



<http://portaildoc.univ-lyon1.fr>

Creative commons : Attribution - Pas d'Utilisation Commerciale  
- Pas de Modification 4.0 France (CC BY-NC-ND 4.0)



<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fr>



**N° de mémoire 2413**

Mémoire d'Orthophonie  
présenté pour l'obtention du  
**Certificat de capacité d'orthophoniste**

Par

**AUBERT Juliette**

**Remédiation de la prosodie émotionnelle dans les troubles du  
spectre de l'autisme**

Etudes de cas : enfant et adolescent

Mémoire dirigé par

**BRUNET Laurie  
DIDIRKOVA Ivana**

Membres du jury

**COSNIER Juliette  
MENGARELLI Flavia**

Année académique

**2023-2024**

**INSTITUT DES SCIENCES ET TECHNIQUES DE READAPTATION**  
**DEPARTEMENT ORTHOPHONIE**

Directeur ISTR  
**Pr. Jacques LUAUTÉ**

**Équipe de direction du département d'orthophonie**

Directeur de formation  
**Solveig CHAPUIS**

Coordinateur de cycle 1  
**Claire GENTIL**

Coordinateur de cycle 2  
**Ségolène CHOPARD**

Responsables de l'enseignement clinique  
**Johanne BOUQUAND**  
**Anaïs BOURRELY**  
**Ségolène CHOPARD**  
**Alice MICHEL-JOMBART**

Responsables des travaux de recherche  
**Mélanie CANAULT**  
**Floriane DELPHIN-COMBE**  
**Claire GENTIL**  
**Nicolas PETIT**

Responsables de la formation continue  
**Johanne BOUQUAND**  
**Charline LAFONT**

Responsable du pôle scolarité  
**Rachel BOUTARD**

Secrétariat de scolarité  
**Audran ARRAMBOURG**  
**Danièle FEDERICI**

# UNIVERSITE CLAUDE BERNARD LYON 1

Président  
**Pr. FLEURY Frédéric**

Vice-président CFVU  
**Mme BROCHIER Céline**

Vice-président CA  
**Pr. REVEL Didier**

Vice-président CR  
**M. HONNERAT Jérôme**  
Délégué de la Commission Recherche Secteur  
Santé

Directeur Général des Services  
**M. ROLLAND Pierre**

## **1 Secteur Santé**

U.F.R. de Médecine Lyon Est Doyen  
**Pr. RODE Gilles**

Institut des Sciences Pharmaceutiques et  
Biologiques  
**Pr. DUSSART Claude**

U.F.R. de Médecine et de maïeutique  
Lyon-Sud Charles Mérieux Doyenne  
**Pr. PAPAREL Philippe**

Institut des Sciences et Techniques de la  
Réadaptation (I.S.T.R.)  
**Pr LUAUTÉ Jacques**

U.F.R. d'Odontologie  
**Pr. MAURIN Jean-Christophe**

## **2 Secteur Sciences et Technologie**

U.F.R. Faculté des Sciences  
Directeur **M. ANDRIOLETTI Bruno**

Institut des Sciences Financières et  
d'Assurance (I.S.F.A.)  
Directeur **M. LEBOISNE Nicolas**

U.F.R. Biosciences  
Directrice **Mme GIESELER Kathrin**

Observatoire Astronomique de Lyon  
Directeur **M. GUIDERDONI Bruno**

U.F.R. de Sciences et Techniques des  
Activités Physiques et Sportives  
(S.T.A.P.S.)  
Directeur **M. BODET Guillaume**

POLYTECH LYON  
Directeur **M. PERRIN Emmanuel**

Institut National Supérieure du  
Professorat et de l'Éducation (INSPé)  
Directeur **M. CHAREYRON Pierre**

Institut Universitaire de Technologie de  
Lyon 1 (I.U.T. LYON 1)  
Directeur **M. MASSENZIO Michel**



## Résumé

Les capacités de réception de la prosodie émotionnelle sont souvent déficitaires chez les enfants et adolescents avec un trouble du spectre de l'autisme (TSA), limitant alors la compréhension des énoncés et les interactions sociales. Malgré cela, les pistes de travail proposées dans la littérature et adressées à des orthophonistes pour améliorer cette réception sont minces, et aucun protocole de remédiation n'est disponible. Cette problématique est donc à l'origine de notre travail, et ce mémoire a pour objectif de proposer un protocole afin de répondre à ce manque. Nous émettons l'hypothèse qu'un entraînement pourrait aider à améliorer la capacité des enfants de plus de 7 ans et des adolescents avec un TSA sans déficience intellectuelle à identifier la prosodie émotionnelle. Nous proposons donc un entraînement spécifique qui consiste à identifier les paramètres objectifs et subjectifs de la prosodie, et à les apparier aux émotions correspondantes. Nous avons mené deux études de cas afin de mettre en œuvre le protocole d'intervention, et d'évaluer son efficacité avec une tâche de désignation, auprès d'un enfant et d'un adolescent. Les résultats de notre étude ont montré une nette amélioration du taux de réussite à la tâche de désignation pour le participant adolescent, et une diminution du temps de réponse. Pour le participant le plus jeune, l'amélioration du taux de réussite est plus discrète mais encourageante. Les résultats mettent également en lumière une influence des émotions sur le taux de réponse. Les participants sont réceptifs à l'intervention et offrent des pistes de réflexions sur la subjectivité des perceptions et sur l'utilisation des outils visuels dans cette prise en soin. Enfin, ce protocole est prometteur, et il serait à répliquer en incluant plus de participants pour en connaître exactement les bénéfices.

Mots clefs : Autisme / prosodie / émotion / enfance / adolescence / intervention orthophonique / Etude de cas

## **Abstract**

Children and adolescents with autism spectrum disorders (ASD) often have deficits in the reception of emotional prosody, which limit their understanding and engagement in social interaction. Nevertheless, scant recommendations exist in the literature regarding interventions by speech and language therapists to enhance comprehension, and a standardized remedial protocol remains absent. Consequently, this lacuna has spurred our investigation, with the overarching objective being the formulation of a comprehensive protocol aimed at addressing this gap. We hypothesize that training could improve the ability of children over 7 years of age and adolescents with ASD without intellectual disability to identify emotional prosody. Therefore, we propose specific training consisting of identifying the objective and subjective parameters of prosody and matching them to the corresponding emotions. We conducted two case studies to implement the intervention protocol and evaluate its effectiveness with a naming task involving a child and an adolescent. The results of our study showed a clear improvement in emotion naming for the adolescent participant, with a reduced response time. For the younger participant, the improvement was more subtle but promising. The results also highlight an influence of emotions on response rate. Participants are receptive to the intervention and provide insights into the subjectivity of perceptions and the use of visual tools in this care approach. Finally, this protocol is promising, and it would be worth replicating it by including more participants to precisely understand its benefits.

Key words: Autism / prosody / emotion / childhood / adolescence / speech therapy intervention / Case study

## Remerciements

Je tiens à adresser mes sincères remerciements aux enfants et à leur famille qui m'ont accordé leur confiance, et qui ont participé activement à la mise en œuvre de ce projet. Je remercie mes directrices de mémoire, Ivana et Laurie de m'avoir fait confiance pendant ces deux années, d'avoir suivi avec beaucoup d'attention ma progression, et de m'avoir permis de rendre fièrement cet écrit qui vient mettre un point final à cinq années de formation orthophonique. Ce quinquennat n'aurait pas été le même sans mes maîtres de stage qui ont toutes su me transmettre leur passion du métier et aiguiller mon sens clinique sans entacher ma personnalité. J'adresse un remerciement tout particulier à Sophie pour m'avoir soutenue pendant toute cette année, et pour m'avoir appris à regarder le soin avec un nouveau prisme. Merci à Ludivine de m'avoir fait confiance afin que ce mémoire voit le jour, de m'avoir tant appris sur ces personnes « extra-ordinaires », comme tu les appelles. Je te remercie de m'avoir donné la force pour continuer à pratiquer ce métier avec conviction, détermination mais aussi avec humilité. Ces cinq années ont été donc pleines d'apprentissages autant théoriques, cliniques, que personnels. Merci au groupe 1, sans aucun doute le meilleur groupe de TD. Un immense merci à ces huit nanas incroyables qui ont comblé ma vie étudiante, Amadine, Cléa, Carole, Emilie, Gabrielle, Laura, Mélanie, Mélissandre, et qui ont accueilli avec tendresse et courage mes montagnes russes émotionnelles. Un merci tout particulier à mon acolyte de voyage et d'aventures en tout genre, Gabrielle sans qui le duo iconique « la paire de copine » n'aurait jamais vu le jour, et sans qui l'aventure suédoise n'aurait pas eu la même saveur. Merci à celles et ceux qui suivent mes péripéties depuis de longue date. Merci à ma famille de s'être laissée embarquer, et d'avoir su garder un œil protecteur. Merci à mes parents d'avoir toujours répondu présent, et de m'avoir fait confiance pour que je m'émancipe pleinement. Vous avez été des piliers essentiels pour ma réussite. Merci à ma petite sœur, Estelle, pour garder sans cesse une main tendue vers ta sœur, pour avoir nourri mes réflexions sur le spectre autistique, et changé mon regard sur l'atypie de l'autisme. Merci à ma colocataire et meilleure amie, Clotilde, pour m'avoir challengée, rassurée, soutenue dans toutes les épreuves, et pour m'avoir écoutée parler d'orthophonie pendant de nombreuses heures. Ces six années de vie commune se terminent avec des souvenirs plein la tête. Merci à tous ces amis qui m'ont permis de prendre l'air, qui ont répondu présent pour célébrer des événements, et qui ont toujours décroché leur téléphone en cas de besoin. Merci à Alexandre, pour avoir contribué de très près à la réalisation de cet écrit. Merci pour ta patience, ta curiosité qui ont soutenu mes réflexions à voix haute, et pour ton optimisme qui m'ont aidée à accomplir de nombreux objectifs sur ces deux dernières années universitaires. Enfin, je remercie Laure-Anne Garié dont la lecture de son ouvrage m'a inspiré ce sujet, et vous, cher jury, de l'intérêt que vous porterez à mon travail.

## Tables des matières

I	Partie théorique.....	1
1	La prosodie émotionnelle .....	2
1.1	Les caractéristiques de la prosodie .....	2
1.2	Les marqueurs caractérisant la prosodie émotionnelle .....	2
1.3	La prosodie émotionnelle au cours du développement .....	4
1.4	Mécanisme de réception de la prosodie émotionnelle.....	4
2	Les troubles du spectre de l'autisme (TSA).....	5
2.1	La terminologie.....	5
2.2	Les émotions dans les TSA.....	5
2.3	Les particularités sensorielles à l'écoute de la voix.....	6
3	La prosodie émotionnelle dans le trouble du spectre autistique .....	7
3.1	Le traitement atypique de la prosodie émotionnelle .....	7
3.1	Des signes cliniques avérés .....	7
3.2	Le retentissement d'une réception de la prosodie émotionnelle empêchée .....	8
3.3	Les pistes de remédiation du trouble prosodique dans la littérature .....	9
4	Une proposition d'intervention pour améliorer la réception de la prosodie émotionnelle chez l'enfant et adolescent avec un TSA.....	10
II	Partie méthodologique .....	11
1	Population.....	11
2	Matériel.....	11
2.1	Les phrases .....	11
2.2	Les photographies.....	12
2.3	Les outils d'intervention .....	12
2.4	Les outils de recueil de données .....	13
2.5	Le questionnaire parental.....	13
3	Procédure .....	13
3.1	Procédure générale sur l'ensemble du protocole : tâche de désignation.....	13
3.2	Procédure spécifique aux mesures pré test et post test.....	13

4	Intervention .....	14
III	Résultats.....	16
1	Participants .....	16
2	Tâche de désignation : appariement prosodie émotionnelle et émotion .....	16
3	Matériel phrastique.....	19
4	Questionnaire parental .....	20
5	Réponses à l'intervention .....	21
5.1	Séance 3 : améliorer l'identification de sons aigus et graves.....	21
5.2	Séance 4 : améliorer l'identification des variations de hauteurs .....	21
5.3	Séance 5 : améliorer l'identification du rythme.....	22
5.4	Séance 6 : améliorer l'identification de l'articulation et de la qualité de voix.....	22
IV	Discussion.....	23
1	Recontextualisation .....	23
2	Réception de la prosodie émotionnelle déficitaire dans le TSA.....	23
3	Effets de l'intervention .....	24
3.1	Sur les impacts fonctionnels du trouble prosodique en réception.....	24
3.2	Sur la réception de la prosodie .....	25
4	Réponse au trouble prosodique et particularités de l'intervention dans le cadre des TSA 26	
4.1	Influence de la perception subjective dans le repérage des émotions.....	26
4.2	Influence des particularités sensorielles dans le repérage des émotions .....	27
4.3	Difficultés dans la cohésion des patterns prosodiques pour identifier l'émotion .....	27
4.4	Guidance visuelle au cœur de l'intervention .....	27
5	Limites et perspectives.....	28
5.1	Evaluation du trouble prosodique .....	28
5.2	Cadre.....	28
5.3	Mesure de l'effet de l'intervention .....	29
5.4	Poursuites de l'intervention .....	29
V	Conclusion .....	30



## I Partie théorique

« L'influence des sentiments et la manière dont on les exprime, les expressions et les gestes restent pour eux comme un livre fermé. » (Frith et al., 2010). Cette citation témoigne des difficultés rencontrées par les personnes avec un trouble du spectre de l'autisme (TSA) lorsqu'il s'agit d'interpréter l'expression des émotions d'un locuteur. Cette difficulté pourrait s'expliquer par un manque de sensibilité pour traiter les différents canaux d'expression des émotions comme les mimiques, les gestes ou encore la prosodie.

La prosodie a trait à la modulation acoustique (Garié, 2021) ou à la « mélodie » de la voix. Elle est décrite comme « not what you say but the way you say it » (Peppé, 2015), autrement dit, la prosodie correspond à la forme, à la manière de dire, et non au contenu verbal. Par sa fonction affective, elle porte un message, celui de l'état émotionnel du locuteur (McCann & Peppé, 2003). La prosodie émotionnelle peut alors complètement changer le sens d'un énoncé, ou apporter une information que le contenu verbal seul ne peut pas transmettre. Ainsi, améliorer l'habileté à reconnaître l'émotion par le canal prosodique pourrait être un atout pour les personnes avec un TSA. Or, la littérature s'intéresse peu à la réception de la prosodie chez l'enfant et l'adolescent avec un TSA, et les pistes pour une remédiation sont minces. Dans ce contexte, nous nous questionnons sur les moyens d'améliorer la réception de la prosodie émotionnelle chez des enfants et adolescents avec un TSA.

L'objectif de ce travail est double. D'une part, nous voulons répondre à cette problématique en proposant un protocole de remédiation qui s'adresse aux orthophonistes prenant en soin des enfants et adolescents avec un TSA. D'autre part, nous voulons évaluer l'effet de cette intervention.

Dans un premier temps, nous chercherons à définir d'une part la prosodie émotionnelle, et d'autre part les troubles du spectre de l'autisme. Nous nous attacherons ensuite à décrire les manifestations du trouble prosodique en réception, et les conséquences de ce trouble chez les enfants et adolescents avec un TSA.

Puis, nous mettrons en lumière quelques pistes de remédiation qui ont inspiré la création du protocole. Nous détaillerons la méthodologie employée pour mettre en œuvre ces deux études de cas. Nous préciserons également la procédure d'évaluation, le contenu du protocole et la progression des séances de passation. Par la suite, nous analyserons les résultats de cette évaluation et nous décrirons les remarques des participants lors de la passation du protocole. Nous discuterons par ailleurs des résultats obtenus au regard de la littérature scientifique, et nous les mettrons en lien avec les particularités des enfants et adolescents avec TSA. Enfin, nous proposerons des perspectives encourageantes pour poursuivre ce travail.

# 1 La prosodie émotionnelle

## 1.1 *Les caractéristiques de la prosodie*

### 1.1.1 *Éléments de définition*

La prosodie se définit de façon générique comme la musique de la voix humaine (Lacheret, 2011). Dans la littérature, la prosodie est décrite différemment en fonction du cadre théorique et de la précision d'analyse (Di Cristo, 2004). Une définition plus étayée de la prosodie est donc complexe. La prosodie se réfère aux propriétés suprasegmentales de la parole (McCann et al., 2007 ; Nadig & Shaw, 2012 ; Paul et al., 2005), et elle s'applique à plusieurs domaines linguistiques comme la phonétique, la phonologie, la syntaxe, la sémantique et la pragmatique (Plaza, 2014). Ainsi, on s'attachera dans ce travail à décrire la prosodie par sa « plurifonctionnalité » (Di Cristo, 2004).

### 1.1.2 *Les fonctions de la prosodie*

L'un des rôles principaux de la prosodie est d'apporter du sens au message oral transmis par un locuteur, en habillant les unités linguistiques (les mots et les phrases). La prosodie endosse alors différentes fonctions, par exemple une fonction syntaxique, pragmatique ou émotionnelle (McCann & Peppé, 2003 ; Patel et al., 2020 ; Paul et al., 2005).

La **fonction syntaxique** de la prosodie permet de segmenter les unités linguistiques (mots, syntagmes, énoncés) à l'oral, et donne un rythme, une accentuation et une intonation, pour que le message soit compréhensible par l'interlocuteur (Di Cristo, 2004 ; Lacheret, 2011 ; Plaza, 2014 ; Rossi, 1997). Ensuite, la **fonction pragmatique** véhicule une intention de communication dans laquelle le locuteur va appuyer un élément de son discours, indiquer une ponctuation (McCann & Peppé, 2003), ou transmettre des informations implicites (Di Cristo, 2004). A titre d'exemple, l'énoncé « ça va » peut sous-entendre « Je vais bien. », ou son strict opposé « Je ne vais pas bien. ». L'énoncé accepte (au moins) deux interprétations différentes, simplement par une modulation prosodique. Dans notre exemple, le locuteur indique son état émotionnel, une émotion positive comme la joie, l'excitation, ou une émotion négative telle que la tristesse, la colère, ou le dégoût. La prosodie porte alors une valeur émotionnelle (Lacroix, 2016), on parlera donc de la **fonction affective** (Benedetti et al., 2022 ; Leung et al., 2017 ; McCann & Peppé, 2003). Parallèlement, ces fonctions sont imbriquées à la **fonction sociale** de la communication (Nadig & Shaw, 2012 ; Perrone-Bertolotti & Gillet Perret, 2020). En effet, la prosodie soutient la communication orale, et elle participe au décodage des émotions (Lacheret, 2011), représentant alors des compétences investies par la cognition sociale (Benedetti et al., 2022 ; Laval et al., 2012).

## 1.2 *Les marqueurs caractérisant la prosodie émotionnelle*

Les émotions transmises par la prosodie sont reconnues, entre autres, par l'identification de patterns qui varient pour chaque émotion. La prosodie se mesure objectivement par

l'observation de paramètres acoustiques tels que la fréquence fondamentale, la durée et l'intensité qui varient au cours du flux de parole (Di Cristo, 2004), auxquels s'ajoute la qualité vocale qui se caractérise par la richesse du spectre (Lacheret, 2011). Ces paramètres acoustiques sont reliés à leurs pendants perceptifs, respectivement, la mélodie, la longueur, la sonie et le timbre (Lacheret, 2011). La sonie correspond à « [l'] intensité d'un son appréciée par la sensation auditive qu'il détermine » (LeRobert, 2021).

Concernant la prosodie émotionnelle, les paramètres les plus fiables seraient la fréquence fondamentale, les variations de hauteur (ou l'étendue fréquentielle) et le rythme (débit et pause) (Devillers & Vasilescu, 2004 ; Ruiz, 2012). Il n'existe pas de consensus clair dans la littérature, le tableau 1 propose donc une synthèse des caractéristiques acoustiques et perceptives de trois émotions (tristesse, colère, joie) portées par la prosodie émotionnelle (Ambert-Dahan, 2020 ; Caelen-Haumont, 2005 ; Garié, 2021). Les auteurs s'accordent pour dire que la tristesse a une fréquence fondamentale (F0) inférieure à celle de la colère et de la joie. Le patron mélodique est plutôt descendant pour la tristesse et montant pour la joie, alors qu'il est irrégulier pour la colère. Le débit de parole est lent pour la tristesse et rapide pour la colère. En revanche, le débit concernant la joie dépend des auteurs, tantôt rapide, tantôt lent. L'étendue fréquentielle est étroite pour la tristesse comparativement à la joie et la colère qui ont une étendue plus large. Pour le timbre et la qualité de la voix, le vocabulaire varie en fonction des auteurs, mais nous observons des similitudes.

**Tableau 1**

*Caractéristiques acoustiques et perceptives des émotions dans les productions vocales*

	Fréquence fondamentale F0	Patron mélodique	Etendue fréquentielle	Débit	Articulation	Timbre/ qualité de voix
Tristesse	<b>basse</b>	<b>descendant</b>	étroite [G] [MA]	<b>lent</b>	hypo-tonique [G] ; [MA]	sonore [D] ; résonnante [MA]
Colère	<b>plus haute</b>	descendant [G] ; irrégulier [D] ; abrupt sur les syllabes accentuées [MA]	large [G] [MA]	<b>rapide saccadé</b>	<b>tendue/ dure</b>	retentissante [D] ; soufflée, voix de poitrine [MA]
Joié	<b>plus haute</b>	<b>montant</b>	large [G] [MA]	plus rapide ou plus lent [G] [MA] ; rapide [D]	dure [D] ; normale [MA] ; pas de consensus [G]	assez retentissante [D] ; soufflée, explosive [MA]

*Notes.* Les données sont issues de Garié (2021) [G], de Davitz (1964) [D] (in Ambert-Dahan, 2020), et de Murray et Arnott (1993) [MA] (in Caelen-Haumont, 2005). En caractère **gras**, ces données sont partagées par tous les auteurs.

### **1.3 La prosodie émotionnelle au cours du développement**

Dès la vie fœtale, on remarque une sensibilité du bébé pour la prosodie (Perrone-Bertolotti & Gillet Perret, 2020 ; Plaza, 2014). En effet, le système auditif est viable dès la 26<sup>ème</sup> semaine de gestation et le fœtus peut percevoir des sons, notamment les contours prosodiques qui seront moins filtrés que d'autres éléments de la parole par le liquide amniotique (Perrone-Bertolotti & Gillet Perret, 2020). Cette sensibilité va se poursuivre dès la naissance. Les bébés s'orientent vers la voix de leur mère grâce aux indices prosodiques perçus. C'est d'ailleurs l'utilisation de la prosodie qui va contribuer à l'interaction mère-enfant (Lacroix, 2016). Il existe une période critique entre 9 et 10 mois dans laquelle l'organisation prosodique se met en place (Martel, 2009) pour que l'enfant de 10 à 12 mois puisse utiliser la prosodie pour communiquer (Lacroix, 2016). Laval et al. (2012) rappellent que les jeunes enfants de 3 à 4 ans peuvent interpréter l'état émotionnel de l'interlocuteur en fonction de sa prosodie émotionnelle, lorsque le contexte situationnel est congruent. Par exemple, la prosodie transmet une émotion positive (la joie) dans un contexte situationnel positif (faire du manège) (Aguert et al., 2010). Les enfants étant sensibles au contexte, ce ne serait qu'entre 7 ans et 9 ans qu'ils arriveraient à s'en détacher pour interpréter l'intention du locuteur par le traitement de la prosodie, au même titre que « les autres éléments de la forme linguistique » (Aguert et al., 2009). Dans une situation où le personnage produit un énoncé avec une prosodie « triste », alors qu'il fait du manège (contexte positif), l'enfant identifie la tristesse comme plus forte que la situation positive associée normalement à la joie.

### **1.4 Mécanisme de réception de la prosodie émotionnelle**

La réception auditive est un processus composé de plusieurs étapes successives. Elle commence par l'étape de perception d'un signal sonore tel que la voix. Ensuite, le récepteur identifie le signal perçu comme pertinent et mobilise son attention dessus. Il repère les modulations prosodiques (les variations de hauteur et le rythme) (Garié, 2021). Puis, il va décoder ces modulations prosodiques pour interpréter le message, et comprendre les informations implicites (Garié, 2021) comme l'état émotionnel du locuteur. D'un point de vue neurophysiologique, les aires associatives de l'audition sont situées dans le gyrus temporal supérieur et le sillon temporal supérieur. Elles sont également en connexion avec le lobe préfrontal qui permet d'orienter la perception auditive (Gomot & Wicker, 2012). Ces aires auditives primaires et secondaires sont également connectées à l'amygdale qui joue un rôle central dans le traitement des émotions (Liebenthal et al., 2016).

## **2 Les troubles du spectre de l'autisme (TSA)**

### **2.1 La terminologie**

Dans ce travail, il a été choisi d'utiliser la terminologie de la CIM-11 (2023) « les troubles du spectre de l'autisme ». Ces troubles appartiennent à la famille des troubles du neurodéveloppement (TND) (Crocq & Guelfi, 2015 ; HAS, 2023), ce qui signifie que les signes cliniques apparaissent précocement, au cours du développement de l'enfant. Les TSA se caractérisent sur deux versants aussi appelés « la dyade autistique » (Crocq & Guelfi, 2015). Le premier versant concerne les déficits persistants dans les compétences sociales et communicationnelles. Le second versant concerne les comportements inadaptés à l'âge et au contexte social, qui sont stéréotypés : « restreints, répétitifs et inflexibles » (CIM-11, 2023). Quant à la notion de spectre, elle renvoie à l'expression hétérogène et variable des syndromes cliniques au sein d'un continuum (Garié, 2021 ; Gepner, 2012 ; Rogers & Dawson, 2020 ; Tardif, 2010). Ce continuum est encadré par deux profils aux extrémités du spectre. Le TSA de haut niveau se caractérise par un développement langagier typique, un quotient intellectuel verbal supérieur à la norme, mais quelques difficultés dans le domaine pragmatique. De l'autre côté, le TSA sévère se caractérise par une absence de développement langagier, souvent accompagné d'une déficience intellectuelle (Lartseva et al., 2015). Une personne avec un TSA peut se situer entre ces deux profils.

### **2.2 Les émotions dans les TSA**

Les compétences émotionnelles, sur le versant expressif et réceptif, ainsi que la régulation et la conscience des émotions sont touchées chez l'enfant avec un TSA (Nader-Grosbois, 2020). Cette symptomatologie est parfois comparable à de l'alexithymie qui rend le décodage, la différenciation et l'expression des émotions difficile et qui peut expliquer dans certains cas le manque d'empathie (Lartseva et al., 2015).

Concernant la régulation émotionnelle, les TSA se caractérisent par des difficultés pour interpréter et adapter les émotions par rapport à l'environnement, ce qui peut engendrer des crises de colère, ou des comportements contextuellement inadaptés (Nuske et al., 2023).

Enfin, sur le versant réceptif, les difficultés à percevoir des émotions et à en tenir compte sont largement répandues dans les TSA (Crocq & Guelfi, 2015 ; Nuske et al., 2013). Cependant, elles n'ont pas le même degré de sévérité entre les individus. En effet, certains enfants avec un TSA vont pouvoir étiqueter une émotion alors que d'autres n'en seront pas capables. Ces différences sont également dépendantes du type de tâche proposé. Certaines études ne montrent pas de différence significative entre les enfants avec un TSA et les autres. Néanmoins, plus la tâche est difficile, plus les performances de traitement d'une émotion semblent faibles (Nuske et al., 2013).

En somme, les personnes avec un TSA ne se sont pas déconnectées des processus émotionnels mais traitent les signaux de façon différente. L'identification d'une émotion implique un ensemble de procédures complexes souvent déficitaires dans les TSA telles que la réception des signaux provenant de plusieurs canaux, leur analyse et la mise en lien avec le contexte (Nuske et al., 2013 ; Tardif, 2010 ; Vermeulen, 2020).

### **2.3 Les particularités sensorielles à l'écoute de la voix**

Les particularités sensorielles décrites dans ce paragraphe sont présentes dès l'enfance, chez les personnes avec un TSA, et elles persistent à l'âge adulte (Nader-Grosbois, 2020).

Sur le plan sensoriel, l'enfant avec un TSA présente « des particularités de filtration (hypo- ou hypersensibilité » (Nader-Grosbois, 2020). En effet, certains enfants montrent une surréaction face à la charge sensorielle (hyper-réactivité), ce qui peut engendrer des comportements d'isolement et de repli (Gepner, 2012 ; Meiss et al., 2015). Dans le cas d'un stimulus vocal, certains enfants avec un TSA peuvent montrer une phonophobie, c'est-à-dire un inconfort pour la voix, spécifiquement pour les voix aiguës et certains sons de la parole. En revanche, l'enfant phonophobique investira d'autres sons, notamment des bruits qu'il peut chercher à reproduire, et qui vont mobiliser toute son attention pour devenir, peut-être, son intérêt restreint (Tan et al., 2012). Dans cette même perspective, Vannetzel et al. (2011) montrent que des enfants avec un TSA, entre 7 et 13 ans, sont souvent en difficulté à l'écoute de la voix, stimulus à caractère social, comparativement au traitement de stimuli non-sociaux, comme des bruits. Par contraste, le profil sensoriel d'hypo-réactivité peut également exister (Crocq & Guelfi, 2015 ; Meiss et al., 2015), voire coexister avec le précédent, et il est comparable à une surdit . Par exemple, l'enfant ne r agit pas   l'appel de son pr nom, ni   des sons particuli rement forts (Tan et al., 2012).

Ces difficult s sont expliqu es par diff rentes  tiologies neurophysiologiques. Deux processus du m canisme de r ception de la prosodie seraient notamment impliqu s, la perception et l'attention. Certains travaux montrent une activit  c r brale similaire lors de l' coute des sons vocaux et non-vocaux, traduisant une indiff renciation au niveau du traitement perceptif des sons (Bidet-Caulet et al., 2017 ; Gervais et al., 2004). D'autres travaux montrent en revanche que les difficult s de traitement de la prosodie rel veraient plut t d'un d ficit attentionnel ( eponien  et al., 2003 ; Rogers & Dawson, 2020). Par ailleurs, la capacit  de r gulation sensorielle (soit l'habilit    distinguer et interpr ter un stimulus auditif ou visuel par rapport   d'autres) implique directement l'attention (de Gaulmyn et al., 2018), et le processus attentionnel a  galement une fonction fondamentale au sein du m canisme de la r ception auditive (Ph lip, 2014). N anmoins, il serait faux d'isoler une de ces deux causes (perceptive ou attentionnelle), car toutes les  tapes de la r ception auditive, allant de la perception   l'interpr tation des modulations vocales, peuvent  tre touch es chez l'enfant avec un TSA

(Garié, 2021; Tardif, 2010), et peuvent être majorées, entre autres, par des difficultés attentionnelles (Garié, 2021).

### **3 La prosodie émotionnelle dans le trouble du spectre autistique**

#### **3.1 Le traitement atypique de la prosodie émotionnelle**

##### **3.1 Des signes cliniques avérés**

Les particularités prosodiques des personnes avec un TSA ont été souvent décrites en production. Elles sont caractérisées par une monotonie, des intonations émotionnelles peu marquées, ou au contraire très expressives. Autrement dit, les observations cliniques montrent une perte du caractère naturel de la parole (Crocq & Guelfi, 2015; Garié, 2021; McCann et al., 2007; McCann & Peppé, 2003). Concernant le versant réceptif, l'identification d'une émotion portée par la voix est déficitaire chez les enfants avec un TSA (Nader-Grosbois, 2020). Ils peuvent manifester aucune réaction, ou montrer des difficultés à percevoir, identifier ou encore interpréter des modulations prosodiques à valeur émotionnelle. (Garié, 2021). Ces difficultés s'observent notamment lorsque les variations prosodiques sont moins marquées. Les émotions telles que la fierté ou la honte (Garié, 2021 ; Stewart et al., 2013), plus élaborées que les émotions « de base » (peur, colère, joie, tristesse, surprise, dégoût) (Caelen-Haumont, 2005) sont également difficiles à reconnaître.

##### *3.1.1 Des difficultés nuancées par la démarche expérimentale*

Des travaux de recherche ont montré des difficultés pour interpréter la prosodie émotionnelle seule, chez des enfants et des adultes avec un TSA (Philip et al., 2010 ; Stewart et al., 2013 ; Xavier et al., 2015). Dans ce mémoire, nous partirons du principe que les difficultés observées chez les adultes sont également observables chez les enfants avec un TSA, car les adultes présenteraient moins de difficultés pour traiter la voix et reconnaître des émotions perçues auditivement (Charpentier et al., 2018). En revanche, d'autres travaux n'ont pas validé l'hypothèse d'un déficit d'identification des émotions au travers de la prosodie émotionnelle chez les enfants avec un TSA par rapport au groupe contrôle (Jones et al., 2011 ; Laval et al., 2012). Des éléments pourraient expliquer la divergence des résultats entre ces études. En effet, le type de stimuli proposé est parfois plus percutant que d'autres (comme les pleurs ou les rires). Ensuite, l'âge des participants et l'échantillon de la population varient entre les études. Enfin, l'hétérogénéité du spectre nous rappelle que les difficultés et leur manifestation sont variables (Garié, 2021 ; Rogers & Dawson, 2020).

L'objectif de la méthode expérimentale est d'isoler l'objet d'étude, en l'occurrence la prosodie émotionnelle, de tout distracteur qui pourrait biaiser les résultats. Les auteurs ont donc utilisé des supports neutres au niveau de la sémantique (du sens) et du contexte situationnel, comme la tenue d'une voyelle (Xavier et al., 2015), la production d'une suite de nombres (Philip et al., 2010), des « mmm » (Stewart et al., 2013), ou des vocalisations (pires et pleurs) (Jones et al.,

2011). De fait, ces études montrent parfois des difficultés quant à la réception prosodique chez les enfants avec un TSA. Cependant, elles ne sont pas écologiques, car elles ciblent spécifiquement les capacités d'identification de la prosodie émotionnelle sans contexte situationnel ou linguistique. En effet, elles ne sont pas représentatives de situations quotidiennes où la prosodie est rarement présentée seule. Elle est plutôt contenue dans un contexte sémantique et situationnel, donnant lieu à une multimodalité des sources d'informations.

Cependant, la multiplicité des modalités semble être à l'origine des difficultés rencontrées chez les personnes avec un TSA pour reconnaître les émotions (Laval et al., 2012). Dans leurs travaux, Laval et al. (2012) ont mis en exergue les difficultés des adolescents avec un TSA à comprendre la prosodie émotionnelle, dans le cas où le contexte donné n'est pas cohérent avec l'émotion (le contexte est positif lorsque la prosodie est négative et vice-versa). Bien que cette étude mette en relation l'interprétation de la prosodie au sein d'un contexte à valeur émotionnelle, elle relève surtout un manque d'interprétation de l'ambiguïté. Les travaux de Stewart et al. (2013) partagent des conclusions similaires, et montrent que les participants avec TSA sont sensibles au contexte linguistique qui guide leur interprétation. De fait, traiter les paramètres prosodiques pour en extraire l'émotion est une étape nécessaire mais insuffisante, et doit être ajustée au contexte linguistique et/ou situationnel (Vermeulen, 2020). A titre d'exemple, parler fort ne signifie pas toujours être en colère. Dans un environnement bruyant, le locuteur s'exprimant sur le ton de la joie augmente l'intensité de sa voix pour faire entendre son message (effet Lombard), mais son intention émotionnelle va rester la même, malgré le changement d'un paramètre acoustique (Vermeulen, 2020).

### **3.2 Le retentissement d'une réception de la prosodie émotionnelle empêchée**

#### *3.2.1 Sur la compréhension du langage*

Le défaut de traitement de la prosodie se répercute directement sur la compréhension du langage oral, car l'enfant n'est pas capable d'extraire du sens à ces modulations vocales (McCann et al., 2007) qui, pourtant, participent significativement à la compréhension non-littérale ou pragmatique du message (Martel, 2009 ; Perret & Klink, 2007) et de façon plus globale à l'acquisition du langage (Lacheret, 2011).

#### *3.2.2 Sur l'interaction sociale.*

La prosodie agit comme un liant entre la communication verbale, non-verbale (les gestes, la posture), et l'intention du locuteur. Dans une situation verbale contenant un message implicite, la prosodie émotionnelle porte le sens de la phrase (Lacheret, 2011). Or, dans le cas où la prosodie émotionnelle n'est pas décodée par l'enfant avec un TSA, ce dernier n'a pas accès à la compréhension implicite du message (Robertson & Baron-Cohen, 2017), ce qui freine l'interaction sociale avec son locuteur (Charpentier et al., 2018).

### *3.2.3 Sur la production de la prosodie.*

La production et la réception de la prosodie fonctionnent ensemble. Ainsi, l'enfant aura davantage de difficultés à construire sa prosodie pour enrichir le message, si sa réception est mauvaise. De fait, en travaillant la modalité perceptive, l'enfant accèdera à l'expression pragmatique du langage (Martel, 2009).

## **3.3 Les pistes de remédiation du trouble prosodique dans la littérature**

### *3.3.1 L'évaluation du trouble prosodique*

L'intervention orthophonique est toujours précédée d'une évaluation (Kremer & Lederlé, 2016). Or, concernant la prosodie, il existe peu d'outils d'évaluation quantitatifs ciblant la prosodie émotionnelle. Nous mentionnerons ici le PEPS-C (Profiling Elements of Prosody in Speech-Communication) révisé en 2015 qui évalue les capacités en réception et en production de la prosodie chez les enfants, disponible en français (Peppé, 2015), et dont les données psychométriques sont robustes (Kalathottukaren et al., 2015). Les difficultés peuvent également être observées par le clinicien dans une évaluation qualitative. En effet, dans la Trame d'Observation et le Soins Structuré de la Pragmatique Langage (TOSS-PL), les cliniciens peuvent estimer sur une échelle de cotation « les difficultés possibles pour percevoir, prendre en compte, reconnaître et/ou interpréter, notamment pour des intonations nuancées » la prosodie (Garié, 2021). En outre, l'évaluation peut se concentrer sur l'impact des difficultés prosodiques en réception, au quotidien. Cette évaluation peut être réalisée grâce à un questionnaire parental, tel que l'OCIF (Outil clinique pour les orthophonistes pour documenter les impacts fonctionnels d'un trouble du langage oral ou écrit) (Fortin, 2021) dont certains items mentionnent la prise en compte et l'interprétation des émotions, et grâce à un entretien avec le patient et sa famille, pour établir une décision partagée (HAS, 2013).

### *3.3.2 Les moyens et outils d'intervention*

Tout d'abord, ce travail doit prendre en compte tous les paramètres objectifs et subjectifs de la prosodie, pour être efficace (Paul et al., 2005). L'orthophoniste peut enseigner ces phénomènes acoustiques de façon formelle. Il explique au patient avec un TSA la terminologie reliée à la prosodie comme la fréquence rattachée à son pendant subjectif la hauteur, la durée rattachée au rythme, l'intensité au volume sonore, etc., en s'appuyant sur des supports divers comme des vidéos ou des enregistrements audios (Garié, 2021). Cette méthode d'intervention est d'ailleurs recommandée spécifiquement pour travailler la fonction affective de la prosodie chez des enfants implantés cochléaire. (Ambert-Dahan, 2020). Elle pourrait donc être adoptée pour des enfants avec un TSA, avec le même but thérapeutique. De plus, l'amélioration des gnosies auditivo-verbales (la capacité à reconnaître des bruits et des sons de la langue) est intéressante. Ces gnosies réunissent des compétences de détection et de reconnaissance des paramètres acoustiques tels que la hauteur, l'intensité et

la durée. Une étude propose d'utiliser du matériel non-vocal comme des instruments de musique, puis elle se dirige vers des sons vocaux, avec des émissions de sons, des chansons, et des phrases (Perret & Klink, 2007).

Par ailleurs, les enfants avec un TSA n'auraient pas de difficultés à reconnaître des émotions transmises par la musique par rapport à leur pairs neurotypiques (Bharathi et al., 2019). L'entraînement musical pourrait alors améliorer l'interprétation de la prosodie émotionnelle (Jiam et al., 2017). En effet, la formation musicale améliorerait les capacités de communication et de réciprocité socio-émotionnelle (Bharathi et al., 2019).

Enfin, ces moyens d'intervention ne ciblent pas seulement la compétence de reconnaissance de la prosodie émotionnelle. Ils sollicitent également la compétence attentionnelle, dont le déficit apparaît parfois comme la cause du traitement perceptif empêché (Rogers & Dawson, 2020).

#### **4 Une proposition d'intervention pour améliorer la réception de la prosodie émotionnelle chez l'enfant et adolescent avec un TSA**

Tout l'enjeu dans la prise en soin des troubles du spectre de l'autisme, est de donner aux enfants des clés pour adopter des comportements sociaux ajustés, afin de faciliter les interactions avec leurs pairs, nécessaires à leur développement langagier, social et cognitif (Société canadienne de pédiatrie, 2019). La réception de la prosodie émotionnelle, déficitaire chez les enfants et adolescents avec un TSA, s'inscrit dans cette prise en soin aux multiples facettes, et pose alors la problématique suivante : comment améliorer cette compétence auprès de cette population ?

Ainsi, dans ce travail nous émettons l'hypothèse qu'un entraînement spécifique des paramètres objectifs et subjectifs de la prosodie permettrait d'améliorer l'identification de la prosodie émotionnelle, chez l'enfant de plus de 7 ans (Aguert et al., 2009) et l'adolescent avec un TSA sans déficience intellectuelle.

A ce jour, aucun entraînement spécifique n'a été élaboré. L'objectif de notre travail est donc de proposer un protocole d'intervention pour entraîner l'identification des paramètres prosodiques et leur appariement avec les émotions ciblées. Avec ce protocole, l'orthophoniste pourra travailler la discrimination et l'identification de certains paramètres prosodiques, comme la fréquence (la hauteur), les variations fréquentielles, le rythme, l'articulation et la qualité vocale, de façon isolée, et à partir de matériels majoritairement présents dans les bureaux des orthophonistes. En parallèle, l'orthophoniste cherchera à donner du sens à ces patterns prosodiques reconnus en les associant avec les émotions ciblées.

Cet objectif devra répondre aux critères de l'Evidence Based Practice (spécifique, mesurable, atteignable, réaliste, et temporellement défini).

Nous formulons plusieurs hypothèses partielles pour vérifier l'hypothèse principale. En effet, à la suite de l'intervention, nous nous attendons à une amélioration du taux de réussite d'une tâche d'appariement de la prosodie émotionnelle entendue dans une phrase, avec l'émotion correspondante. Nous nous attendons également à une amélioration du temps de réponse à cette même tâche. D'autre part, nous nous attendons à ce que la joie soit mieux identifiée (en termes de taux de réponses correctes et de vitesse de réponse) que les autres émotions. Parmi la sélection, la joie est la seule émotion positive, donc elle serait plus saillante que les autres. En interrogeant les parents sur les compétences prosodiques et émotionnelles des participants, nous nous attendons à ce qu'ils aient plus de facilité à déduire les émotions des autres dans leur quotidien, et à entrer en relation avec les autres.

## **II Partie méthodologique**

### **1 Population**

Deux enfants ont suivi le protocole expérimental complet, après avoir passé un entretien. Ils ont été recrutés sur le lieu de stage clinique, en cabinet libéral, entre novembre et décembre 2023. La participation au protocole dépendait de la validation aux critères d'inclusion et d'exclusion. Les participants devaient être âgés de 7 ans ou plus, être verbaux, et avoir un diagnostic de TSA. Le protocole devait également s'inscrire dans le projet de soin de l'orthophoniste. Leur orthophoniste, experte de leurs compétences langagières, estimait que le protocole était réalisable. Les critères d'exclusion sont la présence d'une déficience intellectuelle, ou d'une surdité. Les participants et leurs parents ont reçu une notice d'information expliquant le projet, et ont complété un formulaire de consentement.

### **2 Matériel**

#### **2.1 Les phrases**

##### *2.1.1 Le type de phrases*

Nous avons, dans ce travail, élaboré huit phrases différentes pour multiplier les contextes linguistiques (Annexe C). Chaque phrase a été enregistrée sous trois modalités prosodiques ciblées (joie, colère, tristesse et neutre). La modalité « neutre » sert de modalité contrôle. Ainsi, trente-deux enregistrements ont été réalisés. Certains paramètres ont été contrôlés à l'écriture des phrases. Le contenu sémantique est en accord avec le quotidien d'un enfant et adolescent. Les phrases sont de même longueur et même durée (douze syllabes), et elles ne contiennent que des sons voisés, afin d'en faciliter l'analyse acoustique. Elles ne portent pas de valeur émotionnelle d'un point de vue lexico-sémantique, mais peuvent être exprimées avec les trois émotions choisies et admettent alors différentes interprétations.

Chaque énoncé a fait l'objet d'une analyse sur le logiciel Praat (Boersma & Weenink, 2024) afin d'observer leur pertinence au niveau acoustique, en termes de fréquence fondamentale, d'étendue fréquentielle, de profil prosodique et de débit.

### *2.1.2 Les conditions d'enregistrement et d'écoute*

Toutes les phrases ont été enregistrées dans des conditions environnementales identiques, en champ libre, via l'enregistreur vocal de l'ordinateur Dell®. Les enregistrements se sont succédé, la pièce était calme, la distance entre la locutrice et l'ordinateur est restée la même. La voix est celle d'une femme âgée de 24 ans, dont la fréquence fondamentale moyenne est habituellement de 220 Hz. Les phrases enregistrées sont écoutées par le participant en champ libre, via le haut-parleur de l'ordinateur Dell® dont le volume était à 50 %.

### *2.1.3 Le contrebalancement des phrases*

Le contrebalancement des phrases pour chaque passation permettait l'écoute des huit phrases, présentées aléatoirement, et chacune des quatre émotions est au moins présente une fois.

## **2.2 Les photographies**

Pour la tâche de désignation, quatre photographies ont été imprimées, représentant les trois émotions et la modalité contrôle (joie, tristesse, colère et neutre). Elles sont issues de la base de données FACES Database (Ebner et al., 2010) qui propose systématiquement 2 photographies par émotion, pour un même visage. C'est le participant qui sélectionne la photographie parmi la paire qui lui est présentée. Les photos présentées au participant contiennent uniquement le visage d'une femme pour ajouter de la congruence avec la voix qui est entendue.

## **2.3 Les outils d'intervention**

Une des volontés dans la création du protocole est d'utiliser des outils variés et accessibles par tout orthophoniste. Les outils utilisés lors de l'intervention sont accessibles gratuitement sur internet ou souvent présents parmi le matériel orthophonique comme des instruments de musique (xylophone et maracas). Parmi les sites internet, nous avons eu recours à l'Institut Francilien d'Implantation Cochléaire (IFIC, 2007) (exercice de hauteur de note et exercice de taille d'instrument), au ChromMusicLab (Google, 2016) avec ses deux animations : rouleau de piano mécanique et spectrogramme. Enfin, sur YouTube® est disponible pour l'écoute et la visualisation vidéo un extrait du film Vice-Versa sur la chaîne Disney FR (2015). Nous avons également diffusé aux participants la bande originale du film Vice-Versa « Team Building » (Giacchino, 2015) via la plateforme Spotify®.

## **2.4 Les outils de recueil de données**

Un smartphone Samsung Galaxy S8® a permis de recueillir les entretiens des participants et les passations. Les questionnaires pour les parents ont été imprimés pour être complétés, puis scannés et enregistrés. Enfin, les entretiens avec les participants ont été transcrits sur un document Word.

## **2.5 Le questionnaire parental**

Un questionnaire de fonctionnement à la maison (Outil clinique pour les orthophonistes pour documenter les impacts fonctionnels d'un trouble du langage oral ou écrit, OCIF) (Fortin, 2021), a été complété pour chaque participant par les parents, avant et après la phase d'intervention, en fonction de l'âge de leur enfant (tranche d'âge 6-11 ans ou 12-17 ans). Le questionnaire n'a pas été proposé dans son intégralité. Les parents ont complété les items correspondant à l'intitulé « langage et communication », et les items correspondant à l'intitulé « socialisation ». Le nombre d'items varie en fonction des tranches d'âge.

# **3 Procédure**

## **3.1 Procédure générale sur l'ensemble du protocole : tâche de désignation**

Les participants réalisent une tâche de désignation au cours de laquelle ils doivent désigner la photographie correspondant à l'état émotionnel du locuteur. Ils entendent 8 phrases successivement. A chaque fin d'énoncé, le participant doit désigner une photographie parmi 4 : celle illustrant la joie, celle illustrant la colère, celle illustrant la tristesse et celle n'illustrant aucune émotion (neutre). La photographie neutre fait office de modalité contrôle. Une fois que le participant a donné sa réponse, l'expérimentateur fait écouter l'enregistrement suivant. Cette tâche de désignation dure entre 2 et 3 minutes et elle est enregistrée. Pour chaque désignation d'une émotion qui correspond à la prosodie entendue, le participant obtient 1 point. L'expérimentateur note l'émotion désignée par le participant pour chaque phrase. Le temps de désignation est mesuré pour chaque phrase. Lors des séances qui comprennent l'intervention, cette tâche est passée en fin de séance.

## **3.2 Procédure spécifique aux mesures pré test et post test**

Pour les phases de pré test et de post-test, l'expérimentateur commence par un entretien dirigé avec l'enfant, puis il fait passer la tâche de désignation. L'expérimentateur demande également aux parents de remplir un questionnaire.

### **3.2.1 Entretien avec l'enfant**

Un entretien avec chaque participant a été mené en amont de l'intervention afin de s'assurer que les émotions ciblées soient connues et identifiables par le participant. Les participants ont donc défini spontanément ce qu'est la joie, la colère, la tristesse et la neutralité émotionnelle.

Ils ont ensuite donné la couleur qui correspondait pour eux à l'émotion. L'entretien a également permis d'identifier les photographies, parmi l'échantillon présélectionné, qui étaient les plus représentatives de l'expression des émotions. Pour chaque émotion, le participant avait le choix entre deux photographies, dont l'expression faciale de l'une était plus saillante que l'autre. Cette tâche permettait d'assurer la justesse des réponses, pour que le participant n'ait aucun doute à désigner la cible. Dans le cas où le patient n'est pas en mesure de relier les émotions ciblées avec des éléments sémantiques, contextuels lors de l'entretien, alors le protocole ne peut pas être lancé, et un travail sémantique en amont devrait être réalisé.

### *3.2.2 Questionnaire parental*

Le questionnaire parental a pour objectif d'évaluer la présence d'un comportement. Seuls les items concernant les émotions, la prosodie et la socialisation ont été analysés (4 items). Les autres items cotés agissaient comme distracteurs. La cotation soumise aux parents est différente de celle proposée par le questionnaire. Ainsi, pour chaque item, les parents devaient inscrire un nombre allant de 0 à 4 pour décrire la fréquence d'apparition du comportement, 0 pour « jamais » ; 1 pour « rarement » ; 2 pour « à l'occasion » ; 3 pour « généralement » ; 4 pour « toujours ».

## **4 Intervention**

L'intervention se décline en 4 séances, et chacune met en exergue un marqueur caractérisant la prosodie émotionnelle : la fréquence/la hauteur, les variations de fréquence, le rythme/le débit, l'intensité/le volume, et la qualité vocale. Pour chaque séance, un objectif est énoncé. Pour répondre à cet objectif, l'intervention est hiérarchisée selon trois compétences progressives : la discrimination, l'identification et l'appariement du paramètre acoustique travaillé avec l'émotion correspondante (joie, colère et tristesse).

Afin de décrire précisément l'intervention proposée dans ce protocole et de gagner en lisibilité, le protocole d'intervention a été synthétisé dans le tableau en annexe A.

La première séance d'intervention a pour objectif d'améliorer l'identification de la fréquence, des sons aigus et graves. Le participant doit être capable de discriminer si deux sons ont la même hauteur, ou sont de hauteurs différentes. Puis, il doit identifier si un son entendu est aigu ou grave. Enfin, il associe une hauteur à une émotion selon sa perception.

La deuxième séance consiste à améliorer l'identification des variations de hauteur (étendue fréquentielle). L'étape de discrimination revient à comparer deux mélodies et à dire si elles sont identiques ou différentes. L'étape d'identification a pour but d'évoquer les différences entendues entre deux ou trois gammes en tenant compte de l'étendue de la gamme de fréquences, entre autres, le nombre de notes jouées au xylophone. Enfin, l'étape d'appariement consiste à associer une émotion aux variations de notes ou à une étendue de la gamme de fréquences. Par exemple, la joie est associée à une mélodie contenant beaucoup

de notes, donc à une étendue fréquentielle large. Au contraire, la tristesse est associée à une mélodie contenant peu de notes, donc à une étendue fréquentielle étroite.

La troisième séance consiste à améliorer l'identification du rythme. L'étape de discrimination demande au participant de discerner si deux rythmes joués à la suite sont identiques ou différents. Puis, l'étape d'identification cherche à étiqueter le type de rythme joué : rapide ou lent, avec plusieurs outils (maracas, mélodie jouée par un instrument, extraits d'une bande originale). Enfin, l'étape d'appariement consiste à mettre en relation un rythme avec une émotion, toujours avec plusieurs outils en réception (mélodie jouée par un instrument, extraits d'une bande originale), puis en expression les participants sont sollicités pour parler à la vitesse d'un personnage en colère, triste et joyeux.

Une séance intermédiaire entre la troisième et la quatrième consistait à faire un récapitulatif des éléments vus précédemment. Néanmoins, les deux participants n'ont pas pu être présents sur une séance durant la période prévue, donc cette séance récapitulative n'a pas été proposée.

La quatrième séance vise à améliorer l'identification de la qualité de la voix et de l'articulation. Dans un premier temps, les participants discriminent si deux énoncés sont prononcés avec la même intonation ou avec deux intonations différentes. Ensuite, ils identifient des éléments de résonance, à quel endroit du corps la voix de l'expérimentateur résonne-t-elle. Ils sont guidés par le lexique (voix retentissante ou voix de poitrine, voix explosive ou voix de tête, voix résonante ou voix sonore). Ils identifient également l'articulation. Ils sont toujours guidés par le lexique proposé par l'expérimentateur (articulation tendue, normale, relâchée). Nous laissons une trace visuelle pour chaque intonation décrite. Tout d'abord, nous dessinons un visage de profil de façon schématique avec la couleur choisie par le participant. Ensuite, nous mettons en évidence les zones de résonances (dents de scies au niveau des zones concernées), le type de projection vocale et le type d'articulation (différents degrés d'ouverture de la bouche) (Annexe B). Nous demandons au participant d'apparier une émotion à chaque profil. Nous lui proposons aussi de produire des énoncés avec des intonations différentes avec guidance imitative ou sans, et de porter sa main sur la zone de résonance (poitrine, tête, joue). On termine la séance en repérant les personnages d'un extrait du film animé *Vice-Versa* (Disney FR, 2015), et en décrivant leur résonance juste à l'écoute de l'extrait, puis leur articulation à la visualisation de l'extrait. Pour cet extrait, les répliques des personnages ciblés ont été isolées par l'expérimentateur et la vitesse de lecture a été diminuée.

Afin d'adapter le protocole aux contraintes temporelles de l'exercice de mémoire, un rituel de début de séance a été mis en place dès la deuxième session et consistait à apparier une émotion avec le paramètre travaillé précédemment, en production. Par exemple, à la séance 2, nous demandons au participant de produire une note qui correspond à une émotion

travaillée ; à la séance 3, nous demandons au patient de produire une mélodie qui correspond à une émotion travaillée, etc.

### **III Résultats**

Après avoir présenté les deux participants, nous analyserons l'évolution de leur performance au cours du protocole, de façon descriptive. Les résultats sont présentés séparément compte tenu du faible échantillon de cette étude (2 participants). Nous étudierons d'abord l'évolution du taux de réussite à travers le protocole. Nous chercherons également à comparer les taux de réponses correctes pour chacune des modalités émotionnelles, et nous analyserons les réponses des participants au regard des réponses cibles. Nous présenterons ensuite l'évolution des performances en termes de temps de réponse au cours du protocole. Puis, nous nous intéresserons au matériel phrastique. Nous évaluerons d'abord les taux de réussite et les temps de réponses en fonction des phrases, par une analyse descriptive et paramétrique (ANOVA). Nous observerons ensuite les caractéristiques acoustiques des phrases enregistrées. Ensuite, nous décrirons la fréquence d'apparition des comportements ciblés dans le questionnaire parental avant et après l'intervention. Enfin, nous préciserons pour chaque séance d'intervention les réponses et les remarques des participants.

#### **1 Participants**

Deux participants ont été recrutés sur un lieu de stage en cabinet libéral. Le participant 1 est âgé de 15 ans et 2 mois, sa langue maternelle est le français. Il a actuellement trois diagnostics : trouble développemental du langage (TDL), trouble de l'attention avec hyperactivité (TDAH) et trouble du spectre de l'autisme (TSA). Il suit une prise en soin orthophonique à raison d'une fois par semaine, et une prise soin médicamenteuse (ritaline). Le participant 2 est âgé de 7 ans et 8 mois, sa langue maternelle est le français. Il a deux diagnostics, TDL et TSA. Il suit une prise en soin orthophonique, à raison d'une fois par semaine, et une prise en soin médicamenteuse (risperdal).

#### **2 Tâche de désignation : appariement prosodie émotionnelle et émotion**

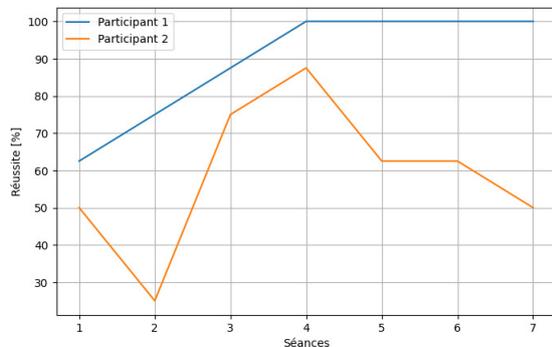
Tout d'abord, nous nous intéressons aux taux de réponses correctes des participants par séance. Les séances 1 et 2 correspondent aux mesures pré-test, les séances 3 à 6 correspondent aux mesures répétées pendant l'intervention, et la séance 7 fait référence à la mesure post-test.

Pour le participant 1, le taux de réponses correctes augmente entre la mesure pré-test (séance 1) et la deuxième séance d'intervention (séance 4) ( $M = 89 \%$ ,  $ET = 15 \%$ ). Puis, il plafonne à 100% de réussite sur les séances suivantes jusqu'à la séance post-test concluant le protocole. Pour le participant 2, les résultats sont variables ( $M = 57 \%$ ,  $ET = 20\%$ ). Néanmoins, l'analyse

visuelle du modèle de régression linéaire (figure 2) montre une légère progression au cours du protocole.

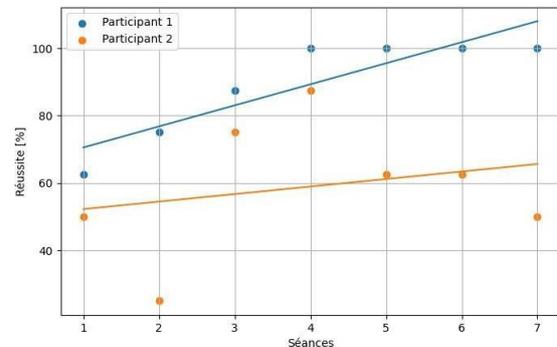
**Figure 1**

*Taux de réponses correctes*



**Figure 2**

*Modèle de régression linéaire*



Le tableau 2 montre le taux de réponses correctes en pourcentage pour chaque émotion, par participant, sur l'ensemble du protocole. Ainsi, nous pouvons classer les modalités prosodiques en fonction de leurs taux de réussite, de la modalité la mieux identifiée à la modalité la moins bien identifiée. Pour le participant 1, la modalité contrôle (neutre) est la mieux identifiée (100 % de réussite), suivie de la modalité « joie » (92.31 %), puis de la modalité « colère » (90.9 %). La modalité « tristesse » est la moins bien identifiée (76.47 %). Pour le participant 2, la modalité contrôle (neutre) a également été la mieux identifiée (80 %), suivie par la modalité « tristesse » (64.71 %), puis par la modalité « joie » (46.15 %). La modalité « colère » a été la plus difficilement reconnue (36.36 %).

**Tableau 2**

*Taux de réponses correctes par émotion (%) sur l'ensemble du protocole*

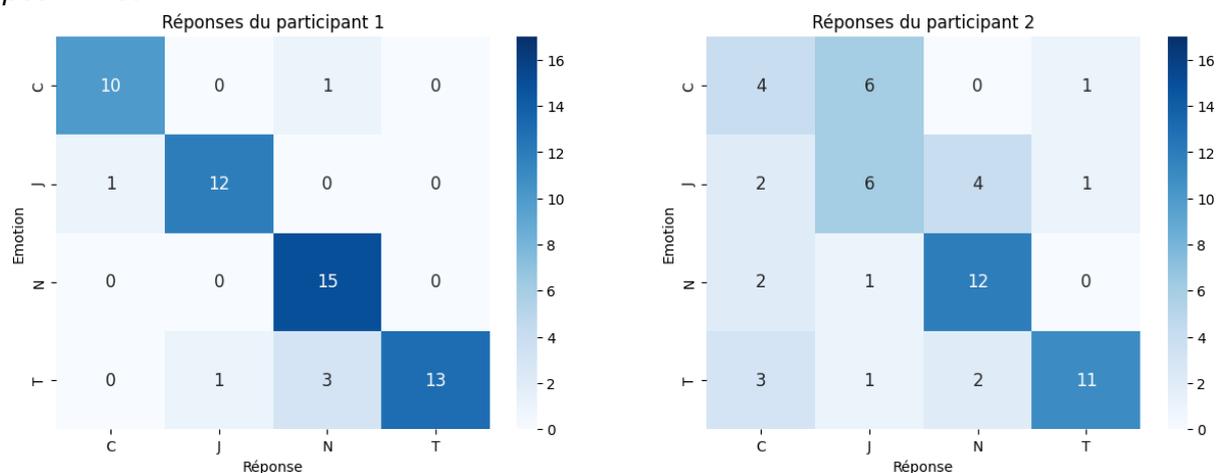
	Participant 1				Participant 2			
	Colère	Joie	Tristesse	Neutre	Colère	Joie	Tristesse	Neutre
<b>Réussite moyenne</b>	90.90	92.31	76.47	100.00	36.36	46.15	64.71	80.00

La figure 3 montre le nombre de réponses accordées aux quatre modalités émotionnelles en fonction de la cible pour chaque participant. Les nuances de bleu correspondent à la concentration des réponses. Les cases plus claires contiennent les concentrations les plus faibles. Les cases plus foncées contiennent les concentrations les plus élevées. Rappelons que pour chaque modalité, le nombre de présentations sur l'ensemble du protocole est différent (colère = 11 fois, joie = 13 fois, tristesse = 17 fois, neutre = 15 fois). Les réponses du participant 1 sont en grande majorité cohérentes avec la cible. Pour les modalités « colère » et « joie », un seul énoncé a été mal identifié pour chacune des émotions. En revanche, le participant associe la modalité « colère » à trois reprises à la modalité « neutre », et une fois

avec la modalité « joie ». Concernant le participant 2, les réponses sont généralement cohérentes avec la cible. Les principales confusions portent sur la colère (identifiée 4 fois sur les 11 présentations) qui est le plus souvent associée à la joie (6 fois). En revanche, la réciproque n'est pas retrouvée. La joie est majoritairement bien identifiée (6 fois sur les 13 présentations), et elle est confondue à 2 reprises avec la colère, mais à 4 reprises avec la tristesse. Enfin, la modalité « tristesse » est correctement identifiée, en majorité (11 fois sur les 17 présentations). Les confusions portant sur cette modalité sont plus hétérogènes et concernent les 3 autres modalités.

**Figure 3**

*Nombre de réponses par émotion relié aux émotions ciblées, sur l'ensemble du protocole pour P1 et P2*

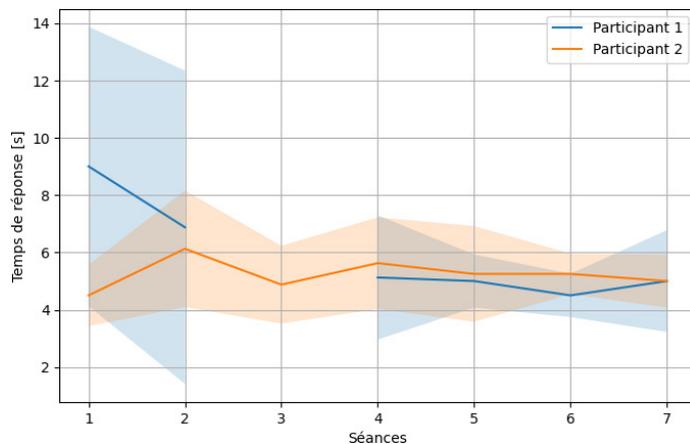


*Notes.* C correspond à la « colère », J à la « joie », T à la « tristesse », et N à « neutre »

Ensuite, nous nous intéressons aux performances des participants en termes de temps de réponse. La mesure à la séance 3 pour le participant 1 n'a pas été faite. Il manque donc une donnée pour avoir une courbe continue. Néanmoins, comme le montre la figure 4, chez le participant 1, le temps de réponse diminue entre la phase de pré-test et la phase d'intervention jusqu'au post test. Par ailleurs, sur cette dernière phase, le temps de réponse se stabilise ( $M = 4.84s$ ,  $ET = 0.28$ ). Pour le participant 2, le temps de réponse est stable ( $M = 5.23s$ ,  $ET = 0.53$ ).

## Figure 4

Temps de réponse moyen (en secondes) pour les 8 énoncés présentés à la tâche de désignation, à chaque séance pour P1 et P2



### 3 Matériel phrastique

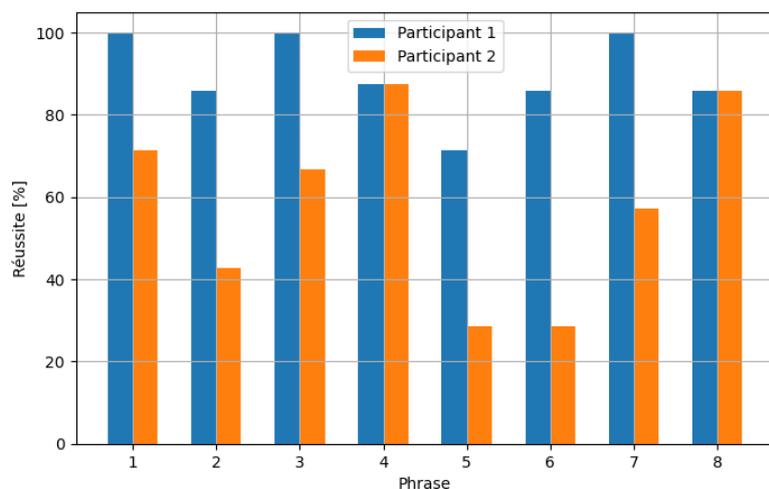
Le contrebalancement des phrases donne lieu à une présentation hétérogène des modalités prosodiques pour chaque phrase. Ainsi, 4 phrases sur les 8 sont entendues au moins une fois dans toutes les modalités prosodiques (joie, tristesse, colère, neutre) sur l'ensemble du protocole. Les quatre autres phrases sont entendues au moins une fois dans 3 modalités sur les 4. Parmi ces quatre phrases, trois phrases n'ont pas été présentées avec l'intonation de la « colère ». A contrario, deux phrases ont été présentées à 4 reprises avec la même modalité, il s'agit de la phrase 2 (avec la colère), et de la phrase 5 (avec la tristesse).

La figure 5 montre le taux de réussite (en %) en fonction des phrases pour les deux participants. Les taux de réussite sont différents en fonction des phrases. Cette variabilité est discrète pour le participant 1 ( $71 \leq P1 \leq 100$ ) et elle est importante pour le participant 2 ( $29 \leq P2 \leq 86$ ). La figure montre également que la phrase 5 est la plus échouée des 8 phrases pour les deux participants. En revanche, aucune phrase ne se démarque comme étant la mieux réussie.

Pour rappel, les participants ont entendu 7 fois chaque phrase au cours du protocole. Pour le participant 1, le score moyen de réussite par phrase est de 6,25 sur les 7 présentations ( $M = 6.25$ ,  $ET = 0.71$ ). Son taux moyen de réussite par phrase est de 89 %. Pour le participant 2, le score moyen de réussite par phrase est de 4,13 sur 7 présentations ( $M = 4.13$ ,  $ET = 1.64$ ). Le taux moyen de réussite est de 59 %. La cinquième phrase de l'échantillon se démarque donc des autres, en étant la plus échouée ( $< -1$  ET) pour les deux participants. Néanmoins, cette démarcation n'est pas significative. En effet, le test ANOVA montre que les phrases n'influencent pas le score des participants ( $p > 0,05$ ).

**Figure 5**

Taux de réussite (%) à la tâche de désignation par phrase, pour P1 et P2



Nous cherchons ensuite à évaluer l'effet de la phrase sur le temps de réponse à l'aide du test ANOVA. Ce test ne montrant pas non plus d'effet significatif ( $p > 0,05$ ), nous ne nous attarderons pas davantage sur cette question.

Les mesures acoustiques des énoncés sont synthétisées dans le tableau 3. Il montre que la tristesse a la fréquence fondamentale (F0) la plus basse, suivie de la joie et de la colère. Le patron mélodique de la tristesse est descendant, alors que celui de la joie est irrégulier (tantôt ascendant, tantôt descendant), et celui de la colère est plutôt descendant et abrupt sur certaines syllabes. L'étendue fréquentielle est étroite pour la tristesse, par rapport à la joie et la colère. Enfin, le débit reste quasiment le même pour les trois modalités.

**Tableau 3**

Analyse acoustique des énoncés pour les modalités tristesse (T), colère (C), et joie (J)

	Fréquence fondamentale F0 (demi-tons)	Patron mélodique	Etendue fréquentielle (demi-tons)	Débit (syllabe/sec)
<b>T</b>	M = 15 ET = 1.48	descendant	M = 9.69 ET = 1.79	M = 2.35 ET = 0.33
<b>C</b>	M = 17 ET = 1.62	descendant, abrupt et irrégulier	M = 12.98 ET = 3.12	M = 2.11 ET = 0.40
<b>J</b>	M = 16 ET = 0.95	montant et descendant	M = 12.73 ET = 2.27	M = 2.47 ET = 0.39

#### 4 Questionnaire parental

Conformément aux réponses du questionnaire, pour le participant 1, le décodage des émotions et des signaux non verbaux reste rare, même après l'intervention. Il cherche généralement à socialiser avec ses pairs, mais ses interactions et participations sociales diminuent. Elles passent respectivement de générales à occasionnelles, et de occasionnelles à rares. Pour le participant 2, son aptitude à décoder le langage augmente légèrement,

passant de jamais à rarement. Il manifeste parfois un intérêt pour les interactions sociales de son âge, il maintient des relations amicales plus fréquentes, mais sa participation à des activités sociales diminue légèrement, passant de toujours à généralement.

**Tableau 4**

*Réponses aux questionnaires parentaux OCIF*

Comportements observables	P1		P2	
	Pré-test	Post-test	Pré-test	Post-test
déduit les émotions et les messages non-verbaux des autres	1	1	0	1
cherche à être en relation avec des gens de son âge	3	3	2	2
maintient de bonnes relations avec ses amis	3	2	1	4
participe à des activités sociales	2	1	4	3

Notes. 0 : jamais ; 1 : rarement ; 2 : à l'occasion ; 3 : généralement ; 4 : toujours

## 5 Réponses à l'intervention

Afin de gagner en lisibilité, nous avons choisi de conserver la numération des séances utilisée précédemment. Ainsi, la séance 3 correspond à la première séance de la phase d'intervention.

### 5.1 Séance 3 : améliorer l'identification de sons aigus et graves

Dans cette séance, nous demandons aux participants d'identifier la hauteur d'un son (aigu ou grave), puis de l'apparier avec une émotion. Les participants n'ont pas de difficultés à différencier les sons aigus des sons graves. Ils appariant clairement la joie avec une fréquence aiguë, et la colère avec une fréquence grave. Pour la tristesse, les résultats sont plus variables. Le participant 2 considère que la hauteur relative à la tristesse est supérieure à celle relative à la joie. Pour le participant 1, elle est située entre les hauteurs associées à la joie et la colère. Par ailleurs, nous observons que le participant 1 sursaute lors de l'entraînement dès que le son est un peu fort, ou lorsqu'il ne s'attend pas à entendre un son.

### 5.2 Séance 4 : améliorer l'identification des variations de hauteurs

Dans cette séance, nous demandons aux participants d'identifier la différence entre 2 mélodies. Est-ce que les notes varient beaucoup ou peu ? L'étendue fréquentielle est-elle large ou étroite ? Puis, les participants doivent apparier une mélodie à une émotion, d'abord en réception à l'écoute d'une mélodie, puis en production avec un xylophone. Les participants identifient bien la différence entre deux gammes (étroite et large) et la décrivent en étant attentifs aux supports visuels (notamment avec le spectrogramme). Lors de l'appariement, le participant 1 ne tient pas compte de l'étendue fréquentielle. Il relie toujours une mélodie à une émotion en fonction de la hauteur de quelques notes, mais pas des variations de hauteur. Selon lui, chaque mélodie a une étendue plutôt étroite dont les notes varient dans l'aigu pour la modalité « joie », dans les mediums pour la modalité « tristesse », et dans les graves pour

la modalité « colère ». Le participant 2 donne des réponses différentes. La mélodie joyeuse a une étendue fréquentielle large et varie sur toute la gamme de notes. La mélodie triste a une étendue étroite, et elle est restreinte à une partie de la gamme, alors que la mélodie « en colère » ne varie pas, elle est jouée avec la note la plus grave. Lorsque les participants jouent une mélodie, ils rajoutent le paramètre de l'intensité à leur production. Ainsi, la colère est jouée avec intensité forte, et la tristesse est jouée, au contraire, avec une intensité faible.

### **5.3 Séance 5 : améliorer l'identification du rythme**

Dans cette séance nous demandons aux participants d'identifier un rythme lent et un rythme rapide, puis de les apparier à une émotion. Les participants identifient la différence entre un rythme rapide et lent. Ils associent un rythme rapide avec la colère et la joie, tandis que le rythme lent est associé à la tristesse. Les participants remarquent que le rythme est aussi corrélé avec l'intensité. Un rythme rapide sera perçu avec plus d'intensité qu'un rythme lent. Lors de la tâche discrimination (pareil ou pas pareil), le participant 1 reste très attentif à la moindre différence entre deux productions, autant du point de vue du rythme, que de celui de l'intensité. De plus, à l'écoute de la bande originale « Team Building », dans Vice-Versa (Giacchino, 2015), les participants ne parviennent pas à identifier l'émotion en lien avec le thème musical. Même une fois la réponse donnée, le participant 1 est en désaccord avec le choix du thème.

### **5.4 Séance 6 : améliorer l'identification de l'articulation et de la qualité de voix**

Dans cette séance, nous demandons aux participants de différencier une articulation « tendue » (hypertonique) d'une articulation « relâchée » (hypotonique). Ils doivent également identifier les zones de résonance (poitrine, tête, bouche), et le type de résonance (retentissant, résonnant, peu résonnant). Ils apparient ensuite la qualité vocale et l'articulation à une émotion. Bien que nous distinguions la qualité articulatoire (tendue, normale et relâchée) de l'aperture, les participants sont plus réceptifs à la description de l'aperture pour identifier les émotions. Dans le cas de la colère et de la joie l'aperture est grande, alors que pour la tristesse l'aperture est plus faible. Concernant la qualité vocale de ces deux émotions, les participants la décrivent en fonction de l'intensité. En effet, la voix retentit ou résonne parce qu'elle est forte. Enfin, en orientant l'attention sur les zones de résonance, dont la poitrine pour la colère, le participant 1 relève spontanément des marques de tensions musculaires au niveau du cou pour exprimer cette émotion. Le participant 2 quant à lui ne relève pas ces marques. Pour finir, lors de l'écoute de l'extrait vidéo, les participants écoutent à plusieurs reprises l'extrait, dont la vitesse de lecture est progressivement diminuée, permettant aux participants d'identifier plus facilement les personnages.

## IV Discussion

### 1 Recontextualisation

L'objectif de ce travail était de créer un protocole d'intervention orthophonique permettant d'améliorer les compétences impliquées dans le mécanisme de réception de la prosodie émotionnelle, pour la joie, la tristesse et la colère. Le protocole a été testé auprès d'un enfant et d'un adolescent avec un TSA. En effet, les capacités de réception de la prosodie émotionnelle sont souvent déficitaires chez les enfants et adolescents avec un TSA (Garié, 2021 ; Nader-Grosbois, 2020 ; Philip et al., 2010 ; Stewart et al., 2013 ; Xavier et al., 2015) et entraînent des limitations en termes de compréhension du langage (Martel, 2009 ; Perret et al., 2012), d'interaction sociale (Charpentier et al., 2018), ou encore de production de la prosodie (Martel, 2009).

Ce protocole d'intervention cible les paramètres ou marqueurs prosodiques qui se sont révélés saillants dans la littérature tels que la fréquence fondamentale et ses variations, le rythme, l'articulation et la qualité vocale (Devilleers & Vasilescu, 2004 ; Ruiz, 2012). L'entraînement proposé vise les compétences d'identification des paramètres acoustiques et perceptifs de la prosodie, et leur appariement avec les émotions ciblées (joie, colère, tristesse).

Tout au long du protocole (avant et après les séances qui précédaient la période d'intervention, et à la fin de chaque séance d'intervention), les participants devaient désigner l'émotion illustrée par une photographie correspondant à l'intonation prosodique d'une phrase entendue. Les photographies correspondant à chaque émotion étaient sélectionnées par le participant au sein d'un échantillon proposé, en amont de l'intervention (Ebner et al., 2010). A chaque recueil de données, 8 phrases enregistrées ont été présentées aux participants dans les mêmes conditions d'écoute. Ces phrases n'ont pas été utilisées comme un outil d'entraînement pendant l'intervention. Nous supposons que l'entraînement améliorerait l'identification de la prosodie émotionnelle pour les trois émotions ciblées, en termes de taux de réussite et de temps de réponse à la tâche de désignation (H1). Nous supposons également que la joie serait l'émotion la mieux identifiée parmi les 2 autres, étant la seule émotion positive (H2). Enfin, en interrogeant les parents au travers d'un questionnaire, nous supposons que les comportements de déduction des émotions et de relations aux autres seraient présents plus fréquemment dans le quotidien des participants (H3).

### 2 Réception de la prosodie émotionnelle déficitaire dans le TSA

Les deux participants ont été vus en bilan initial en orthophonie pour des difficultés de langage. Le diagnostic de trouble développemental du langage (TDL) a précédé celui du TSA. Ce double diagnostic (TSA associé à un TDL) renforcerait les difficultés pour identifier les

émotions quand elles sont présentées sous la modalité auditive seule (Xavier et al., 2015), ce qui correspond à la modalité utilisée dans ces études de cas.

Par ailleurs, le questionnaire parental (OCIF) (Fortin, 2021) montre des difficultés importantes chez les deux participants pour « déduire des émotions et comprendre des messages non-verbaux » (Fortin, 2021). Ces difficultés sont des marqueurs cliniques associés aux TSA (Crocq & Guelfi, 2015). Il est donc évident qu'elles soient retrouvées dans les réponses au questionnaire. Néanmoins, ces difficultés ne sont pas seulement expliquées par un défaut de décodage de la prosodie émotionnelle. En effet, pour déduire les émotions d'un tiers, l'interlocuteur doit tenir compte de la prosodie, mais aussi du contexte linguistique, du contexte situationnel, des éléments gestuels, ou encore des mimiques. Ainsi, la prosodie est un paramètre parmi tant d'autres qui permet de décoder les émotions. De fait, les difficultés mentionnées par les parents s'expliqueraient par un défaut de décodage de plusieurs paramètres (verbaux, non-verbaux et prosodique) qui empêchent la reconnaissance des émotions. Ces difficultés pourraient aussi s'interpréter comme un défaut de la théorie de l'esprit (Duval et al., 2011) qui est la capacité à savoir ce que pense l'autre (Nader-Grosbois, 2020), et à déduire ses émotions. D'ailleurs, un défaut de la théorie de l'esprit est souvent présenté comme à l'origine d'une « cécité » des émotions dans les TSA (Nuske et al., 2013). Concernant les items du questionnaire qui touchent à la socialisation, nous nous attendions à des difficultés majeures. Pour rappel, les TSA se caractérisent par des déficits persistants dans les compétences sociales (CIM-11, 2023). L'interaction sociale serait même parfois empêchée par le manque de compréhension de la prosodie qui véhicule un message implicite (Charpentier et al., 2018 ; Robertson & Baron-Cohen, 2017). Cependant, pour le participant 1, les comportements semblent être présents plutôt fréquemment, ne suggérant pas d'impact majeur dans les interactions sociales. Pour le participant 2, les résultats semblent plus contrastés. Il serait en difficulté pour être en relation avec ses pairs, mais participerait aux activités sociales. Ainsi, l'évolution entre les phases de pré-test et de post-test ne sera pas aussi déterminante pour évaluer l'impact du trouble prosodique sur l'interaction sociale.

### **3 Effets de l'intervention**

#### **3.1 Sur les impacts fonctionnels du trouble prosodique en réception**

Selon les parents, les participants ne déduisent pas plus les émotions et les messages non-verbaux des autres avant et après l'intervention. De plus, les comportements relatifs aux relations aux autres ne sont pas plus présents après la passation du protocole. Nous ne pouvons pas valider l'hypothèse (H3) pour laquelle l'intervention a un effet positif sur la fréquence des comportements sociaux et langagiers ciblés. Rappelons que le questionnaire manque de précision, comme discuté précédemment, ce qui pourrait induire une interprétation différente entre les deux évaluations.

### **3.2 Sur la réception de la prosodie**

Pour le participant 1, les scores montrent une amélioration des performances jusqu'à atteindre une phase de plateau, où le taux de réussite est maximal. Néanmoins, il y a eu une augmentation des performances entre les deux pré-tests. Cette amélioration se poursuit de façon linéaire jusqu'à la phase de plateau. La vitesse de réponse est également améliorée dès la deuxième mesure pré-test, et elle reste stable sur l'ensemble du protocole. Ainsi, le progrès pourrait être davantage lié à l'exposition aux phrases qu'à l'entraînement. Le participant s'habitue aussi à la tâche ce qui lui permet de l'anticiper, et qui expliquerait l'amélioration rapide des temps de réponses. L'hypothèse (H1) est donc partiellement validée. Une succession d'au moins 5 mesures répétées dans une phase de pré-test pourraient répondre à cette question, comme le préconise la méthodologie SCED (Krasny-Pacini & Evans, 2018). Pour le participant 2, l'ensemble des résultats montre une tendance d'amélioration très faible mais encourageante, et traduisent la difficulté d'interprétation de la prosodie émotionnelle. Cette difficulté pourrait être normale et corrélée à l'âge de développement de la prosodie émotionnelle en réception. En effet, le participant 2 est âgé de 7 ans et 8 mois au début de l'étude. Or, ce ne serait qu'à partir de 7 ans que l'enfant interpréterait l'état émotionnel de son interlocuteur par l'information prosodique seule (Aguert et al., 2009). Ainsi, on pourrait considérer que cette compétence est en émergence pour cet enfant. La fluctuation des scores du participant peut aussi être expliquée par un manque de motivation ou de disponibilité au moment de la tâche de désignation. Les changements de comportement sont très fréquents dans les TSA, et les enfants peuvent montrer de grandes différences de performances (Tardif, 2014).

Les participants identifient bien la modalité de contrôle. Ils sont donc capables de discriminer la présence ou l'absence d'émotion en tenant compte de la prosodie. En revanche, quand la prosodie est vectrice d'une émotion, les scores sont plus variables d'une modalité émotionnelle à l'autre. Cela suggère que les émotions ciblées ne sont pas reconnues de manière homogène. Pour le participant 1, la joie est la mieux reconnue des trois, validant l'hypothèse (H2) que l'émotion positive se démarque des deux émotions négatives. Pour le participant 2, la joie n'est pas la mieux reconnue, ce qui invalide l'hypothèse (H2). La tristesse est beaucoup mieux identifiée que les deux autres. Cela peut s'expliquer par les caractéristiques prosodiques de la tristesse qui se distinguent de celles relatives à la joie et la colère. En effet, les variations de hauteur sont beaucoup plus faibles pour la tristesse, et le rythme est plus lent, alors que la joie et la colère partagent des paramètres acoustiques communs. Ces deux émotions ont une fréquence fondamentale plus élevée, des variations de fréquence plus importantes, et un débit plus rapide (Ambert-Dahan, 2020 ; Caelen-Haumont,

2005 ; Garié, 2021). D'ailleurs, la colère est plus souvent étiquetée comme de la joie par le participant 2.

L'interprétation des confusions des participants avec les données de la littérature semble globalement justifiée. En effet, l'analyse acoustique des énoncés est corrélée aux données de la littérature en termes de fréquences fondamentale, étendue fréquentielle et patron mélodique (Ambert-Dahan, 2020 ; Caelen-Haumont, 2005 ; Garié, 2021). En revanche, nous observons peu de différences au niveau du rythme, ce qui est corrélé à la longueur des énoncés relativement courts. Alors, le rythme ne serait pas un paramètre pertinent sur lequel s'appuyer pour améliorer son taux de réponse avec cette longueur d'énoncés. De plus, l'écart entre les fréquences des trois modalités est très fin pour être perçu.

## **4 Réponse au trouble prosodique et particularités de l'intervention dans le cadre des TSA**

### **4.1 Influence de la perception subjective dans le repérage des émotions**

A chaque séance, nous demandons au participant d'apparier un pattern prosodique travaillé avec une émotion. Nous précisons aux participants qu'il n'y a pas de mauvaises réponses, car la perception des différemment paramètres acoustiques est subjective. Par exemple, lors de la première séance, les participants apparient une hauteur avec une émotion. Selon la littérature, une hauteur grave serait associée à la tristesse, alors que la joie et la colère seraient associées à une hauteur plus élevée (Ambert-Dahan, 2020 ; Caelen-Haumont, 2005 ; Garié, 2021). En revanche, les deux participants apparient clairement la joie avec une hauteur aiguë, et la colère avec une hauteur grave. Pour la tristesse, les résultats sont dépendants des participants. Ainsi, l'appariement hauteur-émotion ne correspond pas aux mesures acoustiques de la littérature (Ambert-Dahan, 2020 ; Caelen-Haumont, 2005 ; Garié, 2021). Les indices mélodiques soutenus par la fréquence fondamentale seraient décrits selon une échelle subjective (Bel, 2008). Cette subjectivité perceptive ne nous permettrait donc pas d'établir des règles universelles pour identifier les émotions. Nous laissons donc les participants libres de leur interprétation.

Ainsi, il existerait bien une influence de la subjectivité dans le traitement des perceptions auditives, en tout cas pour associer une hauteur à une émotion. En tenant compte des perceptions de chacun, est-ce que ce paramètre est fiable ? Rappelons que les résultats du participant 1 s'améliorent au cours du protocole, alors que ses appariements (hauteur-émotion ou variation de fréquences-émotion) ne sont pas corroborés aux données de la littérature (Ambert-Dahan, 2020 ; Caelen-Haumont, 2005 ; Garié, 2021). Alors, deux questions émergent. La subjectivité perceptive du participant est-elle suffisamment fiable ? Tiendrait-il compte d'autres paramètres acoustiques plus saillants ?

Les deux participants semblent tenir compte de l'intensité de façon prédominante. L'intensité est un paramètre acoustique souvent identifié par les participants, peu importe la séance. Bien que leurs appariements intensité-émotion soient justes (Ambert-Dahan, 2020 ; Caelen-Haumont, 2005), elles comportent un biais. En effet, l'intensité est très dépendante du contexte environnemental. Dans un environnement bruyant, la voix sera plus forte. Ainsi, nous ne pouvons pas tenir compte uniquement de ce paramètre pour qualifier l'émotion du locuteur (Vermeulen, 2020). Cette raison nous a confortée à ne pas travailler ce paramètre acoustique en séance d'intervention.

#### **4.2 Influence des particularités sensorielles dans le repérage des émotions**

Retrouvés dans le tableau clinique des TSA, les particularités sensorielles peuvent entraver la perception (Nader-Grosbois, 2020), notamment la perception auditive (Tan et al., 2012). En effet, lors du protocole, le participant 1 était très attentif au moindre changement de rythme ou de hauteur, et pouvait sursauter quand le son était trop fort, ce qui correspondrait à un profil d'hyper-réactivité (Gepner, 2012 ; Nader-Grosbois, 2020).

Par ailleurs, les troubles perceptifs observés dans les TSA pourraient être expliqués par la présentation rapide des informations auditives (Lainé et al., 2009). C'est la raison pour laquelle nous diminuons volontairement la vitesse de lecture des vidéos utilisées dans l'intervention qui semble soutenir l'identification des émotions. Le ralentissement du débit de la parole permettrait de pallier les difficultés de traitement des stimuli sensoriels perçus trop rapidement (Lainé et al., 2009 ; Tardif, 2010).

#### **4.3 Difficultés dans la cohésion des patterns prosodiques pour identifier l'émotion**

Le traitement simultané de plusieurs informations est une compétence souvent déficitaire dans les TSA (Nuske et al., 2013). Le manque de cohérence centrale s'observe quand le sujet traite les détails avant le tout (Gepner, 2012 ; Meiss et al., 2015). L'enfant éprouvera alors des difficultés à mettre en cohésion différents éléments entendus pour l'interpréter comme une unité globale. Cette difficulté est retrouvée pendant les séances d'intervention. Dans les extraits de la bande son Vice-Versa (Giacchino, 2015), les paramètres acoustiques se superposent. Par exemple, le thème de la tristesse est une mélodie lente et grave. Les deux participants associent cette mélodie à la colère, car ils sont plus sensibles à la hauteur (aigu, grave) qu'au rythme (lent, rapide). Ils n'arrivent pas à mettre en cohésion les éléments acoustiques (rythme, hauteur, variation de hauteur) pour identifier le thème de l'émotion.

#### **4.4 Guidance visuelle au cœur de l'intervention**

Nous partons de l'idée que certaines personnes avec un TSA ont une « pensée visuelle » (Kunda & Goel, 2011). Pour guider le traitement de l'information auditive seule, nous leur

proposons alors plusieurs supports visuels, comme l'animation de différentes mélodies (rouleau de piano mécanique), et le spectrogramme (Google, 2016). Le spectrogramme permet un feedback visuel simultané de la parole du locuteur et du participant, bien utilisé dans les rééducations vocales en orthophonie (Parmentier et al., 2014). Dans ce cas de figure, il illustre les variations de fréquence et la gamme de fréquence. Plus l'image spectrale est large, plus il y a de fréquences produites (Parmentier et al., 2014).

Les dessins ont également été un support visuel lors de la séance sur l'articulation et la qualité de la voix, afin de donner du sens aux concepts utilisés (zones de résonance, voix explosive, aperture).

Pour résumer, notre étude a montré une nette amélioration du taux de réussite à la tâche de désignation pour un participant, et une amélioration plus discrète mais encourageante pour le deuxième participant. Comme discuté précédemment, les facteurs influençant la progression sont nombreux, mais ils pourraient être expliqués par l'intervention. Cette dernière nous apprend à tenir compte de la subjectivité des perceptions qui guident la reconnaissance des émotions, et à soutenir cette perception par l'indigage visuel.

## **5 Limites et perspectives**

### **5.1 *Evaluation du trouble prosodique***

Bien que le trouble prosodique soit identifié dans le tableau clinique du TSA (Nader-Grosbois, 2020), il n'existe pas d'évaluation des difficultés de réception de la prosodie émotionnelle. Le questionnaire parental (OCIF) (Fortin, 2021) semblait le plus proche de nos attentes pour évaluer le trouble de la prosodie émotionnelle au quotidien et ses impacts. Cependant, il n'a pas permis de mettre en exergue le trouble prosodique seul, étant imbriqué dans un réseau de compétences sous-jacentes (Nuske et al., 2013). Il serait intéressant d'avoir un questionnaire parental plus spécifique, ou un outil d'évaluation de la prosodie émotionnelle, étalonnée en français, et couvrant plusieurs émotions (au-delà des émotions de base) (Kalathottukaren et al., 2015). A ce jour, nous pouvons nous appuyer sur des données présentes dans la littérature pour établir une échelle développementale pour la prosodie émotionnelle, chez l'enfant et l'adolescent (Aguert et al., 2009 ; Martel, 2009 ; Plaza, 2014).

### **5.2 *Cadre***

La durée du protocole était limitée par des contraintes temporelles, alors que plusieurs sous-objectifs étaient menés. Il s'agissait de répondre à un sous-objectif par séance d'intervention qui ciblait un des patterns prosodiques (hauteur, variations de hauteurs, rythme, articulation et qualité vocale), en le travaillant en discrimination, identification et appariement. Ainsi, le protocole pourrait s'organiser sur une durée plus étendue où chaque sous-objectif serait

décliné en deux séances, ce qui permettrait de travailler l'appariement du pattern prosodique avec les émotions, dans une séance dédiée. De plus, nous ne pouvons pas mesurer l'effet d'une seule séance d'intervention. En effet, le protocole aurait pu se concentrer sur un seul pattern prosodique pour évaluer l'effet de l'entraînement. Par exemple, seul l'entraînement portant sur l'identification des variations de hauteurs aurait peut-être suffi à améliorer la reconnaissance de la prosodie émotionnelle.

L'absentéisme des deux participants n'a pas permis de mettre en œuvre la séance récapitulative initialement prévue. Le participant aurait manipulé tous les paramètres travaillés précédemment de façon globale : la fréquence, les variations de fréquences et le rythme. Cette séance aurait fait intervenir la compétence de cohésion des éléments (cohérence centrale) qui est déficitaire dans les TSA (Gepner, 2012).

### **5.3 *Mesure de l'effet de l'intervention***

Les résultats obtenus pour la tâche de désignation montrent une progression linéaire des performances en termes de taux de réussite entre la mesure pré-test et la deuxième séance d'intervention. Cette progression est forte pour le participant 1 et légère pour le participant 2. L'effet re-test pourrait avoir davantage d'influence sur le taux de réussite que l'intervention elle-même. De fait, il serait intéressant de contrôler l'effet test et re-test en adoptant la méthodologie SCED, avec un minimum de deux phases. Une phase sans intervention, sur 5 séances, à distance d'une phase d'intervention (Krasny-Pacini & Evans, 2018). D'autre part, ce protocole a été créé avec l'ambition d'être adopté par tout orthophoniste qui le souhaite, notamment grâce à l'accessibilité des outils. Ainsi, il serait intéressant de faire valider ce protocole par des orthophonistes expertes.

### **5.4 *Poursuites de l'intervention***

Le protocole a été pensé pour améliorer la reconnaissance de trois émotions « de base » par le canal de la prosodie émotionnelle. De fait, il serait intéressant d'observer un transfert ou non des compétences travaillées à d'autres émotions. Dans le cas où le transfert n'a pas lieu, le protocole pourrait s'étendre à d'autres émotions, notamment celles qui sont plus difficiles à appréhender (fierté et honte par exemple) (Garié, 2021 ; Stewart et al., 2013). Par ailleurs, la tâche de désignation n'est qu'une étape dans la reconnaissance des émotions. Le patient pourra dans un second temps dénommer l'émotion sans avoir d'indication visuelle. Ensuite, une fois que le patient est en mesure de se représenter la prosodie des émotions, on lui demandera d'ajouter du sens à la phrase entendue. Par exemple, la phrase « Les amis de Léo arrivent dans la journée » n'aura pas la même signification prononcée avec une prosodie joyeuse, triste ou en colère. Alors, dans une tâche de « décision sémantique », le patient devra admettre ou rejeter la proposition faite en fonction de l'intonation, ou alors choisir entre 2 interprétations.

Pour reprendre notre exemple, le patient aurait le choix entre « la fille est impatiente, elle apprécie beaucoup les amis de Léo » ou « la fille n'apprécie pas les amis de Léo ». Cette tâche fait intervenir des compétences d'implicite. Enfin, l'étape finale serait celle en évocation spontanée dans laquelle le patient devrait expliquer le sens de la phrase.

Le protocole et son évaluation proposés ainsi semblent s'inscrire dans un axe de travail parallèle qui est celui de l'attention auditive. En effet, le mécanisme de réception auditive qui est beaucoup sollicité dans cette étude, implique également l'attention (Garié, 2021). Or, cette attention est souvent déficitaire chez les patients avec un TSA. Alors, il serait intéressant de pouvoir évaluer l'effet de cette intervention sur l'attention auditive.

Pour conclure, ce travail a permis d'apporter de nouvelles pistes de remédiation novatrices pour pallier le déficit de la prosodie émotionnelle en réception, chez les enfants et adolescents avec un TSA. Ces patients sont reçus en orthophonie dans le cadre de troubles de la communication qui pénalisent souvent leur socialisation et leur intégration sociale, et pour lesquels l'axe prosodique ne semble pas travaillé au vu du manque d'informations et de moyens de remédiations dans la littérature. Or, le décodage de la prosodie émotionnelle est essentiel pour traiter un message oral. De fait, cet axe de travail semble intéressant à mettre en œuvre dans le cadre de la prise en soin de cette patientèle. De plus, il pourrait être un support de remédiation pour travailler l'attention auditive.

## **V Conclusion**

Dans ce mémoire, nous avons proposé un protocole de remédiation visant à améliorer l'interprétation de la prosodie émotionnelle chez des enfants et adolescents avec un TSA. Nous avons également créé un moyen d'évaluer cette compétence en amont, pendant, et après l'intervention. Enfin nous avons testé le protocole sur deux participants.

Les résultats de l'évaluation sont encourageants, et montrent une amélioration du taux de réussite à la tâche de désignation, bien que nous ne puissions pas évaluer l'effet significatif du protocole. Pour la conception de ce dernier, nous avons tenu à répondre aux critères de l'EBP (spécifique, mesurable, atteignable, réaliste et temporellement défini). En effet, l'entraînement était porté spécifiquement sur les paramètres : fréquence, variations de fréquence, rythme, et qualité vocale. Les compétences du participant étaient mesurées à plusieurs reprises. L'objectif de remédiation s'est révélé atteignable au vu des résultats pour un participant. Pour le deuxième participant, une période d'intervention plus large aurait éventuellement permis d'observer une tendance d'amélioration plus nette. La mise en œuvre du protocole est réaliste pour des participants ayant un niveau de compréhension suffisant. Enfin, l'entraînement était défini sur une période de 5 semaines à raison d'une fois par semaine.

Le protocole a été créé pour répondre au trouble de la communication sociale rencontré dans les manifestations cliniques des TSA (CIM-11, 2023 ; Crocq & Guelfi, 2015). Ce trouble de la communication touche à la fois la communication verbale, non-verbale et para-verbale (la prosodie), au sein duquel le décodage des émotions apparaît déficitaire (McCann & Peppé, 2003; Nader-Grosbois, 2020 ; Nuske et al., 2023). Les impacts des troubles dans la socialisation deviennent alors majeurs (Charpentier et al., 2018 ; Robertson & Baron-Cohen, 2017). Ainsi, la remédiation de la prosodie émotionnelle en réception, chez l'enfant et l'adolescent avec un TSA, a du sens dans la pratique orthophonique qui œuvre pour améliorer et/ou soutenir la communication. Néanmoins, les outils étant restreints, cet axe de travail ne semble pas évident à mettre en œuvre. Les perspectives de ce mémoire tendent vers la mise à disposition d'un outil d'évaluation rapide, standardisé et contrôlé afin de limiter les effets test et re-test. De plus, le protocole pourrait être jugé et amélioré par un groupe d'experts. Enfin, pour poursuivre la remédiation, un travail sur la compréhension de l'implicite serait intéressant, et ferait intervenir des compétences pragmatiques et langagières plus élaborées.

## Références

- Aguert, M., Bernicot, J., & Laval, V. (2009). Prosodie et compréhension des énoncés chez les enfants de 5 à 9 ans. *Enfance*, 3(3), 341-353. <https://doi.org/10.3917/enf1.093.0341>
- Aguert, M., Laval, V., & Bernicot, J. (2010). Comprendre l'intention communicative du locuteur : Une étude du rôle de l'intonation et du contexte chez des enfants de 5 à 9 ans. *L'Année psychologique*, 110(1), 49-70. <https://doi.org/10.3917/anpsy.101.0049>
- Ambert-Dahan, E. (2020). E.8. Communication et prosodie émotionnelle. In *Surdités de l'enfant et de l'adulte* (p. 308-314). De Boeck Supérieur. [https://www.cairn.sciences.info/article.php?ID\\_ARTICLE=DBU\\_LEYBA\\_2020\\_01\\_0308](https://www.cairn.sciences.info/article.php?ID_ARTICLE=DBU_LEYBA_2020_01_0308)
- Bel, B. (2008). *Subjectivité et émotion dans la prosodie de parole et du chant : Espace, coordonnées et paramètres*. Colloque international émotion, interaction et développement, Grenoble. [https://www.academia.edu/7363139/Subjectivité\\_et\\_émotion\\_dans\\_la\\_prosodie\\_de\\_parole\\_et\\_du\\_chant\\_espace\\_coordonnées\\_et\\_paramètres](https://www.academia.edu/7363139/Subjectivité_et_émotion_dans_la_prosodie_de_parole_et_du_chant_espace_coordonnées_et_paramètres)
- Benedetti, V., Weill-Chounlamountry, A., Pradat-Diehl, P., & Villain, M. (2022). Assessment tools and rehabilitation treatments for aprosodia following acquired brain injury : A scoping review. *International Journal of Language & Communication Disorders*, 57(3), 474-496. <https://doi.org/10.1111/1460-6984.12692>
- Bharathi, G., Jayaramayya, K., Balasubramanian, V., Vellingiri, B., Bharathi, G., Jayaramayya, K., Balasubramanian, V., & Vellingiri, B. (2019). The potential role of rhythmic entrainment and music therapy intervention for individuals with autism spectrum disorders. *Journal of Exercise Rehabilitation*, 15(2), 180-186. <https://doi.org/10.12965/jer.1836578.289>
- 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(13), 1-10. <https://doi.org/10.1186/s11689-017-9194-9>
- Boersma, P., & Weenink, D. (2024). *Praat : Doing phonetics by computer* (Version 6.4.12) [Logiciel]. <http://www.praat.org/>
- Caelen-Haumont, G. (2005). Les états émotionnels et la Prosodie : Paradigmes, modèles, paramètres. In *Phonologie et phonétique : Forme et substance* (Nguyen Noël, p. 397-424). Hermès. <https://hal.science/hal-00142950>
- Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech-Sound-Selective Auditory Impairment in Children with Autism : They Can Perceive but Do Not Attend. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9), 5567-5572.
- 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), Article 1. <https://doi.org/10.1186/s11689-018-9246-9>
- CIM-11. (2023). *Troubles du spectre de l'autisme* [Page web]. Classification internationale des maladies onzième édition. <https://icd.who.int/browse11/l-m/fr#/http://id.who.int/icd/entity/437815624>
- Crocq, M.-A., & Guelfi, J.-D. (2015). *DSM-5 : Manuel diagnostique et statistique des troubles mentaux* (5e éd). Elsevier Masson.
- de Gaulmyn, A., Miljkovitch, R., & Montreuil, M. (2018). Étude clinique des processus sous-jacents de l'attention conjointe de très jeunes enfants avec trouble du spectre autistique. *L'Encéphale*, 44(3), 224-231. <https://doi.org/10.1016/j.encep.2016.12.002>
- Devillers, L., & Vasilescu, I. (2004). Détection des émotions à partir d'indices lexicaux, dialogiques et prosodiques dans le dialogue oral. *Actes des XXVe Journées d'Études sur la Parole*. <https://www.afcp->

parole.org/doc/Archives\_JEP/2004\_XXVe\_JEP\_Fes/actes/jep2004/Devilleers-Vasilescu.pdf

- Di Cristo, A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 23, 67-211.
- Disney FR. (2015, juin 15). *Vice-Versa—Extrait : Pizza* [Page web]. YouTube. <https://www.youtube.com/watch?v=VBFQhx6aRoc>
- Duval, C., Piolino, P., Bejanin, A., Laisney, M., Eustache, F., & Desgranges, B. (2011). La théorie de l'esprit : Aspects conceptuels, évaluation et effets de l'âge. *Revue de neuropsychologie*, 3(1), 41-51. <https://doi.org/10.1684/nrp.2011.0168>
- Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men : Development and validation. *Behavior Research Methods*, 42(1), 351-362. <https://doi.org/10.3758/BRM.42.1.351>
- Fortin, A. (2021). *OCIF : Outil clinique pour les orthophonistes pour documenter les impacts fonctionnels d'un trouble de langage oral ou écrit*. Editions Horizons.
- Frith, U., Gerschenfeld, A., & Roques, S. (2010). *L'énigme de l'autisme* (2e éd). O. Jacob.
- Garié, L.-A. (2021). *Pratique orthophonique avec les enfants et adolescents présentant un TSA*. De Boeck supérieur.
- Gepner, B. (2012). Vers une théorie clinique intégrée des désordres de la constellation autistique. *Développements*, 10(1), 5-36. <https://doi.org/10.3917/devel.010.0005>
- 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), Article 8. <https://doi.org/10.1038/nn1291>
- Giacchino, M. (2015). *Vice-Versa—Team Building*. Walt Disney Records.
- Gomot, M., & Wicker, B. (2012). A challenging, unpredictable world for people with Autism Spectrum Disorder. *International Journal of Psychophysiology*, 83(2), 240-247. <https://doi.org/10.1016/j.ijpsycho.2011.09.017>

- Google. (2016). *Chrome Music Lab* [Site web interactif]. Chrome Music Lab. <https://musiclab.chromeexperiments.com>
- HAS. (2013). *Patient et professionnels de santé : Décider ensemble* [Page web]. Haute Autorité de Santé. [https://www.has-sante.fr/jcms/c\\_1671523/fr/patient-et-professionnels-de-sante-decider-ensemble](https://www.has-sante.fr/jcms/c_1671523/fr/patient-et-professionnels-de-sante-decider-ensemble)
- HAS. (2023). *Trouble du spectre de l'autisme (TSA) : Interventions et parcours de vie de l'enfant et de l'adolescent - Note de cadrage*. [https://www.has-sante.fr/jcms/p\\_3448980/fr/trouble-du-spectre-de-l-autisme-tsa-interventions-et-parcours-de-vie-de-l-enfant-et-de-l-adolescent-note-de-cadrage](https://www.has-sante.fr/jcms/p_3448980/fr/trouble-du-spectre-de-l-autisme-tsa-interventions-et-parcours-de-vie-de-l-enfant-et-de-l-adolescent-note-de-cadrage)
- IFIC. (2007). *Rubrique instruments* [Page web interactive]. Institut Francilien d'Implantation Cochléaire. <https://www.implant-ific.org/s/instruments>
- Jiam, N., Caldwell, M., Deroche, M., Chatterjee, M., & Limb, C. (2017). Voice Emotion Perception and Production in Cochlear Implant Users. *Hearing research*, 352, 30-39. <https://doi.org/10.1016/j.heares.2017.01.006>
- Jones, C. R. G., Pickles, A., Falcaro, M., Marsden, A. J. S., Happé, F., Scott, S. K., Sauter, D., Tregay, J., Phillips, R. J., Baird, G., Simonoff, E., & Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders : Emotion recognition in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 52(3), 275-285. <https://doi.org/10.1111/j.1469-7610.2010.02328.x>
- Kalathottukaren, R. T., Purdy, R., McCormick, S. C., & Ballard, E. (2015). Behavioral Measures to Evaluate Prosodic Skills : A Review of Assessment Tools for Children and Adults. *Contemporary Issues in Communication Science and Disorders*, 42, 138-154. [https://doi.org/10.1044/cicsd\\_42\\_S\\_138](https://doi.org/10.1044/cicsd_42_S_138)
- Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation : A practical guide. *Annals of Physical and Rehabilitation Medicine*, 61(3), 164-179. <https://doi.org/10.1016/j.rehab.2017.12.002>

- Kremer, J.-M., & Lederlé, E. (2016). Chapitre II. Les pratiques principales des orthophonistes. In Presses Universitaires de France (Éd.), *L'orthophonie en France* (8<sup>e</sup> éd., p. 39-85). puf. <https://www.cairn.info/l-orthophonie-en-france--9782130736622-p-39.htm>
- Kunda, M., & Goel, A. K. (2011). Thinking in Pictures as a cognitive account of autism. *Journal of Autism and Developmental Disorders*, 41(9), 1157-1177. <https://doi.org/10.1007/s10803-010-1137-1>
- Lacheret, A. (2011). Le corps en voix ou l'expression prosodique des émotions. *Evolutions Psychomotrices*, 23(90), 25-37.
- Lacroix, A. (2016). La prosodie dans tous ses états. *Bulletin de psychologie*, Numéro 542(2), 83-85. <https://doi.org/10.3917/bupsy.542.0083>
- Lainé, F., Rauzy, S., Gepner, B., & Tardif, C. (2009). Prise en compte des difficultés de traitement des informations visuelles et auditives rapides dans le cadre de l'évaluation diagnostique de l'autisme. *Enfance*, 1(1), 133-141. <https://doi.org/10.3917/enf1.091.0133>
- Lartseva, A., Dijkstra, T., & Buitelaar, J. K. (2015). Emotional language processing in autism spectrum disorders : A systematic review. *Frontiers in Human Neuroscience*, 8, 991. <https://doi.org/10.3389/fnhum.2014.00991>
- Laval, V., Le Sourn-Bissaoui, S., Girard, P., Chevreuil, C., & Aguert, M. (2012). Prosodie émotionnelle et compréhension des actes de langage expressifs chez des enfants et adolescents avec un Trouble du Spectre Autistique. *Revue française de linguistique appliquée*, XVII(2), 77-88. <https://doi.org/10.3917/rfla.172.0077>
- LeRobert. (2021). Sonie. In *Le Robert*.
- Leung, J. H., Purdy, S. C., Tippett, L. J., & Leão, S. H. S. (2017). Affective speech prosody perception and production in stroke patients with left-hemispheric damage and healthy controls. *Brain and Language*, 166, 19-28. <https://doi.org/10.1016/j.bandl.2016.12.001>
- Liebenthal, E., Silbersweig, D. A., & Stern, E. (2016). The Language, Tone and Prosody of Emotions : Neural Substrates and Dynamics of Spoken-Word Emotion Perception. *Frontiers in Neuroscience*, 10, 506. <https://doi.org/10.3389/fnins.2016.00506>

- Martel, K. (2009). Introduction – L'étude de la prosodie en acquisition : Un enjeu clinique. *Enfance*, 3(3), 259-272. <https://doi.org/10.3917/enf1.093.0259>
- 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>
- McCann, J., Peppé, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders*, 42(6), 682-702. <https://doi.org/10.1080/13682820601170102>
- Meiss, E., Tardif, C., Arciszewski, T., Dauvier, B., & Gepner, B. (2015). Effets positifs d'une exposition à des séquences vidéo ralenties sur l'attention, la communication sociale et les troubles du comportement chez 4 enfants autistes sévères : Une étude translationnelle pilote. *Neuropsychiatrie de l'Enfance et de l'Adolescence*, 63(5), 302-309. <https://doi.org/10.1016/j.neurenf.2015.01.004>
- Nader-Grosbois, N. (2020). Chapitre 8. Troubles du spectre de l'autisme. In *Psychologie du handicap* (2<sup>e</sup> éd., p. 295-364). De Boeck Supérieur. <https://doi.org/10.3917/dbu.nader.2020.01.0295>
- Nadig, A., & Shaw, H. (2012). Acoustic and Perceptual Measurement of Expressive Prosody in High-Functioning Autism : Increased Pitch Range and What it Means to Listeners. *Journal of Autism and Developmental Disorders*, 42(4), 499-511. <https://doi.org/10.1007/s10803-011-1264-3>
- Nuske, H. J., Vivanti, G., & Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition and Emotion*, 27(6), 1042-1061. <https://doi.org/10.1080/02699931.2012.762900>
- Nuske, H. J., Young, A. V., Khan, F., Palermo, E. H., Ajanaku, B., Pellecchia, M., Vivanti, G., Mazefsky, C. A., Brookman-Frazee, L., McPartland, J. C., Goodwin, M. S., & Mandell, D. S. (2023). Systematic Review : Emotion Dysregulation and Challenging Behavior Interventions for Children and Adolescents with Autism with Graded Key Evidence-

- Based Strategy Recommendations. *Research Square*. <https://doi.org/10.21203/rs.3.rs-2802378/v1>
- Parmentier, F., Marié-Bailly, I., & Pillot-Loiseau, C. (2014). A l'écoute de sa voix : Apports du spectrogramme comme rétrocontrôle visuel en rééducation vocale. *Glossa*, 116, 18-32.
- Patel, S. P., Nayar, K., Martin, G. E., Franich, K., Crawford, S., Diehl, J. J., & Losh, M. (2020). An Acoustic Characterization of Prosodic Differences in Autism Spectrum Disorder and First-Degree Relatives. *Journal of Autism and Developmental Disorders*, 50(8), 3032-3045. <https://doi.org/10.1007/s10803-020-04392-9>
- 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. (2015). *PEPS-C 2015* [Page web]. PEPS-C. <http://www.PEPS-C.com/peps-c-2015.html>
- Perret, M., Simon, I., De Battista, E., & Launay, L. (2012). Et si l'on rééduquait surtout la voie lexicale ? In A. Devevey, *Dyslexies : Approches thérapeutiques, de la psychologie cognitive à la linguistique* (p. 125-156). Solal éditeur. <https://www.cairn.info/dyslexies-2012--9782353270705-page-125.htm?ref=doi>
- Perret, M.-C., & Klink, L. (2007). Perception des paramètres acoustiques de la voix et compréhension verbale chez trois enfants atteints d'autisme. *Glossa*, 102, 42-50.
- Perrone-Bertolotti, M., & Gillet Perret, E. (2020). Chapitre 3. Le langage oral. In *Neurosciences cognitives développementales* (p. 80-127). De Boeck Supérieur. <https://doi.org/10.3917/dbu.poire.2020.01.0080>
- Phélip, M. (2014). Le paradigme d'écoute dichotique : État des recherches et données développementales. *Revue de neuropsychologie*, 6(1), 59-68. <https://doi.org/10.3917/rne.061.0059>
- Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., Atkinson, A. P., Calder, A. J., Johnstone, E. C., Lawrie, S. M., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum

- disorders. *Psychological Medicine*, 40(11), 1919-1929.  
<https://doi.org/10.1017/S0033291709992364>
- Plaza, M. (2014). Le développement du langage oral. *Contraste*, 39(1), 99-118.  
<https://doi.org/10.3917/cont.039.0099>
- Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience*, 18(11), 671-684. <https://doi.org/10.1038/nrn.2017.112>
- Rogers, S. J., & Dawson, G. (2020). 1-Connaissance actuelles sur l'apprentissage du jeune enfant et sur l'autisme. In *L'intervention précoce en autisme Le Modèle de Denver pour jeunes enfants*. Dunod.
- Rossi, M. (1997, septembre 22). Is syntactic structure prosodically recoverable? *5th European Conference on Speech Communication and Technology (Eurospeech 1997)*. 5th European Conference on Speech Communication and Technology (Eurospeech 1997).  
<https://doi.org/10.21437/Eurospeech.1997-1>
- Ruiz, R. (2012). *Détection de perturbations psychophysiologiques. Application au contexte aéronautique*. [Mémoire d'habilitation à diriger des recherches]. Laboratoire de Recherche en Audiovisuel, Toulouse.
- Société canadienne de pédiatrie. (2019). *La prise en charge et le suivi du trouble du spectre de l'autisme une fois le diagnostic posé* [Page web]. Société canadienne de pédiatrie.  
<https://cps.ca/fr/documents/position/tsa-la-prise-en-charge-et-le-suivi>
- Stewart, M. E., McAdam, C., Ota, M., Peppé, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism*, 17(1), 6-14.  
<https://doi.org/10.1177/1362361311424572>
- Tan, Y.-H., Xi, C.-Y., Jiang, S.-P., Shi, B.-X., Wang, L.-B., & Wang, L. (2012). Auditory abnormalities in children with autism. *Open Journal of Psychiatry*, 02(01), 33-37.  
<https://doi.org/10.4236/ojpsych.2012.21005>
- Tardif, C. (2010). *Autisme et pratiques d'intervention*. Solal.
- Tardif, C. (2014). Développements et variabilités : Les désordres du spectre de l'autisme. In S. Le Sourn-Bissaoui, G. Le Maner-Idrissi, V. Dardier, E. Bonjour, & A. Lacroix (Éds.),

*Développement et variabilités* (p. 121-140). Presses universitaires de Rennes.

<https://doi.org/10.4000/books.pur.61411>

Vannetzel, L., Chaby, L., Cautru, F., Cohen, D., & Plaza, M. (2011). Neutral versus emotional human stimuli processing in children with pervasive developmental disorders not otherwise specified. *Research in Autism Spectrum Disorders*, 5(2), 775-783.

<https://doi.org/10.1016/j.rasd.2010.09.005>

Vermeulen, P. (2020). *Autisme et émotions* (W. De Montis, Trad.; 3<sup>e</sup> éd.). De Boeck Supérieur.

Xavier, J., Vignaud, V., Ruggiero, R., Bodeau, N., Cohen, D., & Chaby, L. (2015). A Multidimensional Approach to the Study of Emotion Recognition in Autism Spectrum

Disorders. *Frontiers in Psychology*, 6.

<https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01954>

## **Annexes**

Annexe A : protocole d'intervention .....	1
Annexe B : support visuel de la séance sur l'articulation et la qualité de la voix.....	5
Annexe C : phrases enregistrées pour la tâche de désignation .....	6
Annexe D : patron prosodique du corpus enregistré.....	6

## Annexe A : protocole d'intervention

Tableau 1 : Protocole de remédiation du trouble prosodique en réception en fonction de l'objectif de travail, des outils et de la progression

Objectif	Outils	Progression		
		Discrimination	Identification	Appariement
Améliorer l'identification de sons aigus et grave	Xylophone	Nous produisons 2 notes différentes ou identiques et le participant doit m'indiquer si ce qu'il entend est pareil ou pas pareil, avec le support sous les yeux. Ici la modalité visuelle est une aide. Puis, le xylophone est caché, seule la modalité auditive est sollicitée.	Nous proposons 2 notes, quelle est la note aiguë : la première ou la deuxième, quelle est la note grave. Le xylophone est visible par le participant pour avoir un appui de la modalité visuelle.	Nous demandons au participant de produire la note qui correspond à la joie, à la tristesse et à la colère.
	Site internet ific : « Hauteur de note » (identification). « Taille des instruments » (identification et appariement).	Le participant doit relever la position de la note différente parmi 3 autres, dans une série.	Nous demandons au participant de m'indiquer si la note différente est plus aiguë ou plus grave que les autres.  Le participant écoute des instruments de musique. Nous lui demandons si cet instrument est plutôt aigu ou plutôt grave.	Nous demandons au participant d'associer un instrument entendu avec une émotion.
Améliorer l'identification	Xylophone	Nous produisons 2 gammes différentes ou identiques et le	Nous demandons au participant d'identifier la différence entre 2	Nous demandons au participant de produire une

des variations de hauteurs ou de l'étendue fréquentielle		participant doit m'indiquer si ce qu'il entend est pareil ou pas pareil, avec le support sous les yeux.	gammes avec ses mots. Nous le guidons en lui demandant ce qu'il observe comme différence sur le xylophone.	mélodie joyeuse, triste et en colère.
	Rouleau de piano mécanique (ChromMusicLab) (Google, 2016)		Le participant écoute 3 extraits de mélodies choisis et il doit me décrire ce qu'il entend et ce qu'il voit sur l'écran avec ses mots.	Nous demandons au participant d'associer une émotion à la mélodie entendue.
	Spectrogramme (ChromMusicLab) (Google, 2016)	A partir de sons préenregistrés, le participant observe ce qu'il se passe à l'écran. Il observe la différence avec et sans son. En présence d'un son le spectrogramme dessine une trace.	Nous demandons au participant de décrire les traces sur le spectre : les couleurs, la hauteur de la trace, les variations (« les vagues » et « les plats »).	Nous demandons au participant de faire la voix d'un personnage en colère, triste, et joyeux, ou de me donner des répliques que nous produis avec les 3 émotions. Puis, nous lui demandons d'associer l'image spectrale à l'émotion.
Améliorer l'identification du rythme	Instruments de musique (maracas et xylophone)	Nous produisons 2 rythmes différents (lent/rapide) ou identiques (lent/lent ou rapide/rapide) et le participant doit m'indiquer si ce qu'il entend est	Nous demandons au participant d'identifier la différence entre les 2 rythmes. Nous le guide en lui demandant ce qu'il observe comme geste et mouvement de la main.	Nous demandons au participant de produire un rythme en fonction des émotions. « ça ressemble à quoi un rythme joyeux ? triste ? en colère ? »

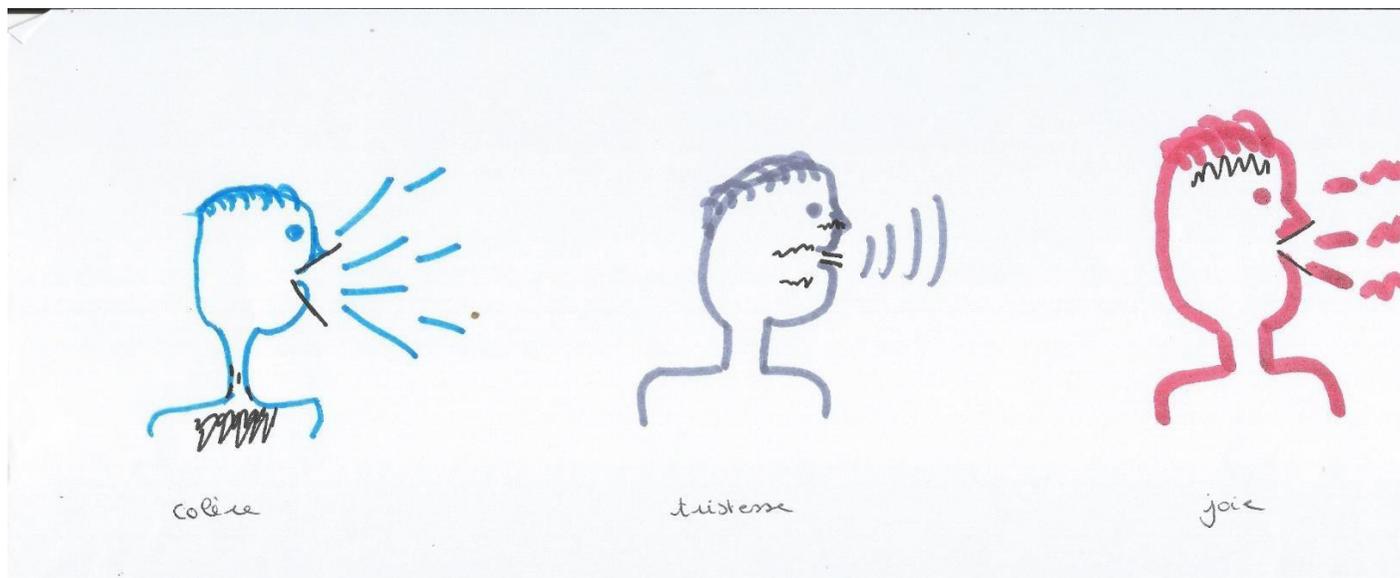
		pareil ou pas pareil, avec le support sous les yeux. Nous lui demandons ensuite de fermer les yeux.		
Rouleau de piano mécanique (ChromMusicLab)			Le participant réécoute les 3 extraits mélodiques et il doit me décrire le rythme, en s'aidant du support visuel. Plus les traits représentant les notes sont courts plus le rythme est rapide, et vice-versa.	Rappel : dans la séance précédente, nous avons associé une mélodie à une émotion. Ici, nous demandons au participant d'identifier le type de rythme pour chaque émotion. Puis, on classe les émotions par ordre croissant de vitesse.
Voix				Nous demandons au participant de mimer ou expliquer le débit de parole d'une personne en colère, joyeuse et triste.
Extrait de la bande originale de Vice-versa : « Team			Nous demandons au participant d'identifier le rythme (lent et rapide) de chaque thème musical en rapport avec une émotion/	Nous demandons au participant de reconnaître les émotions cachées parmi les thèmes musicaux et d'expliquer pourquoi.

	Building » (Giacchino, 2015)			
Améliorer l'identification de l'articulation et qualité de la voix	Voix et dessins	Nous produisons à la suite, un énoncé suivant deux fois, avec la même intonation, ou avec deux intonations différentes, correspondant aux émotions cibles : "bonjour", "ça va", "à dix heure". Nous demandons au participant de m'indiquer si ce qu'il entend est pareil ou différent.	Nous prononçons un énoncé avec une voix de poitrine (retentissante), une voix de tête (explosive) et une voix peu retentissante (résonnante). Nous demandons au patient de me décrire ce que nous fais, la couleur qu'on pourrait donner à cette voix, où se situe la vibration de la voix dans le corps, et comment se projette la voix. Nous dessinons des visages de profil avec la couleur choisie par le participant et on symbolise la zone de vibration (poitrine, tête et bouche), le type de projection de voix (explosive, sonore, peu sonore). On réitère l'expérience, nous demande cette fois au participant de décrire l'articulation et nous le guide par le lexique : <i>est-ce que c'est plutôt tendu, normale ou</i>	Nous demandons au participant de mettre en lien chaque profils avec une émotion. Nous demandons au participant de produire un énoncé de type « bonjour » ou « ça va » avec une émotion. Nous lui demandons de mettre sa main sur la zone de résonance.

			<i>relâché</i> ? On symbolise l'aperture vocale sur les profils des visages.	
	Extrait du dessin animé Vice-versa. (Disney FR, 2015)		A l'écoute des personnages, nous demande au participant d'identifier la zone de résonance. On observe ensuite leur articulation.	<i>Ralentir la vitesse de lecture de l'extrait, isoler les répliques des personnages ciblés.</i> Nous demandons au participant de reconnaître de quel personnage il s'agit : colère, joie ou tristesse.

### Annexe B : support visuel de la séance sur l'articulation et la qualité de la voix

Schéma 1: Illustrations des zones de résonance, de la projection vocale et de l'aperture

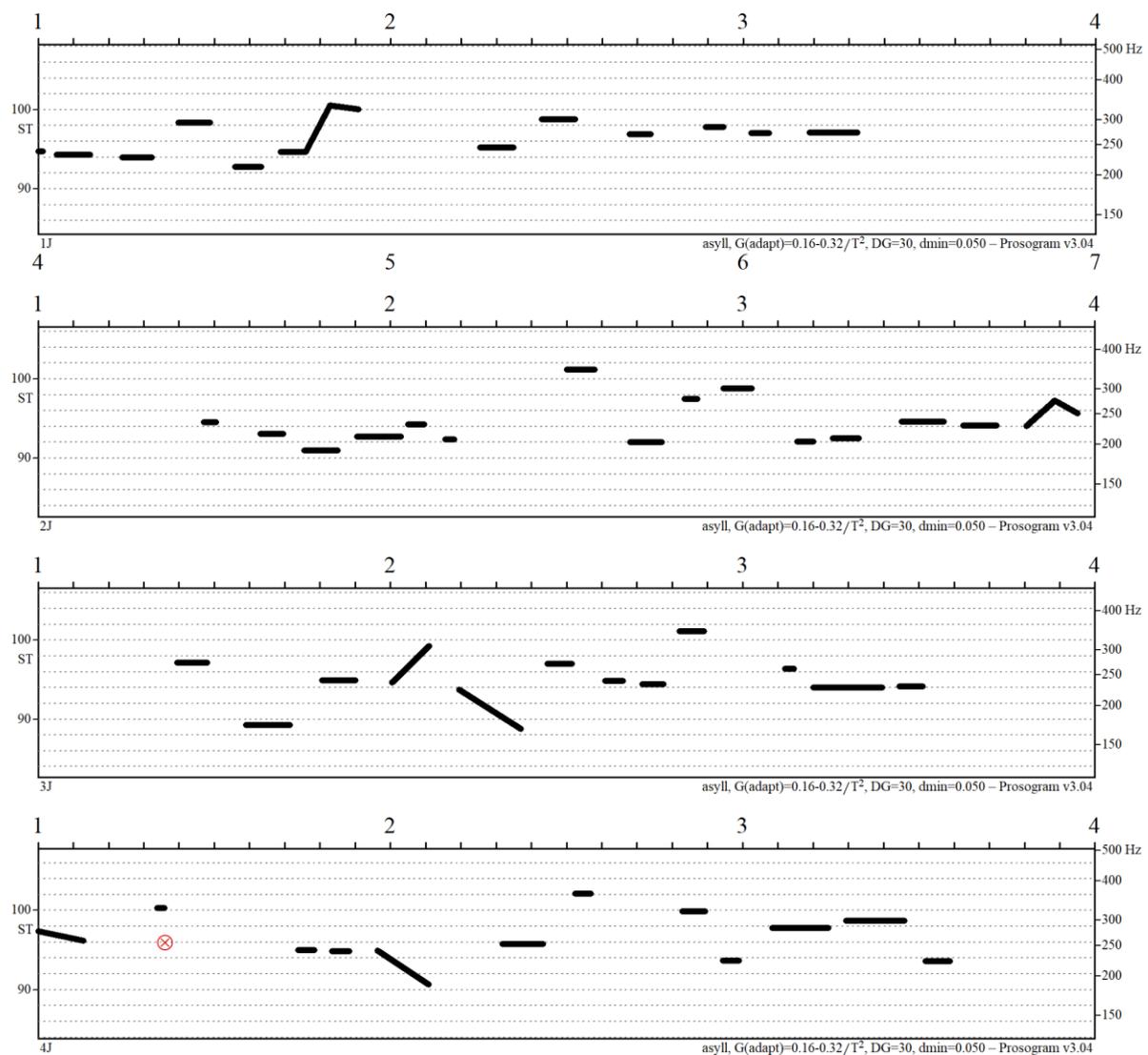


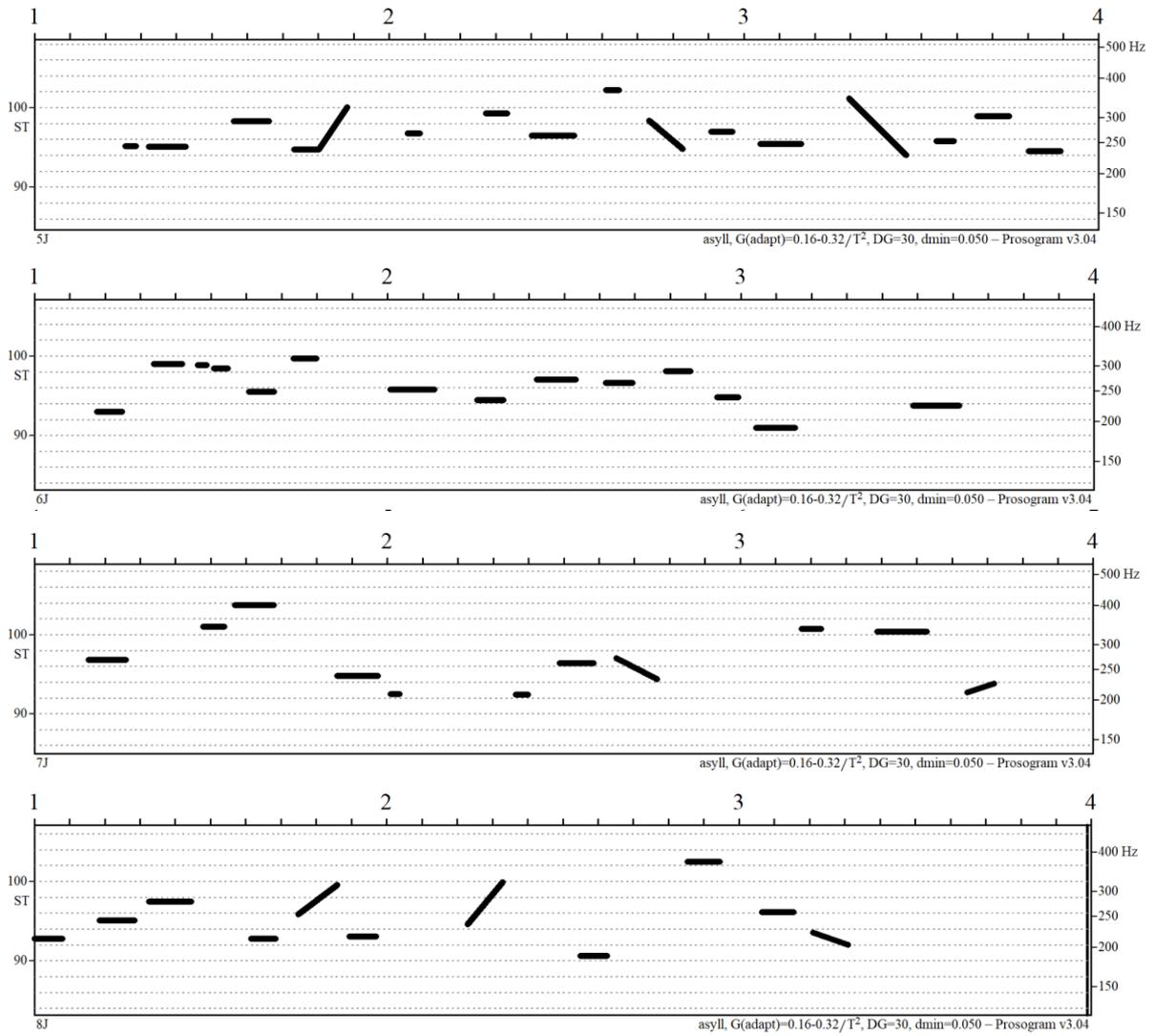
## Annexe C : phrases enregistrées pour la tâche de désignation

1. Les amis de Léo arrivent dans la journée
2. Je bavarde énormément avec mon voisin.
3. Je dois ranger mes jeux vidéo dans la malle.
4. Mardi, je regarde Léo jouer au handball.
5. Le voleur dévalise un magasin en ville.
6. Gabrielle mange des bonbons roses à la guimauve.
7. Mon meilleur ami a déjà vu un iguane.
8. La vendeuse me vend le billet à dix euros.

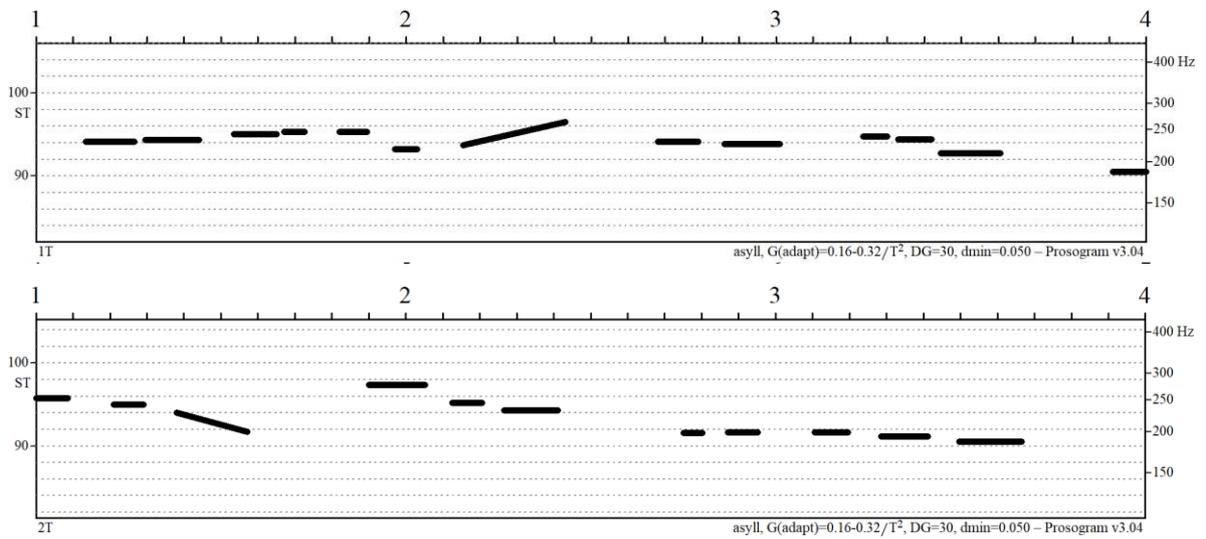
## Annexe D : patron prosodique du corpus enregistré

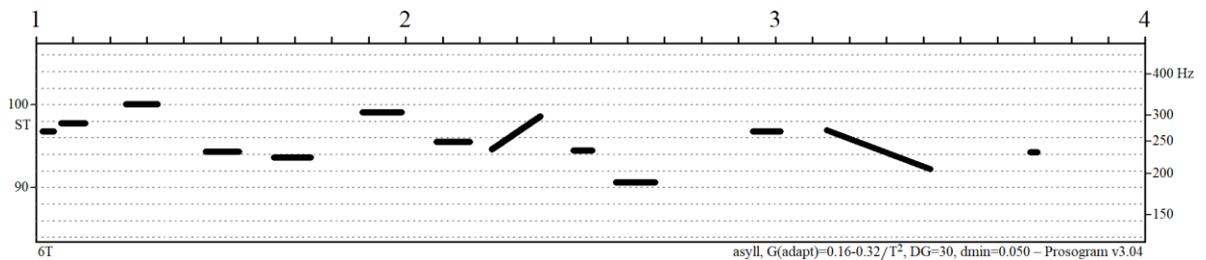
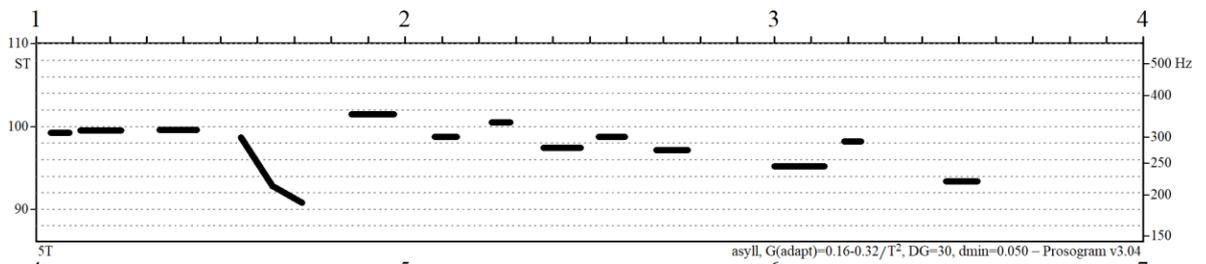
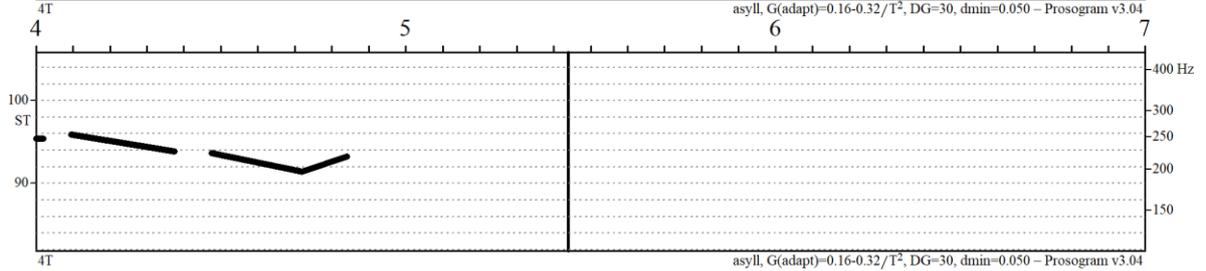
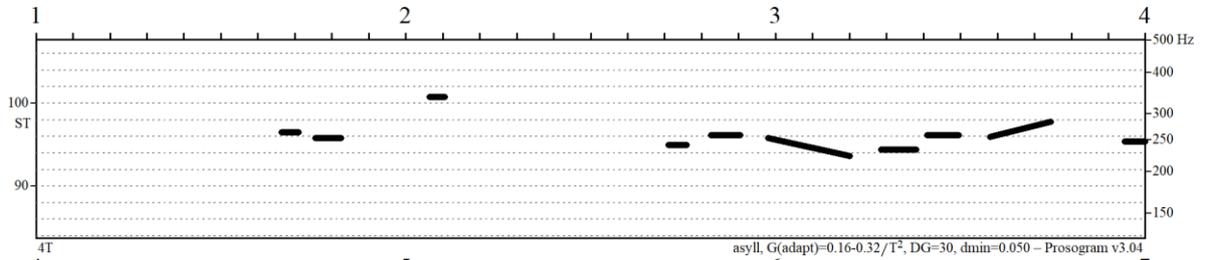
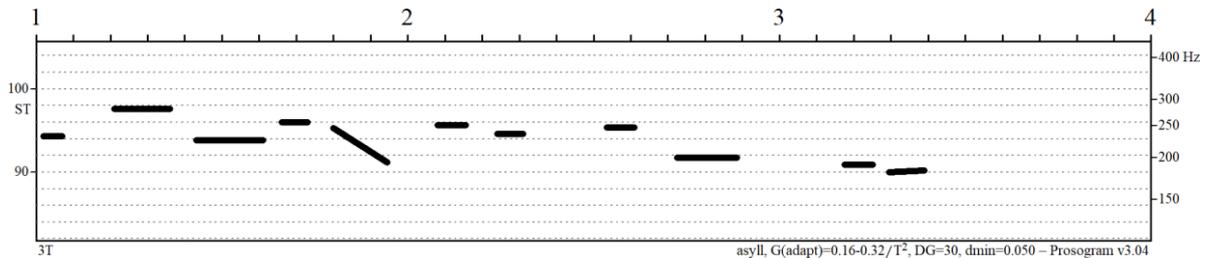
Patron prosodique des 8 énoncés pour la modalité « joie »

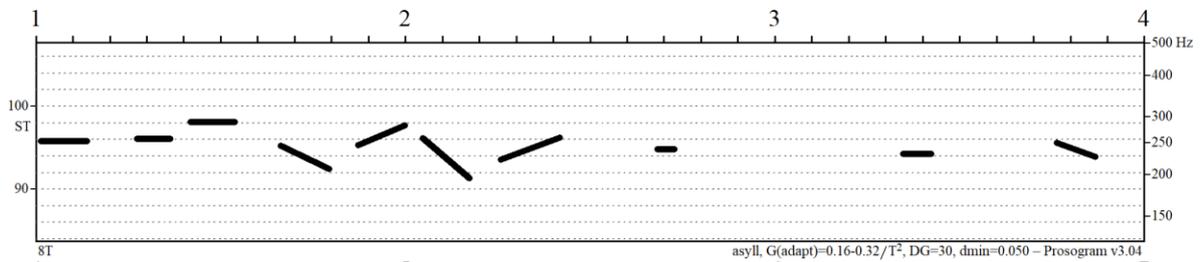
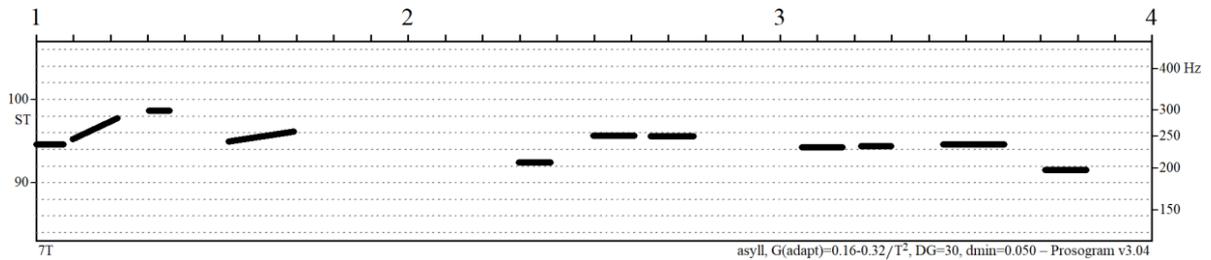




*Patron prosodique des 8 énoncés pour la modalité « tristesse »*







*Patron prosodique des 8 énoncés pour la modalité « colère »*

