doi.org/10.1111/jcpp.12499>
- Todorov, A., Dotsch, R., Wigboldus, D. H. J., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social and Personality Psychology Compass*, *5*(10), 775-791. <https://doi.org/10.1111/j.1751-9004.2011.00389.x>
- Tuller, L., Konrad, I., Ridet, J., & Prévost, P. (2021). *LITMUS-SR-FR-Adults*. Université de Tours.

## Références

- Turner, M. (1999). Annotation : Repetitive behaviour in autism: a review of psychological research. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 40(6), 839-849.
- Uljarevic, M., & Hamilton, A. (2013). Recognition of Emotions in Autism : A Formal Meta-Analysis. *Journal of Autism and Developmental Disorders*, 43(7), 1517-1526. <https://doi.org/10.1007/s10803-012-1695-5>
- Valente, D., Theurel, A., & Gentaz, E. (2018). The role of visual experience in the production of emotional facial expressions by blind people : A review. *Psychonomic Bulletin & Review*, 25(2), 483-497. <https://doi.org/10.3758/s13423-017-1338-0>
- Van de Cruys, S., de-Wit, L., Evers, K., Boets, B., & Wagemans, J. (2013). Weak Priors versus Overfitting of Predictions in Autism : Reply to Pellicano and Burr ( *TICS* , 2012). *I-Perception*, 4(2), 95-97. <https://doi.org/10.1068/i0580ic>
- Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds : Predictive coding in autism. *Psychological Review*, 121(4), 649. <https://doi.org/10.1037/a0037665>
- Van Hoorn, J., Van Dijk, E., Crone, E. A., Stockmann, L., & Rieffe, C. (2017). Peers Influence Prosocial Behavior in Adolescent Males with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 47(7), 2225-2237. <https://doi.org/10.1007/s10803-017-3143-z>
- van Boxtel, A., & van der Graaff, J. (2024). Standardization of facial electromyographic responses. *Biological Psychology*, 185, 108737. <https://doi.org/10.1016/j.biopsycho.2023.108737>
- Verona, E., Patrick, C. J., Curtin, J. J., Bradley, M. M., & Lang, P. J. (2004). Psychopathy and Physiological Response to Emotionally Evocative Sounds. *Journal of Abnormal Psychology*, 113(1), 99-108. <https://doi.org/10.1037/0021-843X.113.1.99>
- Vilaverde, R. F., Horchak, O. V., Pinheiro, A. P., Scott, S. K., Korb, S., & Lima, C. F. (2024). Inhibiting orofacial mimicry affects authenticity perception in vocal emotions. *Emotion*, No          Pagination          Specified-No          Pagination          Specified. <https://doi.org/10.1037/emo0001361>
- Vilidaite, G., Yu, M., & Baker, D. H. (2017). Internal noise estimates correlate with autistic traits. *Autism Research: Official Journal of the International Society for Autism Research*, 10(8), 1384-1391. <https://doi.org/10.1002/aur.1781>
- Waal, F. B., & Ferrari, P. F. (2010). Towards a bottom-up perspective on animal and human cognition. *Trends in Cognitive Sciences*, 14(5), 201-207.

## Références

- Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., Albohn, D. N., Allard, E. S., Benning, S. D., Blouin-Hudon, E.-M., Bulnes, L. C., Caldwell, T. L., Calin-Jageman, R. J., Capaldi, C. A., Carfagno, N. S., Chasten, K. T., Cleeremans, A., Connell, L., DeCicco, J. M., ... Zwaan, R. A. (2016). Registered Replication Report : Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, 11(6), 917-928. <https://doi.org/10.1177/1745691616674458>
- Wagner, J. B., Luyster, R. J., Tager-Flusberg, H., & Nelson, C. A. (2016). Greater Pupil Size in Response to Emotional Faces as an Early Marker of Social-Communicative Difficulties in Infants at High Risk for Autism. *Infancy*, 21(5), 560-581. <https://doi.org/10.1111/infa.12128>
- Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2007). Reading affect in the face and voice : Neural correlates of interpreting communicative intent in children and adolescents with autism spectrum disorders. *Archives of General Psychiatry*, 64(6), 698-708. <https://doi.org/10.1001/archpsyc.64.6.698>
- Wang, J.-E., & Tsao, F.-M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, 37, 162-170. <https://doi.org/10.1016/j.ridd.2014.11.013>
- Wang, L., Ong, J. H., Ponsot, E., Hou, Q., Jiang, C., & Liu, F. (2022). Mental representations of speech and musical pitch contours reveal a diversity of profiles in autism spectrum disorder. *Autism*, 13623613221111207. <https://doi.org/10.1177/13623613221111207>
- Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody—Influence of acoustic parameters and arousal. *NeuroImage*, 39(2), 885-893. <https://doi.org/10.1016/j.neuroimage.2007.09.028>
- Wilhelm, H. (2008). The pupil. *Current Opinion in Neurology*, 21(1), 36-42. <https://doi.org/10.1097/WCO.0b013e3282f39173>
- Williams, C. E., & Stevens, K. N. (1972). Emotions and Speech : Some Acoustical Correlates. *The Journal of the Acoustical Society of America*, 52(4B), 1238-1250. <https://doi.org/10.1121/1.1913238>
- Ylinen, S., & Huotilainen, M. (2007). Is there a direct neural correlate for memory-trace formation in audition? *NeuroReport*, 18(12), 1281-1284. <https://doi.org/10.1097/WNR.0b013e32826fb38a>

## Références

- Yoshimatsu, Y., Umino, A., & Dammeyer. (2016). Characteristics of the Understanding and Expression of Emotional Prosody among Children with Autism Spectrum Disorder. *Autism-Open Access*, 6(4). <https://doi.org/10.4172/2165-7890.1000185>
- Young, A. I., Ratner, K. G., & Fazio, R. H. (2014). Political attitudes bias the mental representation of a presidential candidate's face. *Psychological Science*, 25(2), 503-510. <https://doi.org/10.1177/0956797613510717>
- Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology*, 102(4), 959-974. <https://doi.org/10.1111/j.2044-8295.2011.02045.x>
- Zekveld, A. A., Koelewijn, T., & Kramer, S. E. (2018). The Pupil Dilation Response to Auditory Stimuli: Current State of Knowledge. *Trends in Hearing*, 22, 233121651877717. <https://doi.org/10.1177/2331216518777174>



## **Annexes**

---



## Annexe 1 : Récapitulatif des participants recrutés dans les différentes études

Ici est présenté le récapitulatif des participants vus dans le cadre de cette thèse qui s'intègre plus largement dans le projet ANR projet SEPIA (*Sensory and Emotional Processing In Autism spectrum disorders*). Ce travail de thèse a permis la mise en place de tous les paradigmes associés au projet SEPIA.

		EEG		RevCor	Résonance motrice		Total	
		Étude préliminaire	Version définitive	Version définitive	Étude préliminaire	Version définitive		
Adultes	NT	21	23	35	23	28	130	205
	ASD	/	21	34	/	20	75	
<b>Total</b>		65		69	71		205	

**NB : Si un participant a réalisé plusieurs études il est présent plusieurs fois dans ce tableau. Ici est indiqué le nombre de jeux de données acquis et traités dans le cadre du projet SEPIA.**

Toutes les acquisitions, faites dans le cadre de ce projet, ont été faites avec des paradigmes mis en place dans ce travail de thèse.

Des données chez les enfants neurotypiques et autistes ont été acquises par Zoé Ranty, pendant son stage de M2, que j'ai co-supervisé, s'intitulant « Traitement de la prosodie émotionnelle chez l'enfant : approche comportementale et neurophysiologique », ainsi que pendant sa première année de thèse. Les données acquises, tous paradigmes confondus, représentent au total 99 jeux de données comprenant ceux de 81 enfants neurotypiques et de 18 autistes.

Les données acquises chez l'enfant ne sont pas reprises dans ce travail de thèse.

Le nombre d'inclusions dans le cadre du projet SEPIA ont été mis à jour le 6 août 2024, depuis d'autres acquisitions ont eu lieu.



## Annexe 2 : Étude préliminaire de l'adaptation neurale chez l'adulte neurotypique

### 1. Contexte et Objectifs

Dans l'étude n°1 de cette thèse, il a été mis en évidence qu'une adaptation neurale à des sons vocaux et non-vocaux était possible chez l'adulte neurotypique toutefois les dynamiques de mise en place de ce phénomène n'étaient pas les mêmes. Cette étude a montré que l'adaptation était observable au niveau électrophysiologique dans un paradigme de Roving, néanmoins le protocole (2 séquences de 30 minutes) était trop long pour une population clinique (TSA). Dans l'objectif de tester les effets de la prosodie émotionnelle, dans la mise en place de ce phénomène, il fallait tester une séquence avec des sons de natures variées (émotionnels, neutres, vocaux...) dans une seule séquence rapide. Cette étude préliminaire visait à sélectionner les sons et la durée de séquence les plus appropriés pour examiner l'adaptation neuronale à la prosodie émotionnelle dans le cadre de recherches futures sur l'autisme.

### 2. Méthodologie

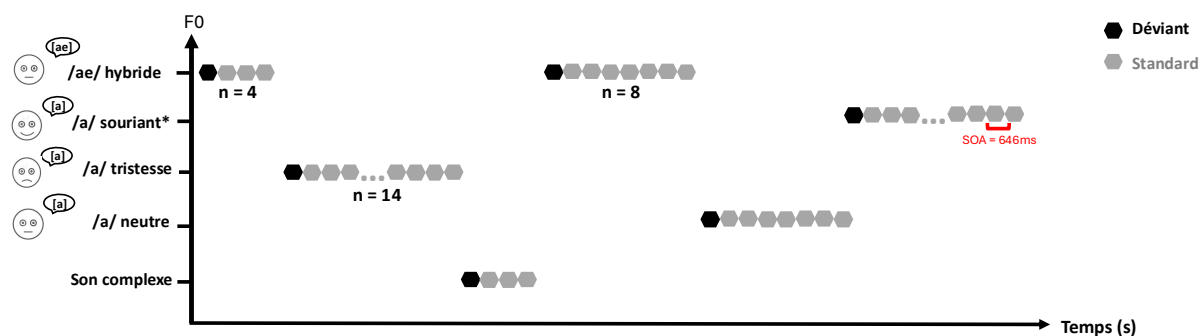
Dans cette étude préliminaire 5 sons de nature différente (tableau S1) étaient présentés dans une seule séquence de Roving d'une trentaine de minutes avec des trains de 4, 8 et 14 répétitions (figure S1). Afin d'avoir des résultats plus robustes seulement les 20 premières minutes de séquence ont été prises en compte. En effet, le processus d'adaptation peut s'atténuer si la séquence de stimulation est trop longue. Les résultats présentés sont ceux après 20 minutes. Le nombre de trains présentés dans le protocole complet et dans les données prises en compte est repris dans le tableau S2.

Nom	Nature	Valence	Spécificité
<b>Complexe</b>	Complexe	Neutre	Son complexe synthétisé à partir du son neutre en reprenant la fondamentale et les 4 premiers formants
<b>Neutre</b>	Vocal	Neutre	/a/ locutrice
<b>Triste</b>	Vocal	Triste	/a/ locutrice
<b>Souriant</b>	Vocal	Souriant	/a/ locutrice – son rendu souriant à partir représentation mentale de la RevCor
<b>Hybride</b>	Vocal	Neutre	Son hybride entre un /a/ et un /e/ – locutrice

*Tableau S1: Récapitulatif des caractéristiques des sons utilisés dans ce paradigme*

	Trains de 4 répétitions	Trains de 8 répétitions	Trains de 14 répétitions	Durée de la séquence
<b>Tâche complète</b>	100 (20 par nature de son)	100 (20 par nature de son)	120 (24 par nature de son)	31 min
<b>Tâche raccourcie</b>	60 (12 par nature de son)	60 (12 par nature de son)	90 (18 par nature de son)	21 min

**Tableau S2: Récapitulatif du nombre de trains dans la séquence de stimulation**



\* /a/ souriant à partir de la tâche de Reverse Correlation (Ponsot et al., 2018)

**Figure S1 : Séquence de stimulation de Roving avec 5 sons de natures différentes pour tester les sons et la longueur de tâche les plus adaptés pour mesurer une adaptation neurale à la prosodie émotionnelle dans l'autisme. SOA : Stimulus Onset Asynchrony**

Dans cette étude préliminaire, 20 adultes neurotypiques ont été inclus (tableau S3). Les critères d'inclusion pour cette étude étaient : i) une absence de maladie infectieuse ou métabolique sévère, ii) absence d'épilepsie ou quelconque antécédent de maladies du système nerveux central, iii) d'antécédents psychiatriques et iv) une absence de délai dans l'acquisition du langage et de la marche. Cette étude était encadrée par le même cadre éthique que les études principales (PROSCEA 2017/23 ; ID RCB : 2017-A00756-47).

<b>Genre</b>	13 ♀ 7 ♂
<b>Age</b>	22,4 ± 6
<b>AQ</b>	17,0 ± 6
<b>EQ</b>	43,3 ± 7
<b>QIv</b>	126,0 ± 17
<b>QInv</b>	93 ± 9

**Tableau S3 : Récapitulatif des caractéristiques des participants à cette étude préliminaire en EEG**

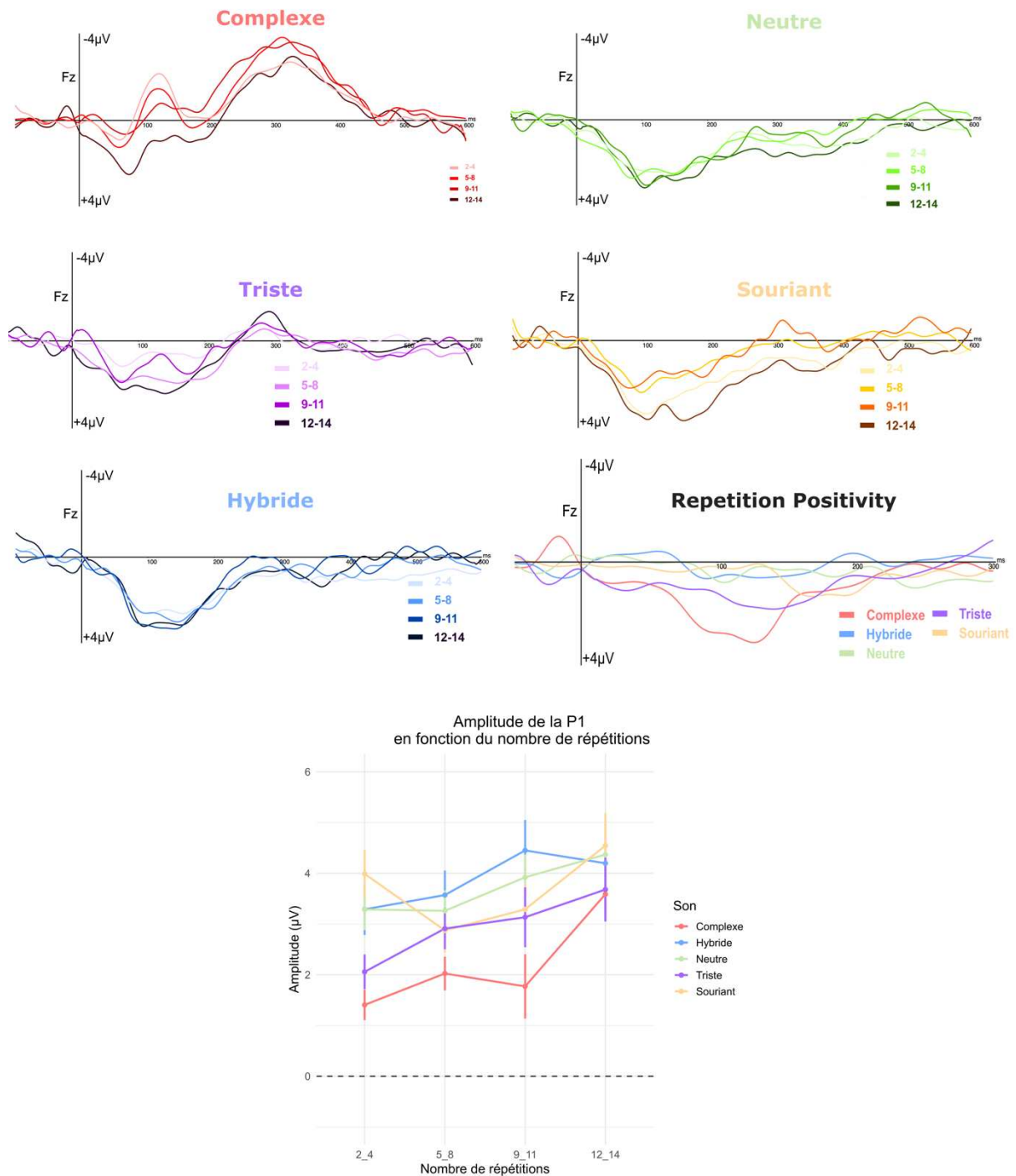
La même méthode et le même matériel d'acquisition que dans les études n°1 et 2 ont été utilisés, seule la séquence de stimulation était différente. Le pré-traitement et le traitement des

## Annexe 2 : Étude préliminaire de l'adaptation neurale chez l'adulte neurotypique

données ont été faits de la même manière que dans l'étude n°2 (voir figure 10). Les PEAs en réponse à 2-4, 5-8, 9-11 et 12-14 répétitions ont été moyennés ensemble. Les amplitudes de réponse (P1) ont été mesurées sur l'électrode Fz pour étudier la mise en place d'une adaptation neurale. Les RP (*Repetition Positivity*) ont été calculées comme la différence entre beaucoup de répétitions (12-14) et peu de répétitions (2-4) pour chaque nature de sons.

### **3. Résultats**

Annexe 2 : Étude préliminaire de l'adaptation neurale chez l'adulte neurotypique



**Figure S2: PEAs en réponse aux différentes natures de son et en fonction du nombre de répétitions**  
 Pour chacun des sons les couleurs vont du plus clair au plus foncé représentant le nombre croissant de répétitions. Les Repetition Positivity ont été calculé (12-14 – 2-4) pour chaque nature de son. Les amplitudes de la P1 ont été mesuré comme l'amplitude maximale dans la fenêtre 60-120ms pour les sons vocaux (neutre, hybride, souriant et triste) et dans la fenêtre 50-110ms pour le son complexe (moyenne ± sem).

Annexe 2 : Étude préliminaire de l'adaptation neurale chez l'adulte neurotypique

Sons	ANOVA Amplitude P1 en fonction du nombre de répétitions	Comparaisons planifiées					
		contrast	estimate	SE	df	t.ratio	p.value
Complexe	F(3,57) = 4.80 ; p < .05 ; $\eta^2 = .15$ *	2_4/5_8	0.619	0.62	57	0.998	0.3224
		2_4/9_11	0.365	0.62	57	0.588	0.5587
		2_4/12_14	2.182	0.62	57	3.519	<b>0.0009 ***</b>
Neutre	F(3,57) = 1.83 ; p > .05 ; $\eta^2 = .05$	2_4/5_8	-0.0263	0.56	57	-0.047	0.9627
		2_4/9_11	0.6341	0.56	57	1.133	0.2618
		2_4/12_14	1.0825	0.56	57	1.935	<b>0.0580 •</b>
Triste	F(3,57) = 2.29 ; p < .1 ; $\eta^2 = .07$ •	2_4/5_8	0.853	0.63	57	1.353	0.1813
		2_4/9_11	1.076	0.63	57	1.707	<b>0.0932 •</b>
		2_4/12_14	1.622	0.63	57	2.574	<b>0.0127 *</b>
Souriant	F(3,57) = 3.94 ; p < .05 ; $\eta^2 = .08$ *	2_4/5_8	-1.115	0.528	57	-2.112	<b>0.0391 *</b>
		2_4/9_11	-0.699	0.528	57	-1.324	0.1909
		2_4/12_14	0.556	0.528	57	1.054	0.2963
Hybride	F(3,57) = 1.76 ; p > .05 ; $\eta^2 = .04$	2_4/5_8	0.282	0.574	57	0.492	0.6248
		2_4/9_11	1.162	0.574	57	2.022	<b>0.0479 *</b>
		2_4/12_14	0.908	0.574	57	1.581	0.1193

Tableau S4 : Résultats des ANOVAs et des comparaisons planifiées de l'amplitude de la P1 en fonction du nombre de répétitions pour chaque nature de son. • p < .1 ; \* p < .05 ; \*\* p < .01 ; \*\*\* p < .001

#### 4. Discussion

La multiplicité des sons utilisés et des comparaisons effectuées dans cette étude préliminaire ont permis de choisir les stimuli les plus adaptés pour l'étude sur le traitement et l'adaptation neurale dans l'autisme à la prosodie émotionnelle et à la voix (étude n°2). La réponse au sourire vocal dans cette étude préliminaire apparaît comme atypique avec une absence de RP et une adaptation très rapide (5-8 répétitions) mais qui ne se maintient pas dans le temps. Au contraire, l'adaptation à une voix triste apparaissait comme stable après 12-14 répétitions malgré un protocole varié dans sa séquence. Enfin, le problème dans ce paradigme était que celui-ci n'était pas équilibré en fonction de la nature des sons. En effet, un seul son complexe contre 4 sons vocaux ne permettait pas de faire la comparaison entre l'adaptation à la voix et à des sons complexes comme dans l'étude n°1, une étape à caractériser dans l'autisme. Ce paradigme nous a permis de tester une de nos hypothèses qui était une absence d'adaptation neurale à un son hybride qui a été mise en évidence dans ce groupe d'adultes neurotypiques. En effet, aucune modulation majeure de l'amplitude de la réponse avec les répétitions n'est observée pour cette nature de son. Quand un son n'est pas familier l'adaptation neurale n'apparaît pas possible dans un protocole de Roving.

Cette étude préliminaire était nécessaire pour choisir les sons et la séquence qui seraient ensuite utilisés pour étudier l'adaptation neurale à la voix et à la prosodie émotionnelle dans l'autisme (voir méthode étude n°2). La séquence choisie était plus courte et mieux équilibrée que celle de l'étude préliminaire.

## Annexe 3 : Étude préliminaire de la résonance motrice chez l'adulte neurotypique

### 1. Contexte et Objectifs

Dans les études n°4 et 5, un paradigme a été utilisé pour étudier la résonance motrice et autonome en réponse à des phrases émotionnelles. Ces études ont pu être mises en place grâce à cette étude préliminaire qui a permis de tester la faisabilité du paradigme et de la tâche.

### 2. Méthodologie

Une première étude de la résonance motrice et autonome a été réalisée chez un groupe d'adultes neurotypiques dans un objectif de faisabilité du protocole. La séquence de stimulation utilisée était exactement la même que celle des études n°4 et 5, mais dans cette étude préliminaire le jugement du degré de sourire des phrases se faisait avec une échelle à 3 niveaux. Les jugements étaient répertoriés comme *Smile*, *Neutral* et *Unsmile* (Figure S3) et les filtres utilisés comme *Smiling*, *Neutral* et *Unsmiling*.

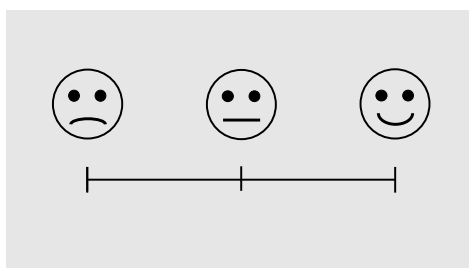


Figure S3: échelle de jugement à 3 niveaux utilisée dans l'étude préliminaire

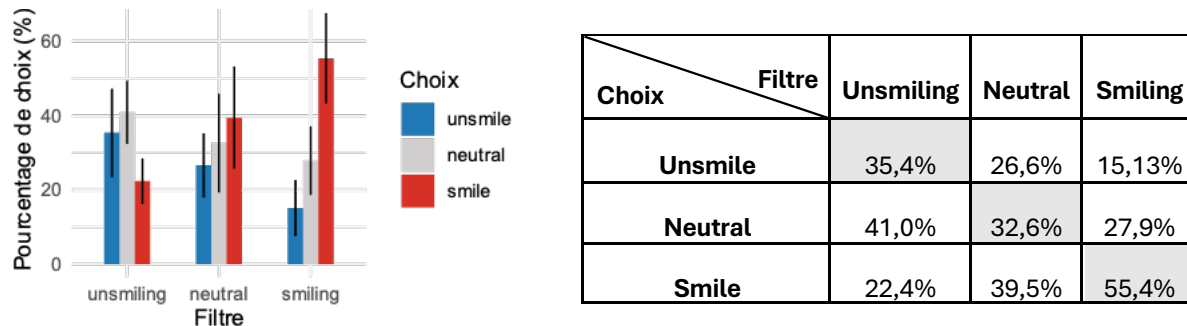
20 adultes neurotypiques ont été inclus dans cette étude préliminaire. i) une absence de maladie infectieuse ou métabolique sévère, ii) absence d'épilepsie ou quelconque antécédent de maladies du système nerveux central, iii) d'antécédents psychiatriques et iv) une absence de délai dans l'acquisition du langage et de la marche. Cette étude était encadrée par le même cadre éthique que les études principales (PROSCEA 2017/23 ; ID RCB : 2017-A00756-47).

Genre	14 ♀ 6 ♂
Age	26,9 ± 2
AQ	12,5 ± 4
EQ	43,9 ± 9

Tableau S5 : Récapitulatif des caractéristiques des participants à cette étude préliminaire en EMG et en pupillométrie

### 3. Résultats

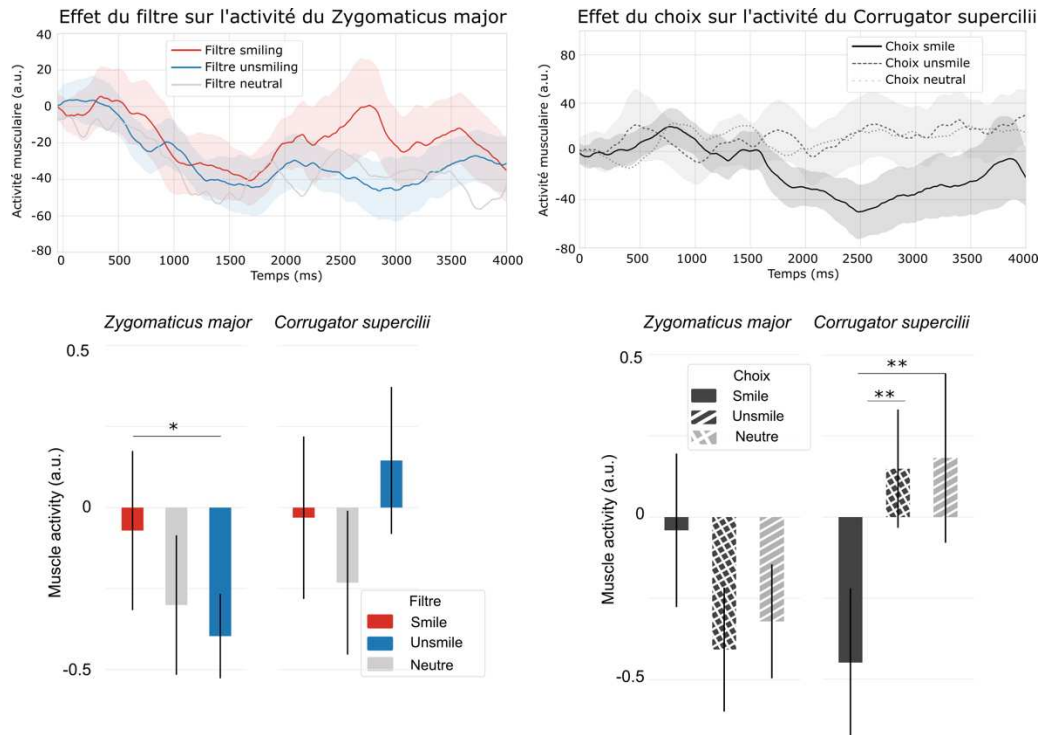
Dans cette version du protocole l'utilisation d'une échelle à 3 niveaux a rendu difficile le choix et la discrimination entre les filtres. Le filtre *Smiling* était bien reconnu tandis que le filtre *Unsmiling* et *Netral* étaient confondus (Figure S4).



**Figure S4 : Matrice de confusion des choix et histogramme des pourcentages de réponse**

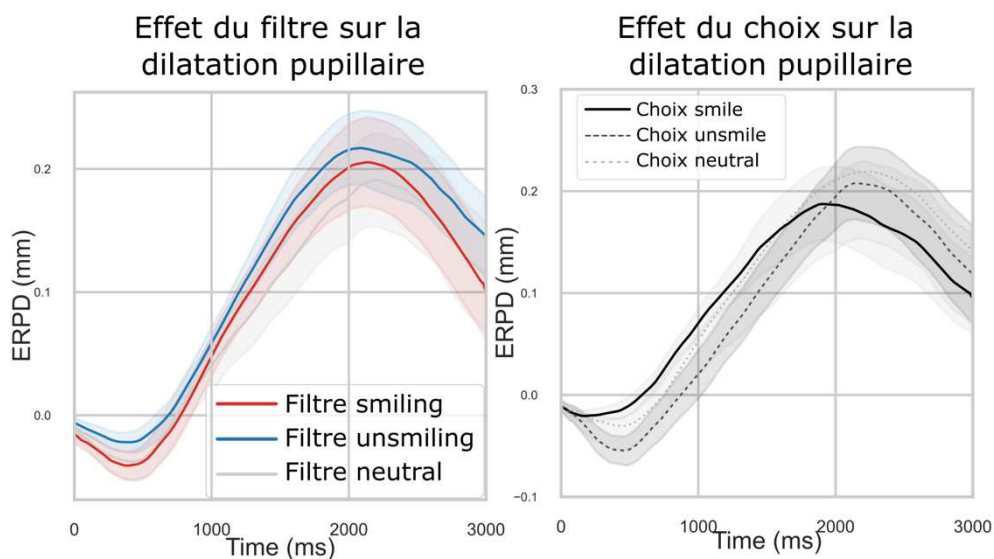
Les amplitudes des réponses musculaires ont été mesurées dans la fenêtre 2300-2800 ms. Les résultats sur les contractions musculaires sont similaires entre cette étude préliminaire et les études n°4 et 5 chez l'adulte neurotypique sur une fenêtre de mesure différente. En effet, un effet du filtre significatif est retrouvé sur l'activité du *Zygomaticus major* ( $F(2,36) = 3.6, p < .05$ ), avec la contraction différente pour le filtre *Smiling* en comparaison aux deux autres. Alors qu'un effet significatif du choix est observé pour le *Corrugator supercilii* ( $F(2,36) = 4.9, p < .05$ ), avec une contraction plus importante quand le choix n'est pas *Smile*.

### Annexe 3 : Étude préliminaire de la résonance motrice chez l'adulte neurotypique



**Figure S5: Réponse musculaire en fonction du filtre et du choix sur le Zygomaticus major et le Corrugator supercilii**

Au niveau de la réponse autonome, aucun effet sur l'amplitude de la dilatation pupillaire n'a pu être mis en évidence tant au niveau de l'effet du filtre que du choix. Une dilatation pupillaire a tout de même été observée. Cependant, un effet sur la latence du maximum a été observé. Le choix a un effet significatif sur la latence du maximum de dilatation ( $F(3,34) = 5.9, p < .01$ ). Cet effet est la conséquence d'une dilatation plus rapide pour les phrases jugées comme souriantes par rapport aux neutres et aux non-souriantes. Une interaction significative entre le choix et le filtre a aussi été mise en évidence sur la latence de la dilatation maximum ( $F(4,68) = 2.9, p < .05$ ). Cette interaction explique l'effet choix mis en avant. En effet, elle est la conséquence d'une dilatation plus rapide pour les phrases neutres jugées comme souriantes.



**Figure S6: Réponse pupillaire au filtre et au choix**

#### 4. Discussion

Cette étude préliminaire a permis de mettre en évidence qu'un protocole purement auditif permettait d'observer une résonance motrice implicite (au filtre) et explicite (au choix) chez des adultes neurotypiques, ainsi qu'un effet sur le SNA avec une dilatation de la pupille en fonction du jugement émis. Les conditions expérimentales permettaient donc de mesurer ces différents phénomènes. Cependant, l'utilisation de cette échelle à 3 niveaux a compliqué la comparaison des effets du choix sur la dilatation pupillaire et la contraction musculaire, car en cas de doutes, les participants optaient pour le jugement neutre. Cela a réduit le nombre d'essais incongruents entre le filtre et le choix.

Pour tester des conditions plus optimales de résonance motrice et autonome chez les adultes neurotypiques et autistes, nous avons opté pour une échelle à 4 niveaux dans les études n°4 et 5 afin de conserver tous les essais.

## Annexe 4 : Données supplémentaires de l'étude n°3

## Supplementary materials

## I. Creation of stimuli

To create non-vocal complex sounds, the values of the fundamental (F0) and the first four formants of each neutral and sad vocal sounds were measured with a speech analysis software<sup>1</sup>. Then non-vocal complex sounds were synthesized using sinewaves of the corresponding frequency values of vocal sounds (Adobe Audition® Software). Then, amplitude of each harmonic was adjusted to match the original vocal sound in frequency and in energy. To achieve the necessary attenuations (spectral slope), each time the frequency doubled we applied a decrease of 12 dB to the amplitude of the harmonic<sup>2</sup>. Finally, to have the most similar sounds as possible, an energy normalization has been performed. Also, fading of 10% of the duration of the sound was applied at the beginning and at the end of each sound. Acoustic sounds characteristics are presented in figure S1.

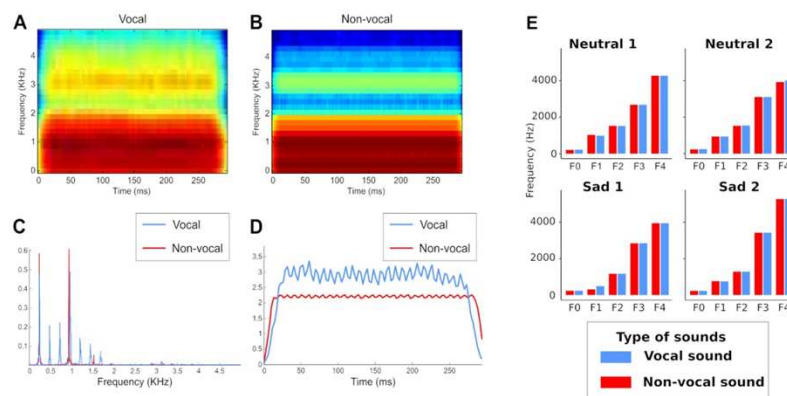


Figure S1: Physical features of a vocal sound and of its non-vocal counterpart. Spectrogram of the vocal (A) and the non-vocal (B) sound. (C) energy of every frequency for the vocal (blue) and the non-vocal (red) sounds. (D) global energy of the vocal (blue) and non-vocal (red) sound as a function of time. (E) Comparison of frequencies of vocal (blue) and non-vocal (red) for the fundamental (F0) and the first four formants (F1, F2, F3 and F4). No statistical difference was revealed between vocal and non-vocal frequencies.

## II. Neural adaptation dynamics to neutral and emotional sounds

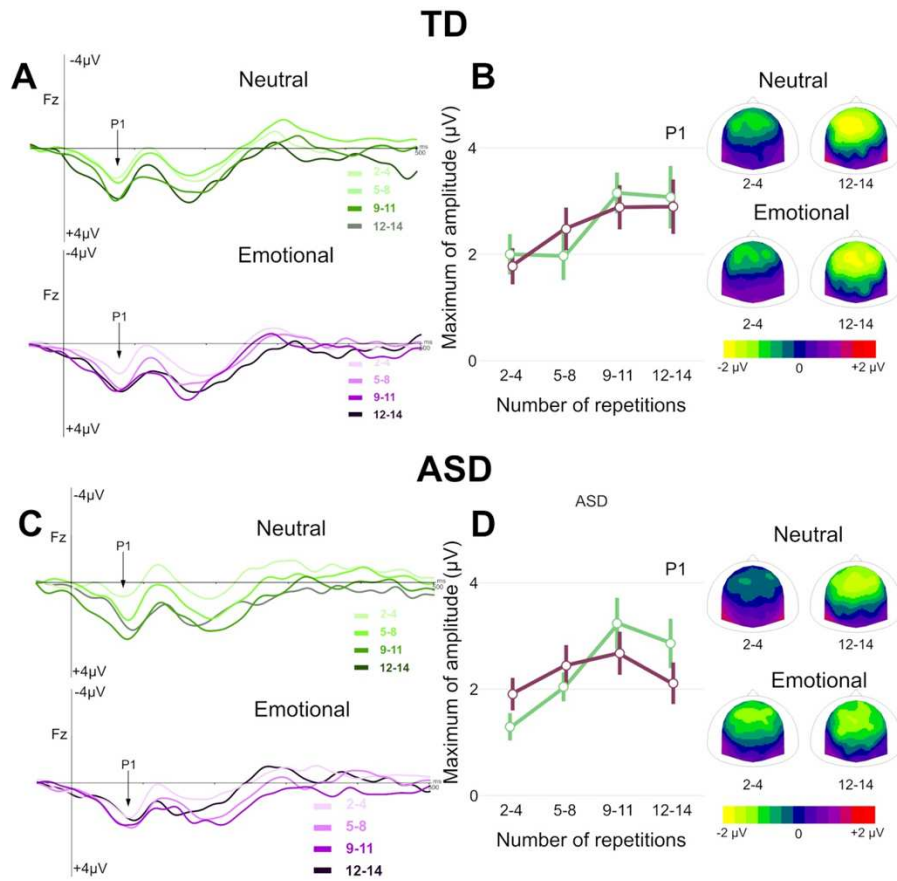


Figure S2: Neural adaptation to vocal sounds according to emotional prosody content (A) Standard ERPs at Fz electrode after 2-4, 5-8, 9-11 and 12-14 repetitions in response to vocal neutral (green – top) and emotional (purple – below) sounds in non-autistic adults (TD). (B) P1 amplitude for neutral and emotional sounds in the TD group plotted for 2-4, 5-8, 9-11 and 12-14 repetitions at Fz (left part) and scalp potential distribution for P1 after 2-4 and 12-14 repetitions for neutral (top) and emotional (below) condition. (C) Standard ERPs at Fz electrode after 2-4, 5-8, 9-11 and 12-14 repetitions in response to neutral (green – top) and emotional (purple – below) sounds in autistic adults (ASD). (D) P1 amplitude for neutral and emotional sounds in the ASD group plotted for 2-4, 5-8, 9-11 and 12-14 repetitions at Fz (left part) and scalp potential distribution for P1 after 2-4 and 12-14 repetitions for neutral (top) and emotional (below) condition.

### III. Repetition positivity to neutral and emotional sounds

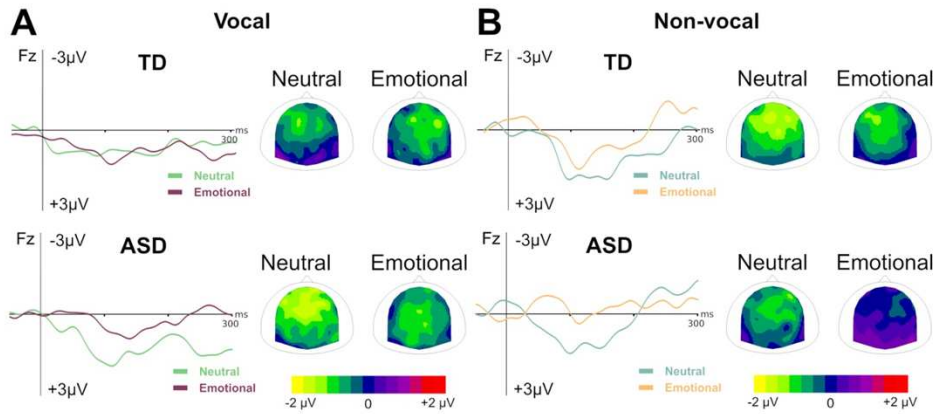


Figure S3: (A) Repetition Positivity (RP) resulting from the difference wave between 12-14 and 2-4 repetitions for vocal neutral (green) and emotional (purple) conditions in TD (top) and ASD (below) with RP scalp potential distribution between 90 and 200ms for both conditions. (B) Repetition Positivity (RP) resulting from the difference wave between 12-14 and 2-4 repetitions for non-vocal neutral (aqua) and emotional (orange) conditions in TD (top) and ASD (below) with RP scalp potential distribution between 90 and 200ms for both conditions.

### References

1. Boersma, P. Praat, a system for doing phonetics by computer. *Glott Int.* **5**, 341–345 (2002).
2. Kreiman, J. & Gerratt, B. Perceptual interaction of the harmonic source and noise in voice. *J Acoust Soc Am* **131**, 492–500 (2012).

## Annexe 5 : Données supplémentaires de l'étude n°4

### Supplementary materials

#### 1. Corpus of sentences used

The corpus consisted of five different sentences, each pronounced by ten different speakers, with neutral semantic content designed to evaluate the effect of prosody [1]. A total of 100 sentences, with an average duration of 1900 ms, made up the sequence. Smiling and unsmiling sentences were repeated twice as often as their neutral counterparts, resulting in 20 neutral, 40 unsmiling, and 40 smiling sentences. The sequence of sentences was arranged to ensure that there were never:

- Two consecutive sentences pronounced by the same speaker,
- Two consecutive sentences with the same linguistic content,
- Three consecutive sentences with the same filter/prosody,
- Three consecutive sentences pronounced by speakers of the same gender.

The unmodified neutral sentences were included in the sequence to balance the filters and reduce habituation to the stimulus.

#### 2. Sample size estimation

This study was the first combined exploratory study on pupil and facial reactivity to subtle emotional prosodic cues. For this reason, no previous similar study reporting size effect was available, preventing us from estimating an appropriate sample size prior to our study. Nonetheless, the number of trials for each participant was large (n=80). To justify the sample size, we conducted an *a posteriori* power analysis based on Arias et al., 2018 [2] reported size effect on muscular activity (with only 60 trials per subject and a sample of 35 participants) on our 25 participants (power = 0.58 for ZM and power = 0.38 for CS). Despite this limitation, this study provides for the first-time insights on emotional contagion and further work should be done to enrich literature on this phenomenon in response to vocal emotion.

#### 3. Validation of vocal smile model

To validate that the vocal smile model used to artificially modify sentences in this study was the same as in Arias et al.'s [2,3] work, the same task as in Ponsot et al. [4] was performed. However, to enhance feasibility, the task duration was reduced significantly by decreasing the number of trials to 300 (compared to 600 trials in the original task). The task was divided into six blocks of 50 trials each. The last two blocks were identical to measure the internal

consistency of the participants' model, estimate its robustness using the double-pass methodology, and assess internal noise [5,6,7].

The methodology was the same as in Ponsot et al. [4]. This task was based on reverse correlation analysis: in each trial, participants compared two different sounds created from the same original sound and chose the one that sounded the most "smiling." The sounds were modified using the open-source CLEESE toolbox [8] from a 500 ms sound the vowel /a/ pronounced by a male speaker. For a detailed description of the sound modifications and analysis procedures, see Ponsot et al. [4]. To compare the two models of vocal smiles, a statistical comparison of spectral modifications was performed on each modified frequency.

The data of two subjects were removed due to excessive deviation from the control model of Ponsot et al. [4].

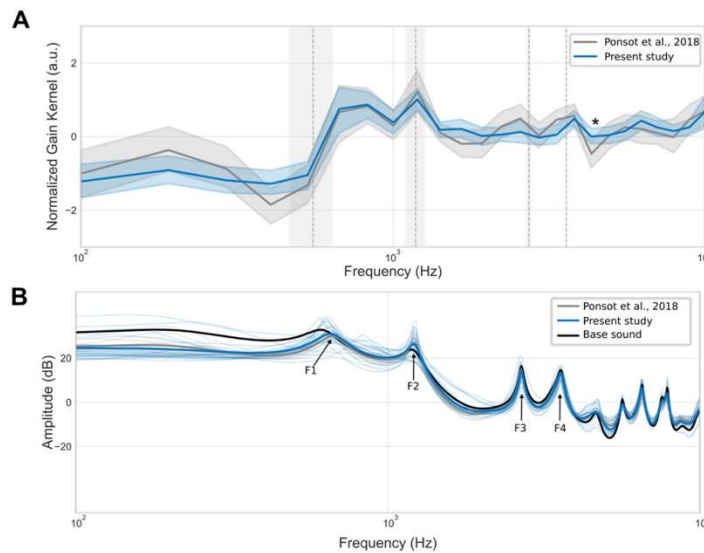


Figure S1: A: Averaged filter underlying the judgment of vocal smile, as derived with reverse-correlation. Asterisks indicate significance differences between the two groups (two-tailed; paired-sample t-tests,  $p < .05$ ). Vertical shaded areas indicate how the first four formants of the voice align with the structure of the Filter. B: Smiling Filter is applied (here with a gain of 0.5) to the original voice, it reveals the internal representation of vocal smile in the two groups compared to the base sound.

The obtained smiling kernel was applied to the original sound and spectral envelope was extracted and compared between groups (Figure S1). An analysis of the acoustic features was performed to determine if formants are modified by the smiling effect. o study the modifications

of frequencies after the application of the smiling filter, the first four formants were extracted using Praat [9]. The results are reported in Table S1.

To determine the influence of vocal smile modifications on the original sound, a comparison analysis was performed between the formant values of the original sound and after modification using a one-sample t-test. The results, along with a comparison between the results of Ponsot et al. [4] and the present study, are reported in Table S1. The only frequency for which there was a significant difference between the two groups was 4332 Hz ( $p < .05$ ). To conclude, as in Ponsot et al., the current sample of participants, even with a smaller number of trials, demonstrated that the mental representation of a vocal smile was the same.

<b>Original sound</b>	<b>F1</b> (555 Hz)	<b>F2</b> (1183 Hz)	<b>F3</b> (2727 Hz)	<b>F4</b> (3591 Hz)
<b>Ponsot et al., 2018</b>	617 ± 83 t(9) = 2.35, p<.05	1187 ± 31 t(9) = .4143, p>.1	2719 ± 28 t(9) = -.88, p>.1	3593 ± 22 t(9) = .35, p>.1
<b>Present study</b>	645 ± 71 t(22) = 6.09, p<.001	1193 ± 27 t(22) = 1.8, p<.1	2733 ± 38 t(22) = .77, p>.1	3601 ± 26 t(22) = 1.9, p<.1
<b>Comparison between Ponsot and Present study</b>	t(15) = -.96, p>.1	t(15) = -.53, p>.1	t(23) = -1.17, p>.1	t(21) = -.91, p>.1

*Table S1: Acoustic characteristics of the sounds before and after the application of the smiling Filter in both groups. Formants frequencies are presented in Hertz (mean ± sd),  $\mu$  corresponds to the frequency value of the formant in the base sound. Results of the 1-sample t-test of the comparison between the value of the formants of the original sound and the modification of smile model.*

#### 4. Motor resonance – effect of the muscle

Congruent responses to a smile could be associated with an increase in Zygomaticus major (ZM) activity compared to Corrugator Supercilii (CS) [10]. To test this assumption, a comparison between the muscular strength of CS and ZM was performed. A repeated-measures ANOVA was conducted to compare mean muscular activity in the 2900-3400 ms time window.

A main effect of muscle was observed ( $F(1,23) = 15.70, p < .001, \eta^2 = .20$ ), with significantly greater activity in CS ( $12.2 \pm 10$  a.u.) compared to ZM ( $-29.6 \pm 6$  a.u.) (Figure S2).

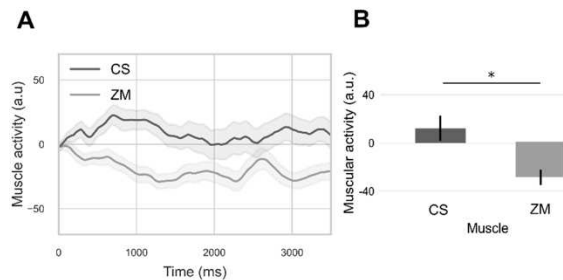


Figure S2: Zygomaticus major (ZM) and Corrugator supercilii (CS) activity. A: Effect of muscle; B: Effect of muscle (CS and ZM) mean response amplitude in the window 2900-3400ms. Shaded areas represent standard error of the mean. \*  $p < .05$

The activity of CS was more pronounced than the activity of ZM. CS activity was more discriminating than ZM activity, which can be explained by the different physiology of these two facial muscles. Despite a common innervation by the facial nerve, they do not depend on the same branch, with CS receiving more nerve impulses than ZM. Additionally, there is less muscular interference in CS than in ZM, making measurement easier and more accurate [11,12,13].

##### 5. Motor resonance – effect of Correct and Incorrect Response

To assess the effect of Response and its interaction with Choice, a repeated-measures ANOVA of muscle activity mean amplitude was performed, with post-hoc pairwise comparisons using Bonferroni correction to further specify significant interaction effects within Correct and Incorrect Responses.

The same time window (2900-3400 ms) was selected for both muscles, and a separate analysis was conducted. A repeated-measures ANOVA was performed to analyze the effects of Choice (Smile vs. Unsmile), Response (Correct vs. Incorrect), and the interaction between these two factors on mean muscular activity.

For ZM, no main effects were observed for either Response ( $F(1,23) = 0.28, p > .1, \eta^2 = 0.002$ ) or Choice ( $F(1,23) = 0.66, p > .1, \eta^2 = 0.01$ ). However, a significant interaction between these two factors was revealed ( $F(1,23) = 4.73, p < .05, \eta^2 = 0.06$ ). To inspect relevant effects,

comparisons within Correct and Incorrect responses were performed with a Bonferroni correction. This analysis revealed a tendency toward a difference between Correct responses according to Choice ( $t(23) = 2.24, p_{\text{corr}} < 0.1$ ). (Figure S3Figure S3). No difference was observed between Incorrect responses ( $t(23) = -1.07, p_{\text{corr}} > 0.1$ ). The significant interaction between Choice and Response confirmed the Filter effect of the main analysis. However, post-hoc comparisons were not significant, which might be the consequence of the subdivision of Filter factor.

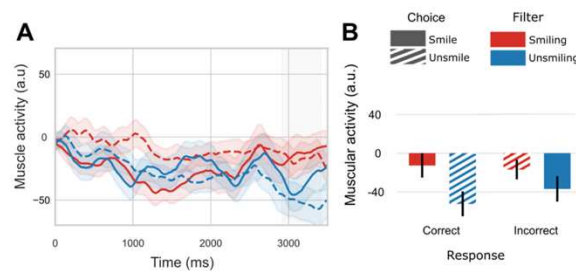


Figure S3: A: Effect of Filter, Choice and Response on ZM activity (without neutral sentences). Shaded areas represent standard error of the mean. B: Barplots represent mean pupil diameter in the window 2900-3400ms (mean  $\pm$  standard error of the mean).

For CS, no main effect was observed for Response ( $F(1,23) = 3.33, p > .05, \eta^2 = .01$ ). A main effect of Choice was revealed ( $F(1,23) = 10.18, p < .01, \eta^2 = .09$ ). No significant interaction between these two factors was demonstrated ( $F(1,23) = 1.89, p > .1, \eta^2 = .01$ ) (Figure S4). Main effect of choice was the same as demonstrated in the previous analysis, confirming that the expression of the emotional explicit motor resonance was supported by CS.

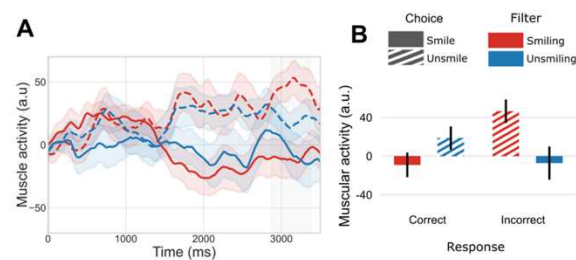


Figure S4: A: Effect of Filter, Choice and Response on CS activity (without neutral sentences). Shaded areas represent standard error of the mean. B: Barplots represent mean pupil diameter in the window 2900-3400ms (mean  $\pm$  standard error of the mean).

## 6. Correlations between physiological measurements

To find a potential link between physiological measures, rating, and EQ scores, a Pearson's correlation matrix was calculated between the discriminating physiological variables (Filter effect on ZM and pupil activity, and Choice effect on CS activity) and accuracy. To highlight these links, for each measure, a difference between responses according to Filter or Choice was calculated, always making this difference in the direction of the effect (e.g., for the ZM, the difference between the response to the Smiling Filter and the response to the Unsmiling filter).

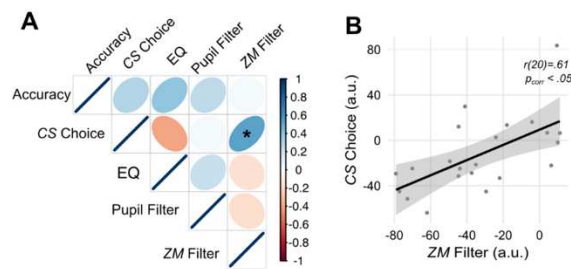


Figure S5 A: Pearson's correlation matrix between physiological measures, rating responses and EQ scores with Bonferroni correction. The scale represents the value of the correlation coefficient ( $r$ ). Accuracy: percentage of good rating of sentences (Smile and Unsmile). CS Choice: difference between the CS response to rated Unsmile and rated Smile sentences (2900-3400ms). Pupil Filter: difference between pupil diameter to Smiling and Unsmiling sentences (2100-2600ms). ZM Filter: difference between ZM response to Smiling and Unsmiling sentences (2900-3400ms). EQ: Empathic quotient; B: Correlation between CS activity according to given Choice and ZM activity according to Filter. \*  $p_{corr} < .05$

No correlation was observed between the measures of fEMG and pupillometry, nor between EQ or Accuracy and any other parameter. However, the correlation between CS activity according to Choice and ZM activity according to Filter was significant ( $r(20) = .61$ ,  $p_{corr} < .05$ ) (Figure S5A). The stronger CS reacted to Choice, the stronger ZM reacted to Filter (Figure S5B).

## 7. References

- [1] Russ, J. B., Gur, R. C., & Bilker, W. B. (2008). Validation of affective and neutral sentence content for prosodic testing. *Behavior Research Methods*, 40(4), 935-939. <https://doi.org/10.3758/BRM.40.4.935>
- [2] Arias, P., Belin, P., & Aucouturier, J.-J. (2018). Auditory smiles trigger unconscious facial imitation. *Current Biology*, 28(14), R782-R783. <https://doi.org/10.1016/j.cub.2018.05.084>
- [3] Arias, P., Bellmann, C., & Aucouturier, J.-J. (2021). Facial mimicry in the congenitally blind. *Current Biology*, 31(19), R1112–R1114. <https://doi.org/10.1016/j.cub.2021.08.059>
- [4] Ponsot, E., Arias, P., & Aucouturier, J.-J. (2018). Uncovering mental representations of smiled speech using reverse correlation. *The Journal of the Acoustical Society of America*, 143(1), EL19-EL24. <https://doi.org/10.1121/1.5020989>
- [5] Burgess, A. E., & Colborne, B. (1988). Visual signal detection. IV. Observer inconsistency. *JOSA A*, 5(4), 617-627. <https://doi.org/10.1364/JOSAA.5.000617>
- [6] Green, D. M. (1964). Consistency of auditory detection judgments. *Psychological Review*, 71(5), 392-407. <https://doi.org/10.1037/h0044520>
- [7] Neri, P. (2010). How inherently noisy is human sensory processing? *Psychonomic Bulletin & Review*, 17(6), 802-808. <https://doi.org/10.3758/PBR.17.6.802>
- [8] Burred, J. J., Ponsot, E., Goupil, L., Liuni, M., & Aucouturier, J.-J. (2019). CLEESE : An open-source audio-transformation toolbox for data-driven experiments in speech and music cognition. *PLOS ONE*, 14(4), e0205943. <https://doi.org/10.1371/journal.pone.0205943>
- [9] Boersma, P. (2002). Praat, a system for doing phonetics by computer. *Glott International*, 5, 341-345.
- [10] Hess, U., Arslan, R., Mauersberger, H., Blaison, C., Dufner, M., Denissen, J. J. A., & Ziegler, M. (2017). Reliability of surface facial electromyography. *Psychophysiology*, 54(1), 12-23. <https://doi.org/10.1111/psyp.12676>
- [11] Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). Psychophysiological science. *Handbook of Psychophysiology*, 2, 3-23.
- [12] Ekman, P., & Friesen, W. V. (1975). *Unmasking the face englewood cliffs*. Spectrum-Prentice Hall.
- [13] Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin*, 95(1), 52-77. <https://doi.org/10.1037/0033-2909.95.1.52>



## Annexe 5 : Liste des publications

- **Publications à comité de lecture international**

**MERCHIE, A.**, Ranty, Z., Wardak, C., Aguillon-Hernandez, N., Bonnet-Brilhault, F., Houy-Durand, E., Aucouturier, J-J., & Gomot, M. Lack of facial motor resonance to vocal emotion in autism: smile is recognized but not embodied, **soumis** à *Nature Human Behavior*

**MERCHIE, A.**, Ranty, Z., Bonnet-Brilhault, F., Houy-Durand, E., Aucouturier, J-J., Gomot, M. (2024). Unveiling Mental Representations of Vocal Smile and Internal Noise in Autism: a Reverse Correlation Study, **en révision** dans *Research in Autism Spectrum Disorders* disponible en pré-print <https://osf.io/preprints/psyarxiv/q4n7z>

**MERCHIE, A.**, Bonnet-Brilhault, F., Escera, C., Houy-Durand, E., & Gomot, M. (2025). Unraveling neural adaptation to vocal and non-vocal sounds in autism. *Clinical Neurophysiology*, 170, 58–66. <https://doi.org/10.1016/j.clinph.2024.12.003>

**MERCHIE, A.**<sup>★</sup>, Ranty, Z.<sup>★</sup>, Aguillon-Hernandez, N., Aucouturier, J.-J., Wardak, C., & Gomot, M. (2024). Emotional contagion to vocal smile revealed by combined pupil reactivity and motor resonance. *Scientific Reports*, 14(1), 25043. <https://doi.org/10.1038/s41598-024-74848-w>, (**★co-first author**)

**MERCHIE, A.**, & Gomot, M. (2023). Habituation, Adaptation and Prediction Processes in Neurodevelopmental Disorders: A Comprehensive Review. *Brain Sciences*, 13(7), Article 7. doi: 10.3390/brainsci13071110

Heurteloup, C.<sup>★</sup>, **MERCHIE A.**<sup>★</sup>, Roux, S., Bonnet-Brilhault, F., Escera, C., Gomot, M. (2021) Neural repetition suppression to vocal and non-vocal sounds. *Cortex*, 148:1-13. doi: 10.1016/j.cortex.2021.11.020 (**★co-first author**)

Díaz-Barriga Yáñez, A., Couderc, A., Longo, L., **MERCHIE, A.**, Chesnokova, H., Langlois, E., Thevenot, C., Prado, J., Learning to run the number line: The development of attentional shifts during single-digit arithmetic (2020). *Annals of the New York Academy of Sciences*, Volume 1477(1), Pages 79–90, ISSN 0077-8923, 1749-6632, doi: 10.1111/nyas.14464

- **Communications orales et affichées en premier auteur**

**MERCHIE, A.**, Ranty, Z., Wardak, C., Houy-Durand, E., Aucouturier, J-J., Gomot, M. Exploring Vocal Smile Integration In Autism: From Acoustic Representation To Facial Motor Resonance, 7th bi-annual European Society for Cognitive and Affective Neuroscience (ESCAN) meeting, Gand (Belgique), Mai 2024 – Poster

**MERCHIE, A.**, Houy-Durand, E., Bonnet-Brilhault, F., Escera, C., Aucouturier, J-J., Gomot, M. Atypical neural adaptation to auditory vocal and non-vocal stimuli in autistic adults, 7eme Journée recherche biomédicale 2023 (Tours, Angers, Orléans, Le Mans), La Riche (France), Décembre 2023 – Poster

**MERCHIE, A.**, Houy-Durand, E., Bonnet-Brilhault, F., Escera, C., Aucouturier, J-J., Gomot, M. Adaptation neurale atypique aux sons vocaux et non-vocaux dans l'autisme, Journée scientifique de la SFR FED Neuroimagerie Fonctionnelle 2023, Tours (France), Juin 2023 – Poster

**MERCHIE, A.**, Escera, C., Gomot, M. Faster neural adaptation to neutral than to emotional prosodic sounds. 21st World Congress of Psychophysiology, Genève (Suisse), Juin 2023 – Poster

**MERCHIE, A.**, Houy-Durand, E., Bonnet-Brilhault, F., Escera, C., Aucouturier, J-J., Gomot, M. Atypical Neural Adaptation to Vocal and Non-Vocal Sounds in Autism. INSAR annual meeting, Stockholm (Suède), Mai 2023 – Poster

**MERCHIE, A.**, Houy-Durand E., Joucla C., Bonnet-Brilhault F., Aucouturier J., Gomot M. Caractéristiques acoustiques et modèle du sourire vocal chez l'adulte autiste. Université d'automne de l'ARAPI, Le Croisic (France), Octobre 2022 – Poster et présentation orale

**MERCHIE, A.**, Roux, S., Bonnet-Brilhault, F., Escera, C., Gomot, M. Effect of the vocal and emotional prosodic content on the neural adaptation to sounds. Federation of European Neuroscience Societies (FENS) Forum 2022, Paris (France), Juillet 2022 – Poster

**MERCHIE, A.** Adaptation neurale aux sons vocaux et non-vocaux dans les troubles du spectre de l'autisme : résultats chez le sujet sain. Rencontre annuelle de la SPNC (Société de Psychophysologie et Neurosciences Cognitives) 2021, Aix en Provence (France), Octobre 2021 – Communication orale

- **Autres communications orales et affichées**

Ranty, Z., **MERCHIE, A.**, Wardak, C., Bonnet-Brilhault, F., Aucouturier, J-J., Auffret, A., Malvy, J., Gomot, M. Sourire vocal et résonance motrice faciale chez l'enfant autiste. Université d'automne de l'ARAPI, Le Croisic (France), Octobre 2024 – Poster et présentation orale

Ranty, Z., **MERCHIE, A.**, Escera, C., Gomot, M. Development of neural adaptation in response to vocal and emotional sounds, Rencontre annuelle de la SPNC (Société de Psychophysologie et Neurosciences Cognitives) 2024, Lille (France), Septembre 2024 – Communication orale

Ranty, Z., **MERCHIE, A.**, Aucouturier, J-J., Wardak, C., Gomot, M. Sourire vocal chez l'enfant : représentation perceptuelle et résonance faciale motrice, Journée scientifique de la SFR FED Neuroimagerie Fonctionnelle 2024, Tours (France), Juin 2024 – Poster

Ranty, Z., **MERCHIE, A.**, Aucouturier, J-J., Wardak, C., Gomot, M. Vocal smile in children: mental representation and facial motor resonance, 7th bi-annual European Society for Cognitive and Affective Neuroscience (ESCAN) meeting, Gand (Belgique), Mai 2024 – Poster



**Annabelle MERCHIE**

## **Traitement de la prosodie émotionnelle dans le trouble du spectre de l'autisme : approche comportementale et neurophysiologique**

### **Résumé en français**

Bien que le diagnostic du trouble du spectre de l'autisme (TSA) nécessite la présence conjointe de troubles de la communication sociale et de comportements répétitifs et spécifiques, peu d'études intègrent à la fois les domaines socio-émotionnel et perceptif. Il reste donc à établir si les difficultés des personnes autistes sont liées à des déficits émotionnels spécifiques, à des particularités sensorielles plus marquées pour les stimuli sociaux, ou à la combinaison des deux. La perception et le traitement des émotions vocales constituent un bon modèle pour explorer chacune des étapes de la boucle *Perception – Représentation – Action* dans l'autisme. Ce travail inclut l'analyse des mécanismes d'adaptation neurale en électrophysiologie, l'examen des représentations internes des émotions vocales en psychophysique, et l'exploration des réactions automatiques induites par les sourires vocaux sur les muscles faciaux et le système nerveux autonome via des mesures électromyographiques et pupillométriques. Alors que les étapes de perception et de représentation des émotions vocales semblent relativement préservées dans l'autisme, les résultats révèlent des différences importantes dans la phase d'Action, suggérant des difficultés d'incarnation des émotions dans l'autisme. Cette approche intégrative originale a permis de dissocier les processus perceptifs et émotionnels, de déterminer à quel niveau les processus physiopathologiques peuvent intervenir, et procure ainsi un nouvel éclairage mécanistique sur les difficultés socio-émotionnelles dans l'autisme.

Mots clés : Autisme, Perception, Émotion, Neurophysiologie, Psychophysique

### **Résumé en anglais**

Both socio-emotional impairments and repetitive behaviors/specific interests are required to establish a diagnosis of Autism Spectrum Disorder (ASD). Yet, little research has attempted to integrate socio-emotional and perceptive domains to explain ASD impairments. It remains unclear whether the difficulties in ASD individuals stem from specific emotional deficits, heightened sensory sensitivities to social stimuli, or a combination of both. Vocal emotion perception and processing serve as a good model for exploring each stage of the *Perception – Representation – Action* loop in autism. This work includes the analysis of neural adaptation mechanisms through electrophysiology, the examination of internal representations of vocal emotions in psychophysics, and the exploration of automatic reactions induced by vocal smiles on facial muscles and the autonomic nervous system through electromyography and pupillometry measurements. While the perception and representation stages of vocal emotions appear relatively preserved in autism, the results reveal significant differences in the Action phase, suggesting difficulties with emotional embodiment in autism. This integrative approach enabled the dissociation of perceptual and emotional processes, pinpointing where the underlying pathophysiological mechanisms may occur, thus providing new mechanistic insights into the socio-emotional difficulties associated with autism.

Keywords : Autism, Perception, Emotion, Neurophysiology, Psychophysics