



HAL
open science

Perception de la prosodie colérique exprimée par des acteurs professionnels et non-professionnels

Camille Onadu, Marion Sayah

► **To cite this version:**

Camille Onadu, Marion Sayah. Perception de la prosodie colérique exprimée par des acteurs professionnels et non-professionnels. Sciences du Vivant [q-bio]. 2021. dumas-03352852

HAL Id: dumas-03352852

<https://dumas.ccsd.cnrs.fr/dumas-03352852>

Submitted on 23 Sep 2021

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

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



Faculté des sciences
médicales et paramédicales
Aix-Marseille Université

Centre de Formation Universitaire
en Orthophonie de Marseille

Perception de la prosodie colérique exprimée par des acteurs professionnels et non-professionnels

Mémoire de fin d'études
En vue de l'obtention du Certificat de Capacité d'Orthophoniste

Camille ONADU et Marion SAYAH
Sous la direction de Caterina PETRONE et Francesca CARBONE

Soutenu le 17 juin 2021



UMR 7309 • CNRS • AMU
**PAROLE ET
LANGAGE**

REMERCIEMENTS

Nous tenons à remercier chaleureusement Caterina Petrone et Francesca Carbone, nos maîtres de mémoire, pour leur disponibilité, leur accompagnement et leurs conseils. Nous leur sommes reconnaissantes de nous avoir partagé leur passion pour la recherche expérimentale et de nous avoir fait confiance pour la réalisation de cette étude.

Merci à Joana Revis de nous avoir aiguillées dans les prémices de notre recherche, et pour ses précieux conseils concernant la soutenance de notre mémoire.

Nous sommes heureuses d'avoir la possibilité de présenter notre travail à Caterina Petrone, Francesca Carbone et Joana Revis. Nous vous adressons nos remerciements pour votre présence et pour l'honneur que vous nous faites en acceptant de faire partie de notre jury de soutenance.

Nos remerciements s'adressent également aux membres du Laboratoire Parole et Langage qui ont contribué à la mise en place des expériences.

Nous n'oublions pas de remercier chaleureusement chaque participant qui ont donné de leur temps pour cette expérimentation qui, sans eux, n'aurait pas pu se réaliser.

Nous exprimons notre reconnaissance à tous nos maîtres de stage qui nous ont permis de construire et d'enrichir notre expérience clinique.

Un grand merci à Madame Ferry : merci de nous avoir accompagnées durant ces années, merci d'avoir toujours écouté nos doutes et nos remises en question, merci d'avoir toujours su trouver les réponses justes avec humour et bienveillance.

Nous tenons également à remercier Florian (alias Floridiot) qui nous a toujours soutenues tout au long de ces cinq longues années d'études, et Lucas (alias Lukidiot) qui a été une épaule sur laquelle se reposer durant la réalisation de ce mémoire. Partager le quotidien avec vous nous rend plus fortes. On vous aime tendrement.

Du fond du cœur, nous attribuons une pensée particulière à nos parents, nos frères et sœurs, à Milky, à nos familles et nos amis pour leur soutien tout au long de ce mémoire, et plus largement tout au long de nos études. Merci D.L, E. et S. pour votre lumière et votre bienveillance.

Marion, mon binôme dans les études comme dans la vie. Je ne te remercierai jamais assez d'avoir rendu ces cinq années d'études aussi magnifiques. Notre complémentarité nous a amenées au bout de ce projet dont on ne peut qu'être fières. Je suis si heureuse qu'on ait réussi à préserver notre bonne entente même dans les moments difficiles. Merci d'avoir toujours été à l'écoute, d'avoir toujours cru en moi. Tu feras une excellente orthophoniste, j'en suis certaine. À notre amitié !

Camille, mon alliée, mon binôme, et avant tout mon amie. Je suis si reconnaissante de cette solide épaule, cette confidente et cette partenaire sans faille que tu as été pour moi pendant cinq ans. Merci d'avoir été plus que présente, que ce soit pour les études, mais pas seulement ; en passant par des moments joyeux, des rires idiots, mais aussi des instants plus solennels, et même difficiles. Notre complicité évidente a su nous donner les clés précieuses pour traverser ensemble cette sacrée aventure, dont ce projet est un bel aboutissement. On l'a fait ! Ta générosité, ton dévouement et ta spontanéité feront de toi une merveilleuse orthophoniste, je n'en ai jamais douté.

SOMMAIRE

REMERCIEMENTS	2
LISTE DES TABLEAUX, GRAPHIQUES ET FIGURES	6
INTRODUCTION	8
CHAPITRE 1 : INTRODUCTION THÉORIQUE ET CONTEXTE SCIENTIFIQUE	9
1. LES ÉMOTIONS	9
1.1. <i>Définitions</i>	9
1.2. <i>Les émotions de base</i>	10
1.3. <i>Dimensions des émotions : valence et arousal</i>	11
1.4. <i>Les canaux de transmission émotionnelle</i>	12
2. LA PROSODIE	13
2.1. <i>Définitions</i>	13
2.2. <i>Les paramètres acoustiques</i>	14
2.3. <i>Les fonctions de la prosodie</i>	16
2.4. <i>La prosodie émotionnelle</i>	17
3. LES ÉTUDES SUR LA PROSODIE ÉMOTIONNELLE	19
3.1. <i>Méthodologie des études sur l'expression de la prosodie émotionnelle</i>	19
3.1.1. <i>Les conditions d'encodage des émotions</i>	19
3.1.2. <i>L'entraînement vocal</i>	22
3.2. <i>Manipulation des indices de la prosodie émotionnelle</i>	24
3.2.1. <i>Mesure et régression des indices</i>	24
3.2.2. <i>Masquage des indices</i>	25
3.2.3. <i>Resynthèse des indices</i>	25
3.3. <i>Méthodologie des études sur la perception de la prosodie émotionnelle</i>	26
4. QUESTIONNEMENTS ET OBJECTIFS DE NOTRE ÉTUDE	27
CHAPITRE 2 : MÉTHODOLOGIE DE L'ÉTUDE	30
1. LA SÉLECTION DES STIMULI	30
1.1. <i>Les sujets</i>	30
1.1.1. <i>Acteurs professionnels</i>	30
1.1.2. <i>Acteurs non-professionnels</i>	31
1.2. <i>Les films</i>	32
1.3. <i>Les émotions</i>	34
1.3.1. <i>La colère</i>	34
1.3.2. <i>L'état émotionnel neutre (condition contrôle)</i>	34
2. LES ANALYSES ACOUSTIQUES	35
2.1. <i>Mesure de fréquence fondamentale (f0) et d'intensité vocale</i>	35
2.2. <i>Résultats des analyses</i>	36
2.2.1. <i>Pour la colère</i>	36
2.2.2. <i>Pour l'état émotionnel neutre</i>	37
2.2.3. <i>Interprétation des résultats</i>	38
3. L'INTERPRÉTATION ET LA MODIFICATION DES STIMULI	39
3.1. <i>Transcription des stimuli et vérification de la neutralité du contenu verbal</i>	39
3.2. <i>Délexicalisation</i>	40
3.2.1. <i>Méthode de filtrage passe-bas</i>	41
3.2.2. <i>Méthode d'analyse et de resynthèse</i>	41
CHAPITRE 3 : EXPÉRIENCES DE PERCEPTION ÉMOTIONNELLE	44
1. TÂCHE DE VALIDATION N°1 : INTELLIGIBILITÉ	44
1.1. <i>La création du test</i>	44

1.2.	<i>Les participants recrutés</i>	45
1.3.	<i>Résultats</i>	45
2.	TÂCHE DE VALIDATION N°2 : IDENTIFICATION	46
3.	EXPÉRIENCE PRINCIPALE : ÉVALUATION ÉMOTIONNELLE	47
3.1.	<i>Sélection des stimuli</i>	47
3.2.	<i>La création du test</i>	49
3.3.	<i>Les participants recrutés</i>	50
3.4.	<i>Résultats</i>	50
3.4.1.	Résultats comparatifs entre les deux groupes d'acteurs	51
3.4.2.	Résultats comparatifs entre les sexes des acteurs	54
3.4.3.	Résultats comparatifs entre tous les acteurs	58
4.	CONCLUSION	61
4.1.	<i>Interprétation des résultats</i>	61
4.2.	<i>Limites</i>	62
4.3.	<i>Ouvertures</i>	63
	BIBLIOGRAPHIE	65
	ANNEXES	70
	RÉSUMÉ	84

LISTE DES TABLEAUX, GRAPHIQUES ET FIGURES

LISTE DES TABLEAUX :

Tableau 1 : Comparaison des paramètres prosodiques des émotions.....	18
Tableau 2 : Acteurs professionnels sélectionnés pour l'étude.....	31
Tableau 3 : Acteurs non-professionnels sélectionnés pour l'étude.....	32
Tableau 4 : Présentation des films sélectionnés.....	33
Tableau 5 : Moyennes obtenues pour la colère et pour chaque paramètre acoustique.....	36
Tableau 6 : Moyennes obtenues pour l'état émotionnel neutre et pour chaque paramètre acoustique.....	37
Tableau 7 : Les 14 stimuli aux transcriptions discordantes.....	40
Tableau 8 : Les 20 stimuli sélectionnés pour l'évaluation émotionnelle.....	48
Tableaux 9 et 10 : Moyennes obtenues pour la valence de la colère et de l'état émotionnel neutre....	51
Tableaux 11 et 12 : Moyennes obtenues pour l'intensité de la colère et de l'état émotionnel neutre...	53

LISTE DES FIGURES :

Figure 1 : Représentation du modèle Circumplex de Russell.....	12
Figure 2 : Spectrogramme et sonogramme du stimulus original (np_caneele_c_1).....	43
Figure 3 : Spectrogramme et sonogramme du stimulus après reconstruction de la fréquence fondamentale (np_caneele_c_1_delex).....	43
Figure 4 : Spectrogramme et sonogramme du stimulus après reconstruction de l'intensité et des caractéristiques spectrales (np_caneele_c_1_intensity_morphed).....	43

LISTE DES GRAPHIQUES :

Graphiques 1 et 2 : Graphiques représentant les moyennes obtenues pour la valence de la colère et de l'état émotionnel neutre.....	52
Graphiques 3 et 4 : Graphiques représentant les moyennes obtenues pour l'intensité de la colère et de l'état émotionnel neutre.....	53
Graphiques 5 et 6 : Graphiques représentant la valence de la colère pour les hommes et pour les femmes.....	55

Graphiques 7 et 8 : Graphiques représentant la valence de l'état émotionnel neutre pour les hommes et pour les femmes.....	55
Graphiques 9 et 10 : Graphiques représentant l'intensité de la colère pour les hommes et pour les femmes.....	56
Graphiques 11 et 12 : Graphiques représentant l'intensité de l'état émotionnel neutre pour les hommes et pour les femmes.....	57
Graphique 13 : Graphique représentant les moyennes obtenues pour la valence de la colère et pour chaque acteur.....	58
Graphique 14 : Graphique représentant les moyennes obtenues pour la valence de l'état émotionnel neutre et pour chaque acteur.....	58
Graphique 15 : Graphique représentant les moyennes obtenues pour l'intensité de la colère et pour chaque acteur.....	59
Graphique 16 : Graphique représentant les moyennes obtenues pour l'intensité de l'état émotionnel neutre et pour chaque acteur.....	59

INTRODUCTION

Les émotions jouent un rôle essentiel dans notre façon de communiquer et sont omniprésentes dans nos interactions sociales. Elles s'expriment à travers nos gestes, nos expressions faciales, ou encore à travers notre voix. Le domaine de recherche sur les émotions s'attache encore à décrire les différentes modalités d'expression et de perception émotionnelles. Dans cette perspective, un groupe de chercheurs et d'étudiants attachés au Laboratoire Parole et Langage d'Aix-en-Provence ont mis au jour en 2019 un projet d'étude nommé COMEDIA (COMparaison des Émotions Des Inexpérimentés et des Acteurs). Mené par Caterina Petrone et Maud Champagne-Lavau, ce projet vise à étudier l'authenticité des émotions en analysant les expressions émotionnelles jouées par des acteurs. Pour cela, COMEDIA fait intervenir de nombreux chercheurs dans différents domaines (linguistes, experts en prosodie, étudiants en sciences cognitives, en cinéma, en orthophonie...) afin de mener à bien plusieurs travaux pluridisciplinaires. Ces derniers tentent notamment de comparer la prosodie et les expressions faciales d'acteurs professionnels et d'acteurs non-professionnels, ainsi que leurs effets sur l'évaluation émotionnelle.

C'est précisément dans le cadre de ce projet que notre travail s'inscrit. À travers notre étude, nous contribuons à entrer dans la recherche sur la prosodie émotionnelle, en tentant de répondre aux questionnements suivants : existe-t-il un lien entre l'expérience cinématographique d'un acteur et sa manière d'exprimer l'émotion de la colère ? Cette expérience professionnelle influence-t-elle aussi notre façon de percevoir la prosodie colérique véhiculée par le jeu d'acteur ?

Afin de répondre à ces problématiques, nous mettrons un point d'honneur à expliciter et à quantifier les différences prosodiques de la parole jouée entre deux groupes d'acteurs (professionnels et non-professionnels) à travers des extraits de films exprimant deux états émotionnels bien distincts : la colère (qui est une des six émotions de base) et l'état émotionnel neutre (qui sera notre condition contrôle).

Ainsi, nous consacrerons notre premier chapitre à définir les concepts d'émotion et de prosodie tout en abordant le contexte scientifique. Dans le second chapitre du mémoire, nous expliciterons la méthodologie de notre étude : la sélection des extraits de films, leurs analyses acoustiques et les modifications apportées. Enfin, dans le troisième et dernier chapitre, nous détaillerons les différentes expérimentations de notre étude.

CHAPITRE 1 : INTRODUCTION THÉORIQUE ET CONTEXTE SCIENTIFIQUE

1. LES ÉMOTIONS

1.1. Définitions

Dans le domaine de recherche des émotions (« émotion » du latin « emovere », dont e- signifie « hors de » et movere signifie « ébranler » ou « mettre en mouvement ») (Granarolo, 2021), l'un des problèmes majeurs rencontrés dans le domaine de recherche des émotions est la difficulté à trouver une définition consensuelle de l'émotion, ainsi que de définir l'ensemble de ses caractéristiques spécifiques (manifestations, structures cérébrales mises en jeu, moyens de transmission...). En 1884, James parle de l'émotion comme d'une réaction physiologique ; lorsqu'on est dans la forêt et qu'un ours apparaît, nos os tremblent à cause de l'ours et on éprouve de la peur parce qu'on sent nos os trembler (James, 1884). Fridja précise cette définition en considérant que les émotions sont relativement de courte durée, et qu'il s'agit d'états non intentionnels entraînant des changements dans le comportement moteur, des changements physiologiques, mais aussi des changements dans les processus cognitifs (Fridja, 1987). En 1994, Ekman insiste sur le caractère aigu et transitoire de l'émotion, ce qui lui confère un rôle hautement adaptatif. Par exemple, face à une injustice ou une humiliation, l'émotion de la colère peut mobiliser très rapidement un individu. Elle implique un mécanisme d'autodéfense, et fournit l'énergie nécessaire pour nous défendre face à ce qui a éveillé notre colère. Ce rôle attribue à l'individu la capacité de mobilisation quasi instantanée de toutes les ressources de son organisme pour faire face à une situation donnée (Ekman et al., 1994).

La difficulté à distinguer les émotions des autres types d'états affectifs rend d'autant plus complexe la définition précise de l'émotion. Ces autres états affectifs peuvent être les humeurs, les sentiments, les positions interpersonnelles, les attitudes ou encore les traits de personnalité affectifs. Préciser cette distinction fait l'objet d'un désaccord entre différents auteurs. Pour LeDoux, émotion et sentiment se confondent ; il explique que ces deux notions sont toutes deux une expérience subjective, un envahissement passionnel de la conscience (LeDoux, 1996). En revanche, Damasio insiste sur la distinction à établir entre ces deux mots (« émotion » et « sentiment »), trop souvent utilisés l'un pour l'autre ; il décrit le sentiment comme un état plus durable que l'émotion (Damasio, 1999).

Scherer va dans le sens de Damasio et précise ce qui différencie émotion des autres états affectifs. Il explique que la distinction est possible en considérant le fait que les émotions sont plus intenses, mais de plus courte durée et amenées à changer plus rapidement que tout autre état affectif. Elles sont aussi bien plus susceptibles d'être fortement axées sur les événements déclencheurs. En effet, elles sont principalement dirigées par les événements extérieurs, les situations ou l'environnement. Enfin, elles ont un plus fort impact sur le comportement : elles induisent systématiquement des réponses comportementales. Par exemple, la colère induit des réponses comportementales très variées : elle s'exprime à la fois par de l'impulsivité, de l'agressivité ou de l'agitation, mais aussi par un comportement de retrait ou de renfermement (Scherer, 2003).

C'est sur la base de toutes ces définitions que nous réalisons notre étude, tout en gardant en tête le contexte de dissensus.

1.2. Les émotions de base

Dans la perspective de catégoriser davantage les émotions, Charles Darwin, en 1872, fut le premier à émettre les hypothèses selon lesquelles il existerait des émotions qui pourraient être universelles (exprimées et reconnues par toutes les cultures et tous les pays) et adaptatives (favorisant la survie de l'espèce en permettant aux individus de répondre de façon appropriée aux exigences environnementales). Pour cela, il s'est intéressé à l'étude de la posture et des mimiques faciales chez les animaux et les humains. Il a alors repéré huit émotions qu'il qualifie d'universelles ou encore d'émotions « discrètes » : anxiété, colère, contrariété, dégoût, honte, joie, souffrance et surprise (Darwin, 1872). Ses travaux ont eu une influence considérable sur les études ultérieures des émotions : par la suite, plusieurs auteurs, et en particulier Ekman, ont défendu et repris cette idée d'un nombre d'émotions universelles. En 1971, ce dernier a conduit plusieurs études auprès de six cultures fondamentalement différentes (Argentine, Brésil, Chili, États-Unis, Japon et tribus Fore de Papouasie) (Ekman, 1971). Il a présenté à chaque participant de ces cultures des photos de visages très expressifs et a évalué leur reconnaissance. De ces investigations, il a conclu que six émotions étaient reconnues dans les cultures présentées précédemment : la colère, le dégoût, la joie, la peur, la surprise et la tristesse. Il a alors considéré ces six émotions comme universelles et les a nommées : « émotions de base » ou « émotions fondamentales ». De ce postulat, Ekman a précisé que pour être considérée comme une émotion de base, une manifestation affective doit notamment répondre aux critères spécifiques suivants : elle doit se déclencher rapidement et être brève,

se retrouver chez tous les primates, être universellement identifiable, être associée à une physiologie propre, et se déclencher de façon involontaire (Ekman, 1999). Par exemple, pour la colère, nous retrouvons un déclenchement rapide face à une agression physique ou morale. Généralement, chez l'individu en colère, nous pouvons retrouver une physiologie caractéristique se manifestant par exemple par une tachycardie, une dyspnée, une hyperthermie ou encore des tensions musculaires.

En dehors de ces émotions primaires, on retrouve les états émotionnels secondaires. Ce sont des réponses émotionnelles à une émotion primaire ; Plutchik en a répertorié jusqu'à 142 (Plutchik, 1980). Ceux-ci sont complexes dans la mesure où ils combinent souvent plusieurs caractéristiques des émotions de base mais sont des états émotionnels plus réfléchis et durables (Caelen-Haumont, 2005). Par exemple, pour la colère, on retrouve un spectre des émotions secondaires allant du mécontentement à la fureur.

C'est sur la base de cette théorie catégorielle des émotions de base que la plupart des études sur l'expression vocale des émotions se sont réalisées. Dans cette optique, nous avons décidé d'élaborer notre étude autour de la colère.

1.3. Dimensions des émotions : valence et arousal

On peut décrire une émotion selon ses dimensions : sa valence et son arousal, deux caractéristiques indépendantes. Cette approche dimensionnelle a été initiée par Wundt (Wundt, 1874). La valence d'une émotion correspond à son aspect négatif (désagréable) ou positif (agréable). Par exemple, les émotions comme la colère et la peur ont une valence négative, alors que la joie a une valence positive. Cette dimension s'étend comme un continuum allant de la douleur extrême ou du malheur, jusqu'au bonheur extrême voire à l'extase. L'arousal (ou « intensité émotionnelle ») d'une émotion correspond quant à lui à son niveau d'excitabilité, d'éveil ou encore d'activité : de peu stimulant voire soporifique à très stimulant ou même à l'excitation frénétique ; en passant par des états intermédiaires. Par exemple, la colère et la peur ont un niveau d'excitabilité très élevé, contrairement à la tristesse qui a un niveau d'excitabilité bien plus faible.

Russell a modélisé cette théorie bidimensionnelle grâce à un cercle et deux vecteurs représentant les valeurs de valence et d'arousal. Au centre de cet espace circulaire se situe le point de neutralité émotionnelle, il correspond à l'absence totale d'un état émotionnel, autrement dit un état émotionnel neutre ; à la périphérie se trouve le degré d'émotion le plus intense (Russel, 1980). (Figure 1)

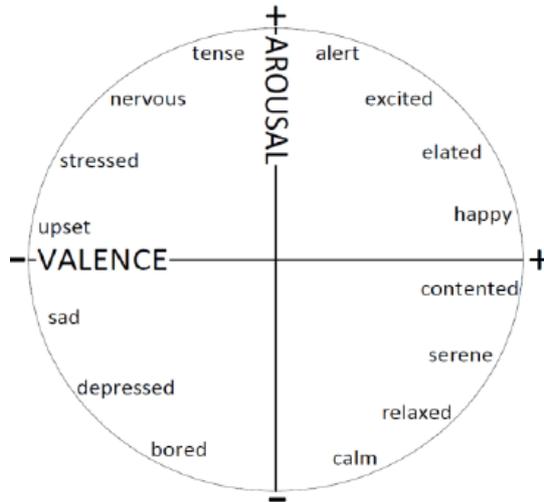


Figure 1 : Représentation du modèle Circumplex de Russell (Russel, 1980).

Certains auteurs ont proposé l'ajout d'autres dimensions. Par exemple, selon Russell et Mehrabian, il est nécessaire d'utiliser une troisième dimension représentant le pouvoir, le contrôle ou encore la dominance. Cette dimension s'étend de soumis (avec la sensation d'absence totale de contrôle sur les événements ou les situations), à puissant (avec la sensation d'être influent et maître de la situation). Ainsi, cette troisième dimension de dominance permet d'améliorer la « résolution émotionnelle ». En effet, la peur et la colère ont des valeurs de valence et d'arousal similaires. Cette troisième dimension permet tout de même de les différencier : la colère a besoin de pouvoir alors que la peur est une émotion de soumission (Russell et Mehrabian, 1977).

Dans ce mémoire, nous allons nous concentrer sur les deux dimensions émotionnelles : valence et intensité pour l'évaluation de la colère. En comparaison à un état émotionnel neutre (notre condition contrôle), la colère est en fait caractérisée par une valence négative et une intensité forte.

1.4. Les canaux de transmission émotionnelle

Le canal de transmission émotionnelle le plus étudié dans la littérature est le corps, probablement parce qu'il induit des manifestations qui sont plus visibles dans la vie quotidienne. En effet, les émotions entraînent un réflexe moteur qui se manifeste à travers nos gestes, nos postures ou notre visage par le biais des expressions faciales. En 1999, Ekman

qui a longtemps étudié ces dernières, est parvenu à saisir des micro-émotions par les mouvements des muscles qui parcourent le visage en moins d'une demi-seconde (Ekman, 1999). L'expression faciale de la colère se manifeste souvent par un abaissement du front, des sourcils intérieurs pointant vers le bas et laissant apparaître des rides entre les deux yeux, une mâchoire fermée ou un teint tendant vers le rouge. Selon Le Breton, les mouvements innombrables du corps lors de l'interaction (les gestes, les mimiques, les postures, déplacements, etc.) s'enracinent dans l'affectivité des individus (Le Breton, 1998).

Toutefois, le corps n'est pas la seule modalité d'expression émotionnelle. On retrouve également un vecteur de transmission tout aussi enclin à transmettre des indices émotionnels précieux : la voix. (Aubergé et Cathiard, 2003). À travers la voix, les émotions sont véhiculées par des variations acoustiques pouvant être perçues de manière auditive et qui ne requièrent pas un contact visuel direct pour être communiquées. Cette grande efficacité rend fondamentale la transmission vocale des émotions pour nos interactions sociales, tant elle affecte et structure nos comportements communicationnels (Hawk et al., 2009). Ainsi, dès lors qu'un processus émotionnel est enclenché, on observe des changements acoustiques spécifiques impliquant diverses modifications dans la voix, plus particulièrement dans la prosodie.

2. LA PROSODIE

2.1. Définitions

La prosodie est, de manière simplifiée, « la musique du langage ». Dans la littérature scientifique, le terme prosodie est notamment utilisé dans les études en linguistique, en phonétique et en phonologie. Bien que les scientifiques ne s'accordent pas sur sa définition exacte, on retrouve tout de même des similarités dans ses différentes descriptions. Elle concerne la structure temporelle, rythmique et mélodique du langage. Entre autres, la prosodie est un phénomène qui se caractérise par la modification des paramètres acoustiques de la voix (Lacheret et al., 1999). C'est un objet complexe de la parole qui regroupe divers phénomènes tels que l'accent, l'articulation, le ton, la mélodie, le rythme ou encore le débit (Audibert, 2008). Grâce à cela, un locuteur peut transmettre davantage d'informations à travers sa voix, notamment ses intentions, son point de vue, et même ses émotions. De plus, la prosodie est propre à chaque langue. Cela s'explique par les normes sociales et les habitudes culturelles qui influencent notre communication, mais aussi par le fait que la

prosodie représente une partie de la grammaire, et qu'elle varie donc selon la langue considérée.

Nous pouvons étudier la prosodie sous deux angles : un niveau phonétique et un niveau phonologique. En ce qui concerne le premier niveau, relié à l'aspect phonétique de la voix, il se traduit par des indices prosodiques repérables dans le signal sonore et est généralement mesuré sur des fenêtres temporelles très larges comme la phrase. Le deuxième niveau, relié à l'aspect phonologique de la voix, concerne la grammaire de la prosodie ; il s'intéresse à comment les différents paramètres acoustiques signalent des phénomènes phonologiques, tels que la structure prosodique d'une langue ou encore le positionnement de l'accent mélodique (Ladd, 2008).

Dans ce mémoire, nous allons nous concentrer sur les aspects phonétiques de la prosodie, puisque ce sont eux qui sont particulièrement liés à la prosodie émotionnelle. En particulier, nous allons détailler ci-dessous les trois indices les plus utilisés pour la description acoustique de la prosodie émotionnelle : la fréquence fondamentale, la durée et l'intensité (Banse et Scherer, 1996 ; Scherer, 2003). De plus, nous porterons aussi notre intérêt sur la qualité de la voix.

2.2. Les paramètres acoustiques

Fréquence fondamentale f_0

La fréquence fondamentale (f_0), exprimée en Hertz (Hz), correspond au nombre d'ouvertures et fermetures par seconde des cordes vocales. C'est elle qui illustre le caractère mélodique plus ou moins aigu des sons ; plus cette fréquence est haute, plus le son est aigu. En moyenne, la fréquence fondamentale est de 110 Hz pour une voix d'homme, et de 220 Hz pour une voix de femme. De cette fréquence fondamentale découlent des mesures fréquentielles utiles dans l'évaluation acoustique de la prosodie : la fréquence fondamentale moyenne (moyenne f_0), les écarts d'amplitude (f_0 range) et les variations de contours de f_0 (Caelen-Haumont, 2005).

Durée segmentale

La durée, exprimée en millisecondes (ms), correspond aux variations temporelles d'un signal sonore. Ces variations correspondent au débit, à l'allongement et au tempo. Le débit se réfère à la vitesse d'articulation mesurée au sein des macro-unités de segmentations (par exemple,

le nombre de syllabes par seconde et par énoncé). On peut mesurer le débit moyen d'une prise de parole en divisant la durée totale de l'énoncé par le nombre de syllabes. En général, le débit conversationnel, autrement dit la vitesse moyenne d'élocution lors d'une conversation, doit se situer aux alentours de 200 mots/minute si les locuteurs veulent être intelligibles. Ensuite, l'allongement correspond à l'augmentation significative de la durée d'une voyelle ou d'une syllabe. Cette augmentation peut illustrer une hésitation, un trait régional ou phonostylistique, ou encore au niveau phonologique un marquage de la fin d'une unité intonative. Enfin, le tempo correspond au nombre de syllabes proéminentes par seconde (Lacheret et al., 1999).

Intensité sonore

L'intensité vocale (appelée aussi énergie ou sonie), exprimée en Décibels (dB), correspond à la puissance mise dans la voix. C'est elle qui illustre le volume sonore plus ou moins faible ou fort ; plus cette intensité est haute, plus le son est fort. Elle dépend de l'amplitude du signal acoustique. Cette amplitude est créée par la variation de pression de l'air exercée sur les cordes vocales. La voix conversationnelle se situe autour de 65 dB, la voix forte (projetée) autour de 80 dB, et la voix chuchotée (à faible intensité) se situe autour de 30 dB (Caelen-Haumont, 2005).

Caractéristiques spectrales

Nous pouvons ajouter à ces trois derniers paramètres les caractéristiques spectrales qui illustrent la qualité vocale. Cette dernière correspond à la coloration auditive ou encore au timbre caractéristique de la voix d'un locuteur (D'Alessandro, 2006). En 1980, Laver affirme que les caractéristiques laryngées et supra laryngées sont considérées comme contribuant à la qualité de la voix (Laver, 1980). En effet, la qualité de vibration des plis vocaux, le positionnement de la langue ou des lèvres, le degré de pression sous-glottique ou les tensions musculaires exercées par les muscles phonatoires durant une prise de parole peuvent être des composantes physiologiques qui caractérisent la qualité vocale d'un individu (Gobl et Chasaide, 2003). Ainsi, pour décrire la qualité de la voix, nous pouvons utiliser les qualificatifs : soufflée, rauque, modale, voilée, murmurée, serrée, craquée, nasale, hyper- ou hypotonique... (Laver, 1980). À titre d'exemple, pour la colère, on peut retrouver une voix criée, serrée, ou encore hypertonique (Gobl et Chasaide, 2003). L'importance fondamentale de la qualité de la voix dans l'expression des émotions a notamment été confirmée par les expériences de Scherer et ses collaborateurs en 1984. En effet, durant cette dernière étude, différentes

procédures de dégradation ou de masquage ont été appliquées à des énoncés de parole afin de masquer les caractéristiques de l'intonation, de la qualité de la voix et du contenu verbal. Cette étude a révélé que les évaluations émotionnelles par les auditeurs semblaient être principalement déterminées par les indices de qualité de la voix (Scherer et al., 1984). En 1986, Scherer affirme même que la clé de la différenciation vocale des émotions discrètes semble être la qualité de la voix (Scherer, 1986).

2.3. Les fonctions de la prosodie

On distingue généralement trois grandes fonctions de la prosodie : les fonctions linguistiques, les fonctions attitudinales et les fonctions émotionnelles (Aubergé, 2002).

Les fonctions linguistiques de la prosodie incluent la structuration de l'énonciation, la modalisation et la focalisation du discours. La structuration (ou démarcation) de la parole a pour but de segmenter les phrases, de les découper en petites portions d'énoncé (mots ou groupes de mots), qui ont souvent partie liée avec l'organisation sémantique et syntaxique. Elle permet ainsi de repérer les unités lexicales d'une phrase (Lacheret et al., 1999).

La modalisation, elle, permet de définir la fonction modale d'une phrase : déclarative, interrogative, exclamative ou impérative. Elle donne ainsi des indices sur l'intention du locuteur et sa manière d'exprimer le message (Lacheret et al., 1999). Enfin, la focalisation (ou emphase) contribue quant à elle à mettre en valeur une ou des parties de l'énoncé. Elle met en avant principalement les unités porteuses de sens et véhiculant l'information principale du discours (Lacheret et al., 1999).

Par exemple, prenons l'énoncé « on mange du saumon ce soir ». La fonction de structuration permet ici de segmenter les unités lexicales de la phrase ; on retrouve donc 6 unités lexicales (autrement dit 6 mots). La fonction de modalisation précise la modalité de la phrase. Ici, avec une intonation montante, l'énoncé sera une interrogation : « on mange du saumon ce soir ↗ ? » ; a contrario, avec une intonation descendante, l'énoncé sera plutôt une déclaration : « on mange du saumon ce soir ↘ ». Enfin, la fonction de focalisation permet ici de mettre en valeur une unité lexicale pour véhiculer un sens précis ; pour « on mange du **saumon** ce soir », le sens est porté sur le contenu du repas, alors que pour « **on** mange du saumon ce soir », le sens est porté sur les personnes qui vont manger.

Les fonctions attitudinales de la prosodie concernent les attitudes volontaires et contrôlées, exprimées par un locuteur dans une situation précise. Elles sont orchestrées par les états

affectifs ou cognitifs du locuteur, ou encore par des stratégies d'interaction. Le locuteur apporte à son discours oral des informations sur ses intentions pendant son interaction verbale. Par exemple, si l'on reprend l'énoncé cité précédemment « on mange du saumon ce soir », nos intentions de véhiculer ce message peuvent être différentes. Dans ce cas précis, on peut vouloir convaincre, imposer ou montrer notre hésitation. Ainsi, nos attitudes et nos stratégies de communication varient en fonction de notre intention. Cette prosodie attitudinale est acquise assez tardivement dans l'interaction sociale, puisque c'est cette dernière qui permet de développer nos capacités en matière d'attitudes dans le langage (Clément, 1999).

Les fonctions émotionnelles de la prosodie sont régies par les événements qui provoquent l'état émotionnel du locuteur (Aubergé, 2002). Puisque notre mémoire porte sur la prosodie émotionnelle, c'est cet aspect que nous allons développer ci-dessous.

2.4. La prosodie émotionnelle

La prosodie émotionnelle véhicule un message acoustique avec une composante émotionnelle. Les études sur la prosodie émotionnelle décrivent des paramètres acoustiques différentiels selon les expressions des différentes émotions de base. Chaque émotion s'exprime différemment au niveau prosodique, et aucun indice acoustique de la prosodie à lui seul n'est suffisant pour identifier une émotion.

En fonction des dimensions de l'émotion (valence, arousal et pouvoir), les caractéristiques acoustiques du signal se modifient, ce qui entraîne des variations prosodiques. En effet, la dimension de l'arousal des émotions (le niveau d'excitabilité) impacte les paramètres de variations temporelles d'un signal : plus le locuteur est émotionnellement excité (c'est le cas pour un état émotionnel colérique), plus le rythme sonore est rapide, et inversement. L'arousal impacte aussi les paramètres fréquentiels : plus le locuteur est émotionnellement excité, plus on observe une augmentation du f_0 moyen et de son amplitude, et inversement (Caelen-Haumont, 2005).

Afin de rendre compte des variations des différentes caractéristiques prosodiques en fonction des émotions de base, on peut comparer les paramètres acoustiques de l'expression de chaque émotion avec une émotion la plus neutre possible. Il s'agit d'un état ne se rapprochant d'aucune émotion et qui sert ici de condition contrôle.

Cette comparaison est illustrée à travers le tableau ci-dessous :

	COLÈRE	JOIE	TRISTESSE	PEUR	DÉGOÛT
Débit	Légèrement plus rapide	Plus rapide ou plus lent	Légèrement plus lent	Beaucoup plus rapide	Vraiment beaucoup plus lent
Moyenne de f0	Vraiment beaucoup plus haut	Beaucoup plus haut	Légèrement plus bas	Vraiment beaucoup plus haut	Vraiment beaucoup plus bas
Amplitude de f0	Beaucoup plus large	Beaucoup plus large	Légèrement plus restreinte	Beaucoup plus large	Légèrement plus restreinte
Intensité	Plus forte	Plus forte	Plus faible	Normale	Plus faible
Qualité de la voix	Soufflée, voix de poitrine	Soufflée, explosive	Résonnante	Voisement irrégulier	Grognement, voix de poitrine
Variations de f0	Abruptes, sur les syllabes accentuées	Régulières, inflexion montantes	Inflexions descendantes	Normales	Grandes, inflexions terminales descendantes
Articulation	Tendue	Normale	Relâchée	Précise	Normale

Tableau 1 : Comparaison des paramètres prosodiques des émotions (adapté par Caelen-Haumont, 2005 ; de Murray et Arnott, 1993).

Nous allons principalement relever dans ce tableau les effets de la colère sur la prosodie et les paramètres qui la composent ; puisque c'est la colère et l'état émotionnel neutre qui sont au centre de notre mémoire. Nous relevons donc, pour un état émotionnel colérique, des paramètres de f0 et une intensité bien plus élevés, un débit de parole légèrement plus rapide, et une voix de poitrine qui semble plus soufflée avec une articulation plus tendue ; tout cela en comparaison à un état émotionnel neutre.

3. LES ÉTUDES SUR LA PROSODIE ÉMOTIONNELLE

3.1. Méthodologie des études sur l'expression de la prosodie émotionnelle

Les études sur l'expression de la prosodie émotionnelle nécessitent l'utilisation d'échantillons vocaux émotionnels. Ces corpus d'échantillons peuvent être créés grâce à différentes techniques. À ce jour, il existe trois grandes catégories de conditions d'encodage des émotions : les émotions authentiques, les émotions induites et les émotions jouées (Scherer, 2003).

3.1.1. Les conditions d'encodage des émotions

Les émotions naturelles obtenues in vivo

La première méthode permettant d'obtenir des échantillons vocaux émotionnels consiste à collecter des extraits audio tirés de contextes de vie quotidienne. Cette collecte de données peut être réalisée sans contrôle d'un examinateur à l'aide d'un appareil d'enregistrement, ou avec le contrôle d'un examinateur qui reste passif en ne donnant aucune directive. Par exemple, l'étude réalisée par Douglas-Cowie et ses collaborateurs, qui visait à montrer les composantes nécessaires à la construction d'un corpus d'échantillons vocaux émotionnels, a utilisé des énoncés représentant des extraits de talk-show sans le contrôle d'un examinateur (Douglas-Cowie et al., 2003). A contrario, pour l'étude de Williams et Stevens, les enregistrements réalisés dans un cockpit auprès de pilotes en interaction, ont été effectués sous le contrôle d'un investigateur (Williams et Stevens, 1972). Ces enregistrements semblent être le contenu de base idéal pour les recherches scientifiques de ce domaine car ils possèdent une validité écologique très élevée étant donné qu'ils sont issus de situations réelles (Scherer, 2003).

Cependant, on retrouve souvent des limites méthodologiques. En effet, ces échantillons vocaux sont fréquemment obtenus pour un très petit nombre de locuteurs car les conditions de collecte in vivo se concentrent généralement sur un nombre restreint de sujets. Cela s'explique par le fait que les situations évaluées sont généralement des interactions entre deux personnes, et rarement avec un groupe. De plus, il est difficile d'obtenir la même quantité de stimuli pour chacun d'eux car le temps et la fréquence de prise de parole ne sont pas contrôlables dans ces conditions. Pour la même raison, enregistrer les mêmes émotions pour chacun des locuteurs s'avère être complexe. Ceci est pourtant absolument essentiel étant donné l'importance d'avoir un échantillon équitable pour obtenir des résultats interprétables

(Banse et Scherer, 1996). De plus, les stimuli obtenus sont généralement très brefs et souffrent d'une mauvaise qualité d'enregistrement car ils ne sont pas enregistrés dans des conditions strictes de laboratoire. (Scherer, 2003).

Par ailleurs, il est nécessaire de remettre en question le fait d'utiliser cette méthode dans le but d'obtenir des émotions purement authentiques. En effet, le niveau réel de spontanéité de ces émotions est remis en cause par l'influence des normes sociales qui peuvent amener le locuteur à contrôler l'expression de ses émotions. Par exemple, le fait d'être observé (par un observateur externe, du public...) peut conduire les intervenants à focaliser leur attention davantage sur leur expression, étant donné que leurs prestations sont amenées à être jugées, et donc fournir une expression peu spontanée. En conclusion, la grande majorité de nos expressions, lorsqu'elles sont observables en public, sont comme des représentations, compte tenu d'un certain conformisme social (Banse et Scherer, 1996).

Les émotions induites

Une autre méthode permettant d'obtenir des échantillons vocaux émotionnels est l'induction expérimentale de l'émotion en laboratoire. C'est une méthode qui modifie, grâce à des manipulations expérimentales, l'état émotionnel des locuteurs qui produisent des stimuli. Cette méthode permet d'obtenir un corpus à la fois plus équitable en termes de quantité de stimuli et de locuteurs, mais aussi de cibler l'état émotionnel étudié.

Une fois les participants recrutés, l'élicitation émotionnelle se réalise sous plusieurs formes. On retrouve tout d'abord une méthode qui consiste à provoquer des inductions émotionnelles plutôt réalistes par le biais du comportement des expérimentateurs. Par exemple, Stemmler et ses collaborateurs, en étant arrogants et offensifs à travers leurs attitudes et leur langage, ont tenté de provoquer une réaction émotionnelle colérique de la part des sujets (Stemmler et al., 2001). Une autre méthode consiste à augmenter la charge cognitive du locuteur pendant qu'il produit son discours. Généralement, cela se fait grâce à des tâches cognitives assez difficiles (par exemple, des tâches de décision lexicale), ce qui tend à induire un stress cognitif. Ce dernier se traduit alors par des changements prosodiques émotionnels dans le discours du participant (Bachorowski et Owren, 1995 ; Scherer et al., 1998 ; Karlsson et al., 2000). Une autre technique très utilisée est celle de l'induction dite de Velten (Velten, 1968). Cette méthode demande aux locuteurs de lire plusieurs fois des phrases à forte charge émotive, ce qui entraîne chez eux des changements affectifs plutôt cohérents. Une autre alternative d'induction émotionnelle consiste en l'utilisation de matériel médiatique, agréable et

désagréable, à présenter aux sujets. Ce matériel prend la forme de diapositives, de photographies, d'images ou d'extraits de films. Cette dernière utilisation consiste à sélectionner des extraits de films susceptibles d'induire des émotions, que l'on présente ensuite à des sujets. Après visionnage de ces extraits, les participants doivent produire un contenu verbal susceptible d'être en lien avec l'émotion présentée (Rottenberg et al., 2007).

Cependant, cette méthode d'induction émotionnelle fait l'objet de controverses concernant son aspect éthique, à la fois parce que l'état émotionnel induit n'est pas nécessairement désiré par le locuteur, mais aussi parce que les méthodes utilisées vues précédemment peuvent parfois être vécues comme trop offensives par les sujets (Scherer, 2003).

Les émotions jouées

Une dernière méthode permettant d'obtenir des échantillons vocaux émotionnels consiste en la simulation d'émotions.

Afin de recueillir en laboratoire ce type d'échantillons joués, les expérimentateurs demandent à des participants (souvent, des acteurs) de produire des expressions vocales véhiculant une ou plusieurs émotions. Le plus souvent, un scénario typique écrit sous forme de script doit être lu ou appris par cœur par le locuteur. La particularité de ces scénarios est qu'ils illustrent une histoire qui doit être véhiculée avec une émotion en particulier. Autrement dit, les sujets sont amenés à utiliser des énoncés avec lesquels ils doivent simuler un état émotionnel bien spécifique (souvent parmi les 6 émotions de base). La linguistique comme la phonétique de ces énoncés sont prédéfinies avec ou sans contenu émotionnel (Aubergé et al., 2004). D'autres corpus de stimuli vocaux ont été recueillis sans l'appui d'un script, mais en demandant aux participants d'improviser en racontant une histoire, en développant un thème ou en décrivant un événement passé. Toutes ces improvisations doivent être jouées avec une émotion bien précise (Leinonen et al., 1997 ; Amir et al., 2000).

L'un des avantages importants du recueil d'expressions émotionnelles simulées en laboratoire est qu'il est généralement plus facile de demander à des participants de jouer une émotion particulière, plutôt que d'induire indirectement cette émotion chez eux. De plus, les questions éthiques ne représentent pas un obstacle ici puisque le consentement du participant est recueilli.

Dans cette même perspective de recueillir un échantillon de stimuli joués, les chercheurs ont aussi la possibilité de construire leur corpus émotionnel en dehors des laboratoires. Pour ce faire, plutôt que de recruter des participants et d'enregistrer leurs expressions jouées, ils peuvent, par exemple, directement sélectionner des extraits vocaux déjà réalisés. C'est notamment le cas en prenant des extraits de films dans lesquels des acteurs s'expriment avec une prosodie émotionnelle spécifique. L'intérêt d'utiliser des stimuli cinématographiques réside dans le fait d'éviter toute directive donnée en laboratoire concernant l'émotion à produire. De ce fait, pour la réalisation d'un film, les instructions pour susciter l'expression émotionnelle d'un acteur diffèrent de celles données en laboratoire (Enos et Hirschberg, 2008).

3.1.2. L'entraînement vocal

Bien que l'on sache peu de choses sur les impacts potentiels du jeu d'acteur sur l'expression et la perception des émotions, une grande partie de la recherche scientifique sur la prosodie émotionnelle utilise comme corpus des stimuli obtenus auprès d'acteurs (Drolet et al., 2012). Ce type de recherches part d'un postulat : l'expérience dans le domaine de la comédie aurait un impact positif sur la capacité à exprimer des émotions plus authentiques, plus réalistes. En effet, la formation et l'expérience d'un acteur sont un entraînement spécifique de la manipulation vocale, et plus précisément de l'expression de la prosodie émotionnelle. De plus, les différentes techniques de jeu des acteurs sont influencées par des sentiments propres ; leur formation serait donc censée accroître le réalisme de leur jeu, en s'appuyant sur leurs états affectifs intérieurs (Jürgens et al., 2015). À travers ces techniques de jeu basées notamment sur la méthode Stanislavski, créée par Constantin Stanislavski (professeur d'art dramatique russe), les acteurs apprennent à développer une action, une intention pour chaque script (Stanislavski, 1971). Cela se réalise sans nécessairement mentionner à l'acteur l'émotion désirée ; l'acteur doit plutôt se concentrer sur l'action à réaliser. Puis, en coopération avec le metteur en scène, ils s'imprègnent des traits de caractères du personnage à incarner et s'immiscent dans le contexte d'action de la scène. Ainsi, au fil de leur répétition, les acteurs formés arrivent à adopter pleinement leur personnage et les circonstances du scénario et s'engagent de ce fait plus spontanément dans l'action. Celle-ci induit à son tour l'expression émotionnelle résultante, à travers le corps mais aussi à travers la voix. Cette approche basée sur la méthode Stanislavski est généralement complétée par des techniques d'entraînement à la réactivité émotionnelle physique, mentale et vocale. Enos et Hirschberg soutiennent que ces approches, majoritairement utilisées dans la formation d'acteur, peuvent conduire à des expressions de la prosodie émotionnelle plus authentiques (Enos et Hirschberg, 2008).

Cependant, l'entraînement au jeu d'acteur ne serait pas qu'un avantage dans l'expression de la prosodie émotionnelle. En effet, au cours de leur formation, les acteurs apprennent des techniques de discours performatif et celles-ci ont notamment pour but d'être comprises de tous et de produire un discours émotionnel dans un but artistique (Aubergé et al., 2004). Ainsi, les représentations des acteurs sont alors influencées par des stéréotypes conventionnels de l'expression de l'émotion. D'autre part, au théâtre, les acteurs doivent transmettre leur message jusqu'au dernier rang ; cela nécessite de produire un discours de forte intensité et bien intelligible. Cette condition engendre bien souvent une surexpression de l'articulation et de la prosodie (Kracauer, 2005).

Au vu des désavantages liés à la formation d'acteur, l'idée d'utiliser des stimuli obtenus auprès de non-acteurs a commencé à émerger. Dans cette perspective, Jürgens en 2015 a tenté de définir quel groupe de sujets, parmi les acteurs et les non-acteurs, pouvait se rapprocher le plus intensément des expressions émotionnelles naturelles. Pour ce faire, l'auteur a sélectionné 80 stimuli tirés d'interviews dans la base de données d'une station de radio allemande dans lesquels les interviewés parlaient d'une situation chargée en émotion (colère, joie, tristesse ou peur). Ensuite, elle a aussi recruté 42 acteurs et 39 non-acteurs qui ont dû reproduire à leur manière une à trois fois les enregistrements authentiques. Ces énoncés devaient être représentés avec une des quatre émotions énoncées plus haut, qui étaient mentionnées dans la fiche descriptive de la situation de l'énoncé. Les acteurs, malgré leur expérience, avaient pour consigne d'utiliser le moins possible leur expression scénique, autrement dit devaient s'éloigner de leurs habitudes de jeu. Elle a ensuite comparé les caractéristiques acoustiques de stimuli d'émotions authentiques (issus des interviews de radio allemande) avec les caractéristiques acoustiques des expressions vocales jouées (par les acteurs et par les non-acteurs). L'analyse acoustique a été effectuée sur deux niveaux : le niveau 1 se concentrait sur les voyelles (a, e, i) simples et le niveau 2 concernait la séquence de parole. Les données acoustiques retrouvées chez les acteurs différaient de celles des non-acteurs et des émotions authentiques. En revanche, les profils acoustiques des stimuli des non-acteurs ressemblaient fortement à la structure de la parole authentique (Jürgens et al., 2015).

Finalement, la question d'utiliser préférentiellement des sujets acteurs ou des non-acteurs dans les études sur les expressions émotionnelles reste encore en suspens. À ce jour, aucune étude dans ce domaine n'a cherché à comparer les acteurs entre eux. Afin de préciser et d'approfondir les recherches déjà menées, il pourrait être pertinent de se demander s'il existe,

en fonction du type d'acteur, des variations dans les expressions. Existerait-il des différences dans leurs performances vocales émotionnelles suivant s'ils ont suivi telle ou telle formation, s'ils ont joué davantage au théâtre ou dans des films cinématographiques, ou encore s'ils ont plus ou moins d'expérience ? Sur la base de ces questionnements, Jürgens préconise que les études futures devraient se concentrer sur les différences individuelles dans l'encodage des émotions afin de démêler les effets de l'âge, du sexe, de l'utilisation de la voix, de l'humeur actuelle du locuteur sur le jeu de l'émotion, ou encore de l'expérience d'acteur non-professionnel, facteurs que son étude n'a pas pu prendre en compte (Jürgens et al., 2015).

3.2. Manipulation des indices de la prosodie émotionnelle

Comme nous l'avons vu précédemment, les études sur l'expression de la prosodie émotionnelle utilisent comme support des échantillons vocaux émotionnels : l'émotion s'exprime à la fois par le contenu verbal et par la prosodie émotionnelle de l'énoncé. Aussi, il est important de préciser que le lexique des stimuli donne des indices émotionnels aux auditeurs. Par exemple, des énoncés contenant des expressions comme « arrête », « tu m'énerves » ou encore « ça suffit » donnent des indices en faveur de l'émotion de la colère. Cependant, dans ces stimuli, prosodie émotionnelle et mots affectifs se confondent souvent et il est particulièrement difficile de mesurer séparément leur impact. Il paraît donc nécessaire de manipuler ce contenu verbal afin qu'il n'influence plus le jugement des auditeurs, qui écouteront les extraits dans les expériences de perception de la prosodie émotionnelle. Trois méthodes sont utilisées pour cette manipulation : la première concerne la mesure et la régression des indices, la deuxième le masquage des indices, et la troisième la resynthèse des indices (Scherer, 2003).

3.2.1. Mesure et régression des indices

Afin de mettre en avant les indices de la prosodie émotionnelle, la façon la plus simple de s'y prendre consiste en l'extraction des paramètres acoustiques d'échantillons vocaux d'émotions jouées ou naturelles, puis la confrontation de ces extraits avec les jugements des auditeurs sur l'émotion ou l'attitude sous-jacente du locuteur. Cette méthode permet d'obtenir des informations sur les caractéristiques vocales susceptibles de déterminer l'inférence des juges. Une fois ces extractions et confrontations réalisées, il est nécessaire d'utiliser un modèle statistique de régression multiple qui consiste à faire diminuer les paramètres acoustiques

représentant une inférence dans le jugement de l'émotion. Cette régression permet d'estimer la contribution relative de chacun de ces paramètres (Banse et Scherer, 1996). Cette manipulation a établi qu'une grande partie des inférences du jugement de l'émotion était expliquée par environ dix paramètres acoustiques, parmi lesquels figurent la fréquence fondamentale (f_0) moyenne, la déviation standard de la f_0 , l'intensité ou la durée des périodes voisées.

Cependant, cette méthode reste limitée lorsque le locuteur ne produit pas une émotion suffisamment extrême. Dans ce cas, l'analyse statistique ne donnera pas de résultats significatifs. Enfin, les résultats dépendent fortement de la précision de l'extraction et de la transformation des paramètres (Scherer, 2003).

3.2.2. Masquage des indices

Dans cette deuxième approche, les indices verbaux sont masqués, déformés ou supprimés. Cela permet d'explorer les effets qui découlent des inférences d'émotions sur les juges écoutant ces supports modifiés. Cette méthode de masquage des indices s'illustre au travers de différentes techniques. Chacune de ces dernières supprime ou déforme certains types d'indices acoustiques. Par exemple, nous retrouvons le filtrage passe-bas qui consiste à laisser passer les basses fréquences et à atténuer les hautes fréquences du signal sonore sans affecter le contour de f_0 . Ainsi, il contribue à supprimer l'intelligibilité et à modifier le timbre de la voix. L'épissage aléatoire, lui, supprime des informations de contour et de séquence de la parole sans altérer l'ensemble des informations spectrales (autrement dit, le timbre de la voix) (Scherer, 2003).

Toutes ces techniques vont permettre de modifier les indices acoustiques qui représentent des inférences dans la prosodie émotionnelle. Par conséquent, elles rendent possible l'utilisation de stimuli émotionnels quel que soit leur contenu verbal d'origine.

3.2.3. Resynthèse des indices

La technique de resynthèse des indices est une méthode plus récente qui s'est inspirée de l'avancée des technologies de la parole et des voix de synthèse. Elle permet de reconstruire des stimuli synthétisés à partir de paramètres acoustiques extraits de stimuli existants. Par exemple, dans l'objectif d'éliminer le contenu verbal (délexicaliser) d'un stimulus tout en gardant ses caractéristiques acoustiques ; nous commençons par extraire les paramètres

acoustiques caractéristiques de la prosodie de ce stimulus. Puis nous reconstruisons un nouveau stimulus en copiant les paramètres acoustiques extraits précédemment sur des extraits vocaux non verbaux issus de base de données de logiciels tels que MBROLA. Cette copie/resynthèse permet ici d'obtenir les mêmes caractéristiques acoustiques tout en n'ayant plus le contenu verbal (Ramus et Mehler, 1999). Cette dernière approche est à l'heure actuelle la plus prometteuse. Les avantages de cette technique seront passés en revue dans le chapitre 2 (cf. partie 3.2).

3.3. Méthodologie des études sur la perception de la prosodie émotionnelle

Les méthodes d'encodage des émotions vues précédemment ont fait l'objet de nombreuses études comparatives au niveau de leur perception. Ces études se concentrent davantage sur deux des méthodes d'encodage : les émotions spontanées et les émotions jouées. Les chercheurs se sont demandé s'il existait des différences significatives dans la perception d'une émotion lorsqu'elle est spontanée ou lorsqu'elle est simulée. Peut-on distinguer facilement une émotion authentique d'une émotion jouée ?

Pour répondre à cette question, Scheiner et Fischer ont recueilli pour leur étude 80 enregistrements de discours authentiques à partir d'archives radiophoniques allemandes et de quatre catégories d'émotions (20 stimuli pour chacune des catégories suivantes : colère, tristesse, joie et peur). Les auteurs ont ensuite fait rejouer le contenu de chacun de ces enregistrements par un acteur professionnel. L'ensemble des stimuli authentiques et joués (160 au total) ont été présentés à des participants de trois cultures différentes (Allemagne, Roumanie et Indonésie). Ces participants devaient tenter de reconnaître la condition d'encodage des stimuli (choix alternatif : authentique ou jouée). Les résultats ont montré que les participants n'ont pas réussi à faire la distinction entre une émotion jouée et une émotion authentique (seulement légèrement au-dessus du hasard) (Scheiner et Fischer, 2011).

Dans la même perspective, Jürgens a utilisé des stimuli authentiques, des stimuli joués par des acteurs, et des stimuli joués par des non-acteurs (cf. méthodologie expliquée dans la partie 3.1.4). Ces stimuli représentaient aussi la colère, la tristesse, la joie et la peur. L'ensemble des stimuli authentiques et joués (240 au total) ont été présentés à des participants pour une épreuve de perception. Les résultats ont montré que la condition d'encodage (authentique ou jouée) a été reconnue dans 59,8% des cas (Jürgens et al., 2015).

Les taux de reconnaissance de la condition d'encodage dans ces deux études étant assez faibles, cela indique une capacité restreinte des participants à juger la condition d'encodage. De ce fait, ces mêmes études ont aussi tenté de lever le voile sur le lien entre condition d'encodage et reconnaissance des émotions.

L'étude de Scheiner et Fischer vue précédemment a également mesuré la reconnaissance de l'émotion (choix alternatif : colère, tristesse, joie ou peur). Bien que les participants n'aient pas réussi à faire la distinction entre les deux conditions d'encodage, celles-ci ont tout de même influencé la reconnaissance de l'émotion : dans les trois cultures, la colère a été reconnue avec plus de précision lorsqu'elle était jouée, et la tristesse lorsqu'elle était authentique.

De plus, les expressions des acteurs ont été reconnues avec plus de précision que celles des non-acteurs uniquement pour les stimuli de colère, de peur et de joie. Le sexe du participant n'a eu aucun effet sur la reconnaissance de l'encodage (Scheiner et Fischer, 2011).

Dans cette même perspective, Jürgens a aussi mesuré la reconnaissance de l'émotion à travers son étude. Les différentes émotions ont été correctement reconnues dans 44,7% des cas. De plus, là encore, la colère jouée a été mieux reconnue que la colère authentique, tandis que la tristesse et la joie ont été mieux reconnues à travers une expression authentique (Jürgens et al., 2015). Finalement, on a pu observer à travers ces deux études que l'émotion jouée la mieux reconnue est la colère.

4. QUESTIONNEMENTS ET OBJECTIFS DE NOTRE ÉTUDE

Ainsi, nous avons pu constater que certaines des études scientifiques sur les émotions mentionnées précédemment ont cherché à comparer la parole authentique (autrement dit, celle de la vie quotidienne) et la parole jouée. Cependant, ces recherches ont rencontré un succès mitigé pour deux raisons principales. La première raison, c'est que cela peut induire des limites d'un point de vue éthique concernant l'enregistrement de la parole authentique. En effet, comment est-il possible d'envisager d'enregistrer quelqu'un en colère sans lui demander sa permission au préalable ? La deuxième raison est simple : si nous demandions à un participant sans entraînement vocal de produire une émotion, sa production vocale, et plus précisément sa prosodie émotionnelle, serait très probablement surjouée et peu naturelle.

Dans la perspective de contourner ces deux limites, d'autres études se sont alors servi d'extraits radio, d'entretiens, de talk-show... Bien que cela permette d'obtenir des stimuli authentiques, il est tout de même nécessaire de garder à l'esprit que cela reste de la parole pas tout à fait spontanée. En effet, face à une caméra ou un micro, nous avons tendance à

contrôler notre parole pour nous conformer à une certaine exigence sociale. De plus, nous avons pu observer, au travers de ces études, qu'extraire des échantillons dans de telles conditions entraîne une qualité d'enregistrement fluctuante et un nombre restreint d'interlocuteurs. Ces projets ont tout de même servi à démontrer une certaine similarité des caractéristiques acoustiques de la prosodie émotionnelle entre celle véhiculée par les voix des individus filmés et celle par les voix des acteurs en chambre sourde dans les laboratoires.

Nous avons pu constater que c'est généralement dans ces laboratoires que les études cliniques sur les émotions jouées se réalisent, et cela grâce à l'enregistrement de sujets qui sont très souvent, voire toujours, des acteurs. Ces derniers sont amenés à simuler des émotions variées dans des conditions souvent très contrôlées. Choisir des acteurs pour tenter de recueillir un discours émotionnel le plus naturel possible semble être une stratégie raisonnable, étant donné que ce sont des sujets ayant reçu une formation poussée dans l'art de la simulation. Cependant, les vraies questions que cette méthode a soulevées sont les suivantes : la formation d'un acteur lui permet-elle d'exprimer une émotion vraiment réaliste, authentique, et proche de ce que l'on exprime émotionnellement dans la vie quotidienne ? Ou, au contraire, est-ce qu'elle n'induit pas un discours sous la base d'un apprentissage plutôt stéréotypé, surjoué, et loin de celui de la vie quotidienne ?

Finalement, très peu d'études ont cherché à répondre à ces questionnements en comparant les performances d'acteurs et de non-acteurs. Bien qu'il n'y ait pas encore de consensus, les travaux les plus récents ont émis l'hypothèse que, contre toute attente, les sujets non-acteurs pourraient être généralement plus susceptibles de se rapprocher d'une expression émotionnelle authentique. Nous avons pu constater que, jusqu'à aujourd'hui, les études cherchent encore à élucider les origines de ces disparités dans les capacités de jeu émotionnel, et que finalement, les auteurs préconisent de s'intéresser davantage aux caractéristiques interindividuelles des participants qui sont, en l'occurrence, des acteurs.

C'est pour cette raison que nos problématiques sont les suivantes :

Existe-t-il un lien entre l'expérience professionnelle d'un acteur et sa manière d'exprimer une émotion ? Est-ce que cette expérience professionnelle impacte aussi notre façon de percevoir la prosodie émotionnelle véhiculée par un acteur ?

C'est grâce aux études réalisées précédemment ainsi qu'aux questions qu'elles soulèvent que nous avons décidé d'utiliser, pour notre projet, des groupes d'acteurs professionnels ayant une grande expérience cinématographique, et des acteurs non-professionnels débutant dans le domaine du cinéma.

Choisir des acteurs non-professionnels est d'autant plus motivé par le fait que, dans le cinéma, ces acteurs qui n'ont que très peu d'expérience cinématographique sont de plus en plus recrutés. Cette tendance nouvelle montre que les réalisateurs leur portent un réel intérêt : ils

trouvent certainement chez eux quelque chose de plus authentique dans leur posture, dans leur voix, et donc dans leur façon d'exprimer les émotions. Par exemple, dans le film primé *Bande de filles* réalisé par Céline Sciamma, l'ensemble des rôles principaux a été attribué à des acteurs non-professionnels tels que Karidja Touré, Assa Sylla ou encore Simina Soumaré. Au travers de cette étude, nous avons tenté d'explicitier et de quantifier les différences entre les deux groupes d'acteurs grâce à des extraits de films exprimant deux états émotionnels bien distincts : la colère (qui est une des six émotions de base) et l'état émotionnel neutre (qui sera notre condition contrôle). À travers cela, nous avons cherché à explorer et à élucider la question de l'impact de l'entraînement vocal sur l'expression des émotions.

CHAPITRE 2 : MÉTHODOLOGIE DE L'ÉTUDE

La première étape méthodologique consiste en la construction d'un corpus d'énoncés représentatifs de la prosodie colérique jouée par des acteurs expérimentés (professionnels) et des acteurs non expérimentés (non-professionnels). Ce corpus comporte également des énoncés dont la prosodie est neutre, c'est-à-dire qui n'expriment aucune émotion particulière.

Dans un premier temps, afin de constituer ce corpus, une étudiante en cinéma a sélectionné de courts extraits de films francophones joués par des acteurs eux-mêmes choisis selon des critères de sélection spécifiques. Nous procéderons ensuite à une analyse acoustique de ces énoncés. Enfin, le contenu verbal de ces énoncés fera l'objet d'une délexicalisation. Dans le chapitre 3, cette délexicalisation sera vérifiée par le biais d'un test d'intelligibilité (tâche de validation n°1). De plus, un test d'identification émotionnelle sera réalisé dans le but de vérifier la classification dans les catégories « colère » et « état émotionnel neutre » de chaque stimulus (tâche de validation n°2). Enfin, nous réaliserons une étape d'évaluation des dimensions émotionnelles.

1. LA SÉLECTION DES STIMULI

1.1. Les sujets

1.1.1. Acteurs professionnels

Le premier groupe de sujets choisis pour notre étude sont des acteurs expérimentés dans le cinéma (désignés comme professionnels). Ils répondent tous à un critère en particulier : leurs extraits choisis pour l'étude ne doivent pas être issus de leur première apparition cinématographique. Chacun de ces acteurs francophones doit donc avoir déjà joué dans plusieurs films, dans des rôles assez importants pour justifier d'une expérience cinématographique riche. Ces acteurs avaient entre 1 et 157 films à leur actif avant de jouer dans le film choisi. Ce critère nous assure d'une expérience de jeu d'acteur conséquente, autrement dit d'un entraînement vocal avancé.

Ces acteurs sont au nombre de 14 : 9 d'entre eux sont des hommes, et 5 sont des femmes. Ils étaient âgés de 15 à 63 ans lors de leur participation au film sélectionné.

ACTEURS PROFESSIONNELS (14)					
HOMMES (9)			FEMMES (5)		
Nom de l'acteur	Âge au moment du film	Nombre de films joués avant le film choisi	Nom de l'actrice	Âge au moment du film	Nombre de films joués avant le film choisi
Giorgio Albertazzi	38 ans	3	Nathalie Baye	33 ans	20
Bernard Blier	63 ans	157	Béatrice De Staël	52 ans	3
Jean Carmet	59 ans	129	Valérie Donzelli	37 ans	12
Michel Fortin	42 ans	28	Léa Seydoux	31 ans	27
Claude Hébert	38 ans	2	Rosalie Symon	15 ans	1
Gérard Lanvin	31 ans	10			
Marco Perrin	52 ans	36			
Jean Rougerie	50 ans	15			
Jean Yanne	39 ans	26			

Tableau 2 : Acteurs professionnels sélectionnés pour l'étude.

1.1.2. Acteurs non-professionnels

Le deuxième groupe de sujets choisis pour notre étude sont des acteurs non expérimentés dans le cinéma (désignés comme non-professionnels). Ils répondent tous à un critère en particulier : leurs extraits choisis pour l'étude doivent être issus de leur première apparition cinématographique. Ce critère de sélection nous permet de nous assurer d'une expérience cinématographique pauvre, à la différence des acteurs professionnels. Cependant, ces acteurs pouvaient avoir déjà eu une autre expérience au théâtre, en doublage de voix, en publicité... Ce critère de sélection s'explique par le fait que ces expériences en dehors du cinéma sont particulièrement difficiles à contrôler.

Ces acteurs sont au nombre de 9 : 3 d'entre eux sont des hommes, et 6 sont des femmes. Ils étaient âgés de 18 à 52 ans lors de leur participation au film sélectionné.

ACTEURS NON-PROFESSIONNELS (9)					
HOMMES (3)			FEMMES (6)		
Nom de l'acteur	Âge au moment du film	Nombre de films joués avant le film choisi	Nom de l'actrice	Âge au moment du film	Nombre de films joués avant le film choisi
Ghislain Ghesquière	Non connu	0	Farida Benkhetache	Non connu	0
Dylan Robert	18 ans	0	Séverine Caneele	25 ans	0
Emmanuel Schotté	41 ans	0	Emilie Dequenne	18 ans	0
			Simina Soumaré	Non connu	0
			Assa Sylla	18 ans	0
			Karidja Touré	20 ans	0

Tableau 3 : Acteurs non-professionnels sélectionnés pour l'étude.

1.2. Les films

Les extraits audio sélectionnés pour l'étude ont été tirés de 12 films tous francophones, ce qui nous a permis d'interpréter correctement la voix originale de l'acteur. Ces films ont été retenus selon le critère de sélection suivant : chaque film a été primé ou nommé pour une récompense cinématographique (César, Palme d'Or, Grand Prix...), et/ou l'acteur jouant l'extrait a été lui-même primé ou nommé pour une récompense cinématographique pour ce film. Ce critère de sélection permet de s'assurer de la qualité à la fois du film et de l'acteur, qu'il soit expérimenté ou non.

NOM	DATE DE SORTIE	PRIX OU RÉCOMPENSES	ACTEURS SÉLECTIONNÉS
L'année dernière à Marienbad	1961	Prix Gassman (2004)	Giorgio Albertazzi
Nous ne vieillirons pas ensemble	1972	Prix d'interprétation masculine au Festival de Cannes (1972)	Jean Yanne
Buffet froid	1979	César du meilleur scénario adapté en 1980 + 3 nominations aux Césars (1980) pour la photographie, le décor et le montage	Bernard Blier Jean Carmet Michel Fortin Marco Perrin Jean Rougerie
La Drôlesse	1979	Nommé pour la Palme d'or du meilleur film (1972)	Claude Hébert
Une étrange affaire	1981	César de la meilleure actrice dans un second rôle (1982) + Prix Gabin (1982)	Nathalie Baye Gérard Lanvin
L'Humanité	1999	Grand Prix du Jury du Festival de Cannes (1999) + Prix d'interprétation féminine et masculine au Festival de Cannes (1999)	Ghislain Ghesquière Emmanuel Schotté Séverine Caneele
Rosetta	1999	Prix d'interprétation féminine au Festival de Cannes (1999)	Emilie Dequenne
La graine et le mulet	2001	Grand Prix du Jury à la Mostra de Venise (2007) + Prix Méliès (2007) + César du Meilleur Film (2008)	Farida Benkhetache Rosalie Symon
La Reine des Pommes	2010	Prix du jury des jeunes au Festival International des Jeunes Réalisateurs de St Jean- de luz (2009) + Grand prix du Festival du film de Cabourg (2011)	Valérie Donzelli Béatrice de Staël
Bande de filles	2014	Prix spécial de l'Académie des Lumières (2015) + Révélation de l'année au Grand Prix Elle Cinéma (2014) + Black Reel Awards 2016 : nomination comme meilleure Révélation Féminine + 3 nominations aux Oscars dont meilleur espoir féminin	Karidja Touré Assa Sylla Simina Soumaré
Juste la fin du monde	2016	Palme d'Or + Meilleur Espoir + Meilleure Actrice (pour l'actrice sélectionnée), Grand Prix Festival de Cannes (2017) et 3 Césars pour le film	Léa Seydoux
Shérazade	2018	César du Meilleur Espoir Masculin (2019) + Prix Lumières : Lumière de la révélation masculine (2019)	Dylan Robert

Tableau 4 : Présentation des films sélectionnés.

1.3. Les émotions

Les différents extraits de films joués par les acteurs présentés ci-dessus devaient représenter soit une émotion de colère, soit un état émotionnel neutre.

Comme expliqué précédemment, les stimuli ont été choisis par une étudiante en cinéma et vérifiés par deux linguistes experts des émotions. Les extraits vocaux issus de cette première sélection sont au nombre de 73. Ils représentent de courtes phrases de 5,85 mots en moyenne.

1.3.1. La colère

La colère a été définie comme une des six émotions de base (émotions fondamentales) par Ekman (Ekman, 1992). Pour cette étude, notre choix de la colère a été influencé par l'étude de Scherer dans laquelle la colère et la tristesse sont généralement les mieux reconnues à travers la voix (Scherer, 2003). De plus, il existe deux types de colère : la colère chaude qui est de nature impulsive et intense, et la colère froide qui est plus réfléchie et moins explosive. Ici, nous avons décidé de nous concentrer sur la colère chaude. En effet, d'un point de vue prosodique, ses contours très marqués nous permettront d'obtenir des résultats plus probants en comparaison avec les contours d'un état émotionnel neutre.

Dans notre étude, 38 stimuli représentant la colère ont été sélectionnés par une étudiante en cinéma (annexe 1). Ces stimuli représentent de courtes phrases de 5,23 mots en moyenne.

1.3.2. L'état émotionnel neutre (condition contrôle)

Un contenu émotionnel neutre signifie qu'il ne contient pas d'émotion particulière. Nous avons donc sélectionné des extraits dans lesquels la prosodie s'avère neutre, sans intention d'expression d'une émotion quelconque. Cet état émotionnel neutre servira de condition contrôle et permettra une analyse distinctive et une comparaison avec les autres stimuli représentant la colère, émotion particulièrement marquée.

Dans notre étude, 35 stimuli représentant un état émotionnel neutre ont été sélectionnés par la même personne que pour les stimuli de la colère (annexe 2). Ces stimuli représentent de courtes phrases de 6,51 mots en moyenne.

2. LES ANALYSES ACOUSTIQUES

Nous nous sommes tout d'abord assurées de la qualité acoustique des différents extraits : tous n'ont peu ou pas de bruits de fond et sont de bonne qualité sonore.

Ensuite, chaque extrait a subi les mêmes analyses acoustiques, c'est-à-dire sur son intensité, puis sur sa fréquence fondamentale minimum, moyenne et maximum, ce qui nous a permis d'obtenir aussi le f_0 range (étendue vocale) en soustrayant la fréquence fondamentale minimum à la fréquence fondamentale maximum. Par la suite, nous avons calculé la moyenne et la déviation standard de chaque paramètre acoustique des différents extraits classés en quatre catégories (colère pour les professionnels, colère pour les non-professionnels, neutre pour les professionnels et neutre pour les non-professionnels).

Afin de réaliser les différentes analyses acoustiques des stimuli, nous avons utilisé le logiciel Praat (Boersma et al., 2021). Praat est un logiciel libre scientifique conçu pour la manipulation, le traitement et la synthèse de sons vocaux. Pour la collecte des différentes données et les calculs, nous avons utilisé Microsoft Excel.

2.1. Mesure de fréquence fondamentale (f_0) et d'intensité vocale

Nous avons extrait les données de fréquence fondamentale (f_0) et d'intensité pour chaque stimulus. Nous avons ensuite calculé à l'aide d'Excel les moyennes et les déviations standard de chaque donnée (f_0 minimum, f_0 maximum, moyenne de f_0 , f_0 range et intensité moyenne). Les valeurs de déviation standard nous donnent une indication sur la variabilité autour de la moyenne : plus la déviation standard est élevée, plus les données obtenues pour un échantillon sont variées.

Ainsi, ces différentes analyses acoustiques, nous permettront de procéder à une description et une comparaison des caractéristiques prosodiques spécifiques pour chacune des quatre catégories (colère pour les professionnels, colère pour les non-professionnels, neutre pour les professionnels et neutre pour les non-professionnels). De plus, pour une analyse plus fine, nous séparerons les hommes et les femmes.

2.2. Résultats des analyses

2.2.1. Pour la colère

Les résultats obtenus pour la colère sont présentés dans le tableau ci-dessous (tableau 5) :

MOYENNES OBTENUES POUR LA COLÈRE (σ = Déviation standard)				
PARAMÈTRES ACOUSTIQUES	HOMMES		FEMMES	
	PROFESSIONNELS	NON- PROFESSIONNELS	PROFESSIONNELS	NON- PROFESSIONNELS
f0 minimum	201 (σ = 73)	213 (σ = 31)	216 (σ = 109)	228 (σ = 59)
f0 maximum	386 (σ = 119)	332 (σ = 55)	470 (σ = 75)	459 (σ = 77)
Moyenne f0	307 (σ = 94)	269 (σ = 40)	355 (σ = 79)	351 (σ = 55)
f0 range	185 (σ = 88)	120 (σ = 45)	254 (σ = 59)	230 (σ = 91)
Intensité	75 (σ = 2)	78 (σ = 3)	69 (σ = 4)	70 (σ = 6)

Tableau 5 : Moyennes obtenues pour la colère et pour chaque paramètre acoustique.

Ainsi, pour la colère, nous remarquons à travers le tableau 5, que les valeurs moyennes sont sensiblement similaires entre les deux catégories d'acteurs (professionnels et non-professionnels) et pour chacun des paramètres acoustiques.

Nous relevons tout de même que les différences entre les acteurs professionnels et non-professionnels dans l'expression de la colère sont plus marquées pour les hommes que pour les femmes. En effet pour les hommes, nous retrouvons des valeurs de f0 maximum, f0 moyenne et de f0 range avec un écart plus important que pour les femmes, entre professionnels et non professionnels. Aussi, les acteurs professionnels ont en moyenne une étendue vocale (f0 range) supérieure à celle des acteurs non-professionnels. Pour l'intensité, nous ne retrouvons pas de grandes différences. Enfin, les déviations standard obtenues sont généralement plus élevées pour les acteurs professionnels que pour les acteurs non-

professionnels, que ce soit pour les hommes ou pour les femmes. Ces derniers résultats témoignent d'une plus grande variabilité vocale pour les acteurs professionnels dans l'expression de la colère.

2.2.2. Pour l'état émotionnel neutre

Les résultats obtenus pour l'état émotionnel neutre sont présentés dans le tableau ci-dessous (tableau 6) :

MOYENNES OBTENUES POUR L'ÉTAT ÉMOTIONNEL NEUTRE (σ = Déviation standard)				
PARAMÈTRES ACOUSTIQUES	HOMMES		FEMMES	
	PROFESSIONNELS	NON- PROFESSIONNELS	PROFESSIONNELS	NON- PROFESSIONNELS
f0 minimum	111 (σ = 35)	111 (σ = 23)	160 (σ = 42)	144 (σ = 36)
f0 maximum	223 (σ = 92)	225 (σ = 114)	293 (σ = 112)	271 (σ = 45)
Moyenne f0	141 (σ = 43)	149 (σ = 40)	203 (σ = 34)	206 (σ = 32)
f0 range	112 (σ = 95)	113 (σ = 97)	134 (σ = 127)	127 (σ = 58)
Intensité	68 (σ = 6)	71 (σ = 3)	64 (σ = 6)	68 (σ = 9)

Tableau 6 : Moyennes obtenues pour l'état émotionnel neutre et pour chaque paramètre acoustique.

Pour l'état émotionnel neutre, nous retrouvons encore des valeurs moyennes sensiblement identiques entre les deux catégories d'acteurs (professionnels et non-professionnels) et pour chacun des paramètres acoustiques. Ce phénomène est encore plus marqué que pour la colère.

Nous remarquons également que les déviations standard obtenues sont généralement plus élevées pour les acteurs professionnels que pour les acteurs non-professionnels, que ce soit

pour les hommes ou les femmes. Là encore, comme pour la colère, les résultats témoignent d'une plus grande variabilité vocale pour les acteurs professionnels.

2.2.3. Interprétation des résultats

Pour chaque paramètre acoustique relevé, nous avons comparé les résultats obtenus pour les professionnels et pour les non-professionnels. Ainsi, nous n'avons pas relevé de différences significatives entre les deux groupes d'acteurs ; ce phénomène s'applique d'autant plus pour l'état émotionnel neutre que pour la colère. Nous pouvons en conclure que les acteurs, qu'ils soient professionnels ou non, expriment sensiblement de la même façon un état émotionnel neutre. Cependant, les valeurs obtenues pour la déviation standard de la colère sont plus élevées que celles retrouvées pour l'état émotionnel neutre. Toutes ces observations témoignent d'une palette plus large d'expression de la colère que de l'état émotionnel neutre, que ce soit pour les professionnels et les non-professionnels.

De plus, nous avons également comparé les résultats obtenus pour la colère et pour l'état émotionnel neutre. Ainsi, quel que soit le sexe ou l'expérience de l'acteur, les valeurs acoustiques ont révélé l'existence d'une interaction entre l'émotion et le modèle acoustique : nous retrouvons des valeurs acoustiques différentes en fonction des états émotionnels exprimés. En effet, pour la colère, l'ensemble des paramètres acoustiques s'est avéré plus élevé que pour l'état émotionnel neutre ; ces résultats sont conformes à la littérature à travers laquelle la grande variabilité des contours acoustiques est liée aux émotions les plus éveillées comme la colère, tandis qu'une faible variabilité est relevée pour les expressions moins éveillées, comme l'état émotionnel neutre (Scherer, 2003 ; Jürgens et al., 2015).

En ce qui concerne l'intensité, nous nous attendions à obtenir une différence bien plus importante entre la colère et l'état émotionnel neutre. Ce phénomène s'explique probablement par des changements de position du microphone ; il aurait donc été plus pertinent de calculer des mesures d'intensité relative.

3. L'INTERPRÉTATION ET LA MODIFICATION DES STIMULI

Après avoir effectué les différentes analyses acoustiques, nous avons également travaillé sur le contenu verbal de chaque extrait avec l'aide de différents examinateurs professionnels et étudiants en linguistique. En effet, certains stimuli présentaient une intelligibilité imprécise. Afin de vérifier cette intelligibilité, nous avons tout d'abord effectué une transcription du contenu verbal des stimuli. Par la suite, nous avons vérifié leur neutralité émotionnelle et une délexicalisation a été réalisée sur chaque stimulus afin de masquer leur contenu verbal.

3.1. Transcription des stimuli et vérification de la neutralité du contenu verbal

Le contenu verbal des différents extraits analysés représente de courtes phrases ou groupes de mots allant de 0,46 à 1,91 seconde.

Afin de vérifier leur intelligibilité et leur neutralité linguistique, ces stimuli ont été transcrits par quatre examinateurs : nous-mêmes, une linguiste et un étudiant en linguistique. Cette transcription a été réalisée de deux façons. Tout d'abord, les extraits ont été transcrits de manière phonétique, c'est-à-dire chaque examinateur a transcrit les sons tels qu'il les a entendus. Puis, une seconde transcription a été effectuée : les mots entendus ont été rapportés avec la bonne orthographe. Certaines transcriptions ont été différentes selon les examinateurs : elles sont au nombre de 14 sur les 73 stimuli totaux (tableau 6). Ces deux transcriptions ont permis d'une part de vérifier l'intelligibilité commune, le sens des énoncés, et d'autre part la présence éventuelle de mots à connotation émotionnelle.

Dans l'objectif que les stimuli n'influencent pas le jugement de la prosodie, il fallait s'assurer que le contenu linguistique de l'extrait ne contenait pas d'indice émotionnel. Plus précisément, les mots entendus et transcrits ne devaient pas faire partie du champ lexical d'une émotion en particulier. Le contenu lexical des stimuli sélectionnés était variable et présentait parfois des mots avec une connotation émotionnelle, comme « vas-y arrête » ou encore « exécution » qui ont une valence très négative. Il a donc été nécessaire de masquer le contenu lexical afin d'éviter des interférences avec le jugement de la prosodie (Kramer, 1964).

TRANSCRIPTIONS DIFFÉRENTES		
1	Vas-y arrête !	Oh vas-y arrête !
2	Tu tires dans le tunnel	T'es toujours dans le tunnel
3	Elle est à vous	Elle est pas X
4	J'ai autre chose à faire	Et j'ai autre chose à faire
5	La grève, elle est finie	Anna, elle va les finir
6	Elle est à mon père	Bah y'a mon père
7	Ce n'est pas vrai que nous ayons besoin de l'argent	Ce n'est pas vrai que nous ayons besoin de l'X
8	Pas comme ça	Pas content
9	En général, j'arrive trop tard	Mon général, j'arrive trop tard
10	Tu veux pas que je descende Filo ?	Tu veux pas que je descende Filou ?
11	Dis, on sort on va à la brasserie	Dis aux hommes on va à la brasserie
12	Elle est à mon père	Bah y'a mon père
13	Je suis pas malade moi	Mais je suis pas malade moi
14	Pas vraiment	Oh vraiment

Tableau 7 : Les 14 stimuli aux transcriptions discordantes.

Ainsi, en s'appuyant sur les données de la littérature et en réponse à l'intelligibilité imprécise et au manque de neutralité émotionnelle de certains extraits audio sélectionnés, nous avons décidé de délexicaliser l'ensemble des stimuli.

3.2. Délexicalisation

Étant donné que le contenu verbal des stimuli présente une connotation émotionnelle, il paraît nécessaire de manipuler le contenu verbal afin qu'il n'influence pas le jugement des auditeurs qui écouteront les extraits dans notre expérience d'évaluation de la prosodie émotionnelle.

Comme nous l'avons précisé précédemment (cf. Chapitre 1, partie 3.2) dans la littérature, il existe différentes techniques de masquage ; la plus utilisée est celle du filtrage passe-bas (Scherer, 2003).

3.2.1. Méthode de filtrage passe-bas

Concrètement, cette méthode permet d'éliminer l'information acoustique sur une plage de fréquences donnée. Normalement, les fréquences comprises entre 400 et 600 Hz sont supprimées car la compréhension du contenu lexical dépend fortement de ces fréquences (Fletcher, 1953 ; Licklider et Miller, 1951). Après réalisation de cette méthode, les sons semblent être comme écoutés à travers un mur (Starkweather, 1956).

Cependant, cette méthode a laissé apparaître plusieurs désavantages. Tout d'abord, elle ne permet pas de savoir quelles propriétés du signal sont éliminées et lesquelles sont préservées car le système est basé sur une approximation. Ensuite, cette méthode ne s'applique pas sur les fréquences hautes, alors que la fréquence fondamentale (f_0) est souvent supérieure au seuil de 400/500 Hz, d'autant plus pour les voix féminines et dans le cas de la colère (Kramer, 1964 ; Ramus et Mehler, 1999). Aussi, du fait que certaines informations segmentales sont préservées, il est possible que certains mots soient encore reconnus (Kramer, 1964). De plus, le filtrage passe-bas entraîne une distorsion importante de la qualité de la voix. En effet, la proportion d'énergie diffère de phonème en phonème ; par exemple, la voyelle /a/ a beaucoup plus d'énergie dans les basses fréquences que la voyelle /i/. Le filtrage passe-bas donne ainsi une amplification injustifiée de la voyelle /a/ (Ramus et Mehler, 1999). Enfin, la parole filtrée par le biais de cette méthode semble être « étouffée » avec un effet très peu naturel.

Toutes ces constatations nous ont amenées à utiliser une autre méthode pour la délexicalisation des stimuli.

3.2.2. Méthode d'analyse et de resynthèse

En raison des limites du filtrage passe-bas démontrées précédemment, il a été décidé d'appliquer une technique d'analyse et de resynthèse. Cette méthode, appliquée ici par Francesca Carbone, permet de créer des stimuli délexicalisés dont la fréquence fondamentale, l'intensité et les caractéristiques spectrales ont été reconstruites à partir des fichiers audio originaux.

Concernant la reconstruction de la fréquence fondamentale, elle a été réalisée grâce au logiciel MBROLA (Dutoit et al., 1996). C'est un logiciel qui est fréquemment utilisé en psycholinguistique pour les travaux de recherche sur la perception de la prosodie émotionnelle (Ramus et Melher, 1999 ; Audibert, 2008). Le logiciel MBROLA consiste en la synthèse de stimuli par la concaténation de diphones ; tout cela en utilisant une base de données de diphones du français. Les voyelles et les consonnes des fichiers audio originaux subissent des modifications : les fricatives sont remplacées par le phonème /s/, les occlusives par le phonème /t/, les liquides par le phonème //, les nasales par le phonème /n/, les glissantes par le phonème /j/, et les voyelles par le phonème /a/. Cette modification, appelée « saltanaj », permet de préserver le contour intonatif des stimuli originaux tout en supprimant le contenu lexical.

Concernant la reconstruction de l'intensité, elle a été réalisée à l'aide de la Toolbox de Praat grâce à la fonction « Copy intensity contour ».

Enfin, concernant la reconstruction des caractéristiques spectrales, elle a été réalisée grâce à une procédure de morphing vocal par le système STRAIGHT (Speech Transformation and Representation based on Adaptive Interpolation of weiGHTEd spectrogram) (Kawahara, 2006) dans Matlab. STRAIGHT est capable d'analyser puis de resynthétiser diverses qualités vocales sur la base d'un algorithme : la fréquence fondamentale (f_0), les fréquences de formants, la durée, la densité spectro-temporelle et la périodicité. La qualité de la voix est alors intégrée aux combinaisons de ces paramètres. Ainsi, un taux de morphing vocal de 0,5 a été appliqué à chaque stimulus.

Cette méthode présente plusieurs avantages. Tout d'abord, les reconstructions de la fréquence fondamentale f_0 et de l'intensité sont précises et évitent la perte non contrôlée d'informations prosodiques comme c'est le cas pour le filtrage passe-bas (Ramus et Mehler, 1996). De plus, les caractéristiques spectrales et la qualité de la voix des locuteurs sont gardées. Enfin, les stimuli resynthétisés sont perceptivement naturels.

Pour illustrer la méthode de resynthèse appliquée par Francesca Carbone, nous montrons pour exemple les différentes étapes de la méthode sur un stimulus (np_caneele_c_1) à travers les figures 2, 3 et 4. Chacune des figures présente le spectrogramme et le sonogramme du stimulus. Ce dernier est un extrait représentant la colère et joué par l'actrice Séverine Caneele, classée dans la catégorie des acteurs non-professionnels. Son contenu verbal est « C'est rien du tout ».

La figure 2 présente le stimulus dans sa forme originale, avec son contenu verbal. La figure 3 correspond au stimulus obtenu après reconstruction de la fréquence fondamentale avec MBROLA. Enfin, la figure 4 représente le stimulus dans sa forme finale, après les étapes de reconstruction de l'intensité avec Praat et des caractéristiques spectrales avec STRAIGHT.

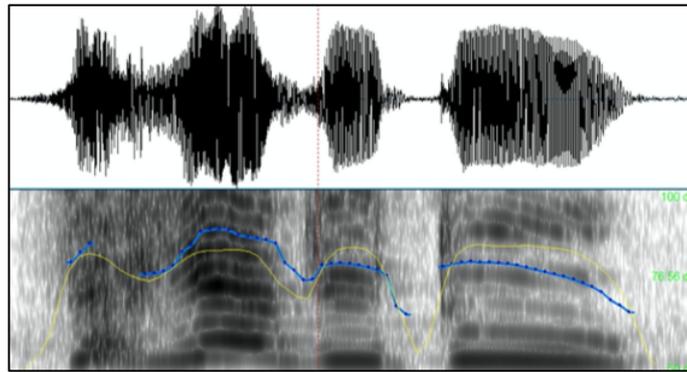


Figure 2 : Spectrogramme et sonogramme du stimulus original (np_caneele_c_1)

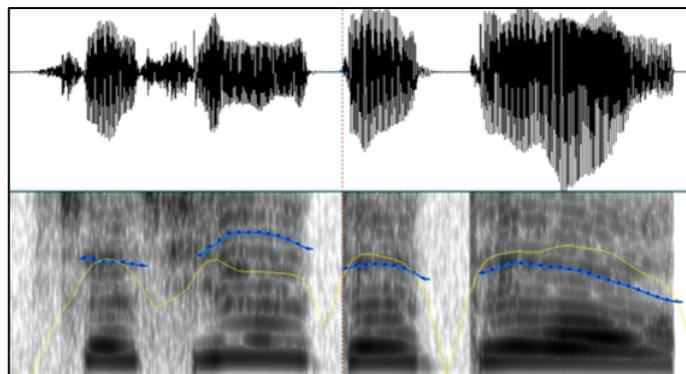


Figure 3 : Spectrogramme et sonogramme du stimulus après reconstruction de la fréquence fondamentale (np_caneele_c_1_delex)

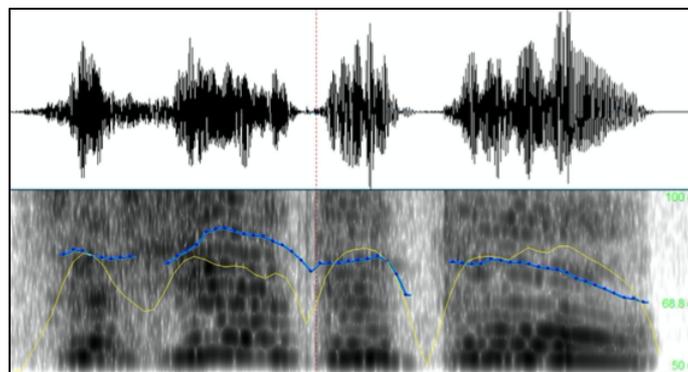


Figure 4 : Spectrogramme et sonogramme du stimulus après reconstruction de l'intensité et des caractéristiques spectrales (np_caneele_c_1_intensity_morphed)

Grâce au travail d'analyses acoustiques, puis d'interprétation et de modification des stimuli sélectionnés, nous pouvons désormais aborder les deux tâches de validation réalisées (intelligibilité et identification émotionnelle) avant de détailler le déroulement de l'évaluation principale de perception émotionnelle.

CHAPITRE 3 : EXPÉRIENCES DE PERCEPTION ÉMOTIONNELLE

Dans un premier temps, deux tâches de validation du matériel ont été réalisées. Le premier test a été mené par nous-mêmes. Il s'agit d'un test d'intelligibilité qui permet de vérifier l'inintelligibilité des stimuli apportée par la délexicalisation. Le deuxième test a été mené par une étudiante en sciences cognitives. Il s'agit d'un test d'identification visant à vérifier que la prosodie des stimuli est représentative de la colère ou de l'état émotionnel neutre. Enfin, seuls les stimuli les moins intelligibles et dont l'émotion a été bien reconnue ont été retenus pour l'épreuve principale d'évaluation émotionnelle à laquelle nous avons contribué. Cette dernière consistait en l'évaluation des dimensions émotionnelles des stimuli (valence et intensité). Toutes ces épreuves ont été validées par le comité d'éthique d'Aix-Marseille Université.

1. TÂCHE DE VALIDATION N°1 : INTELLIGIBILITÉ

1.1. La création du test

Le test d'intelligibilité, que nous avons mené, consistait en la vérification des effets de la délexicalisation sur le contenu verbal des stimuli. Il permettait de préciser la nature et le nombre de mots intelligibles dans chaque énoncé. Pour ce faire, nous avons fait écouter les 73 extraits audio à des participants (évoqués plus bas). Dans un souci de ne pas rendre l'expérience trop contraignante et coûteuse en temps, ces stimuli ont été séparés en deux listes différentes pour deux tests distincts : la liste n°1 comportait 35 stimuli, et la liste n°2 en comportait 38. Chaque test durait environ 10 à 15 minutes. Au vu de la situation sanitaire, il a été décidé que le test d'intelligibilité soit réalisé en ligne ; la plateforme en ligne a été mise en place par Nicolas Audibert, maître de conférence en phonétique au LPP – Laboratoire de Phonétique et Phonologie, Paris 3.

Une fois les tests mis en place, nous avons envoyé un lien d'accès ainsi qu'une notice d'information aux personnes recrutées pour les tests (annexe 4). Avant de commencer l'épreuve, il était nécessaire que tous les participants aient lu et accepté le consentement.

Voici la consigne donnée aux participants : « *Votre participation consiste à entendre des extraits audio de courte durée et modifiés par l'expérimentateur. Après avoir écouté chaque extrait, vous allez retranscrire le contenu verbal de l'énoncé (la phrase entière, une partie ou un seul mot de la phrase) que vous avez entendu. Dans le cas où l'énoncé semble*

complètement inintelligible, vous laisserez le champ vide. Cette tâche se déroulera en ligne. Vous devez être installé dans une pièce calme, assis, face à votre écran d'ordinateur. ».

Concernant le déroulement du test, les participants ne pouvaient écouter qu'une seule fois chaque stimulus afin d'éviter les effets de répétition, les assimilations et inférences qui en découlent. Une fois le stimulus entendu, ils devaient écrire les mots auditivement compréhensibles, s'il y en avait, puis passer au stimulus suivant. Si aucun mot n'était perçu, le participant laissait une réponse vide.

1.2. Les participants recrutés

Plusieurs critères de sélection ont été définis pour le recrutement des participants. Ceux-ci devaient être des hommes ou des femmes adultes âgés de 18 à 55 ans. Leur langue maternelle devait être le français. Ils devaient également n'avoir aucun problème auditif diagnostiqué. Enfin, deux dernières questions leur étaient posées : une concernant leur diplôme le plus élevé obtenu ou pour lequel ils étaient encore en formation, et une deuxième concernant la fréquence d'utilisation de leur langue maternelle. Ces deux questions ne faisaient pas partie des critères d'inclusion mais permettaient de vérifier la diversité dans l'échantillon des participants et de réaliser une description démographique de ce dernier.

Au total, 54 personnes ont participé à ce test d'intelligibilité. 7 d'entre elles ont finalement été exclues car elles ne rentraient pas dans les critères d'inclusion. Finalement, 47 participants, dont 25 avaient répondu à la liste 1 et 22 à la liste 2, ont été retenus pour le traitement des résultats. L'échantillon des participants représentait 38 femmes et 9 hommes avec une moyenne d'âge de 24,7 ans. La plupart des participants étaient des volontaires ou des étudiants devant valider des crédits dans le cadre de leurs études.

1.3. Résultats

Tout d'abord, nous avons défini la nature des mots de chaque énoncé en étiquetant les mots soit comme « mot lexical » (aussi appelé mot de contenu), soit comme « mot grammatical » (aussi appelé mot-outil). Les mots de contenu sont des unités lexicales qui jouent un rôle sémantique dans la phrase (les noms, les adjectifs qualificatifs, les verbes, les adverbes...). Les mots grammaticaux sont, quant à eux, des unités lexicales dont le rôle est principalement syntaxique (les déterminants, les prépositions, les pronoms, les conjonctions...). Une fois cette catégorisation définie, nous avons calculé les pourcentages d'intelligibilité de mots de contenu

par réponse des participants car c'est ceux-là qui sont porteurs de sens et qui peuvent inférer dans la reconnaissance de l'émotion. En moyenne, pour l'ensemble des stimuli, les mots de contenu ont été intelligibles à 25,4% (annexe 5).

Toutefois, nous avons choisi d'utiliser un critère plus strict, c'est-à-dire de considérer le pourcentage d'intelligibilité de mots totaux plutôt que celui de mots de contenu. Ce choix est motivé par le fait que certaines phrases ne présentaient pas de mots de contenu mais contenaient tout de même des mots grammaticaux à connotation émotionnelle. En effet, si nous prenons l'exemple du mot « pas » dans l'énoncé « pas comme ça » : ce mot n'est pas étiqueté comme « mot de contenu » alors même qu'il a une valence négative. Le critère de mots totaux est donc plus strict car il englobe tous les mots, y compris ceux avec une valence négative. Pour cela, pour chacune des réponses obtenues, nous avons effectué un calcul de pourcentage du nombre de mots correctement transcrits en fonction du nombre total de mots du stimulus. Ce calcul a donné un pourcentage d'intelligibilité pour chaque réponse. Nous avons ensuite calculé la moyenne de tous les pourcentages obtenus pour le même stimulus, ce qui nous a permis de déterminer le pourcentage moyen d'intelligibilité de chaque stimulus. Ainsi, 6 stimuli sur les 73 présentés ont été perçus comme totalement inintelligibles (0% d'intelligibilité moyenne) par l'ensemble des participants. En moyenne, tous les stimuli ont été intelligibles à 26,8% (annexe 5).

Pour l'épreuve principale de l'étude, nous avons retenu les stimuli avec un taux d'intelligibilité inférieur à 30%. Ce seuil assez bas, nous permet ainsi de garantir que les stimuli sélectionnés pour l'expérience principale sont suffisamment inintelligibles. Ainsi, 39 stimuli avec un taux d'intelligibilité inférieur ou égal à 30% ont été retenus : 17 correspondaient à des extraits d'acteurs professionnels et 22 à des extraits d'acteurs non-professionnels.

2. TÂCHE DE VALIDATION N°2 : IDENTIFICATION

Nous n'avons pas participé à cette deuxième tâche de validation, qui a été menée en ligne par une étudiante en sciences cognitives dans le cadre du projet COMEDIA présenté au début de l'étude. Cependant, cette tâche d'identification émotionnelle représente un élément important dans la validité du matériel ; c'est pourquoi nous allons l'introduire dans notre mémoire, et en faire une brève description avec une présentation des résultats obtenus. Il s'agit d'un test d'identification visant à vérifier que la prosodie des stimuli est représentative de la colère ou de l'état émotionnel neutre.

Ainsi, l'étudiante en sciences cognitives a recruté 39 participants (21 femmes et 18 hommes) qui ont écouté les 73 stimuli une seule fois et ont dû déterminer l'émotion véhiculée par la

prosodie pour chacun d'entre eux. Pour cela, les participants devaient sélectionner l'émotion perçue, sous présentation des choix alternatifs : « colère » ou « neutre ». Les critères d'inclusion étaient les mêmes que pour la tâche précédente.

L'analyse des résultats a également été réalisée par l'étudiante en sciences cognitives, ce qui nous a permis d'observer plusieurs phénomènes. Pour les deux groupes d'acteurs, nous avons pu relever pour la colère un pourcentage d'identification correcte plus important pour les acteurs professionnels (83,5%) que pour les acteurs non-professionnels (61,9%). Pour l'état émotionnel neutre, nous ne relevons pas de nette différence entre les professionnels (92,4%) et les non-professionnels (93,9%). Ainsi, de manière générale, l'état émotionnel neutre a été mieux reconnu que la colère.

Pour l'épreuve principale de l'étude, nous avons retenu les stimuli avec un taux d'identification émotionnelle supérieur à 70%. Ce seuil élevé nous permet ainsi de garantir que les stimuli sélectionnés pour l'expérience principale véhiculent bien la colère ou l'état émotionnel neutre. Ainsi, 58 stimuli avec un taux d'identification émotionnelle supérieur ou égal à 70% ont été retenus : 34 correspondaient à des extraits d'acteurs professionnels et 24 à des extraits d'acteurs non-professionnels.

3. EXPÉRIENCE PRINCIPALE : ÉVALUATION ÉMOTIONNELLE

Cette dernière étape consistait en une épreuve d'évaluation des dimensions émotionnelles : la valence et l'intensité (arousal). Pour rappel, la valence d'une émotion définit son caractère positif ou négatif, et l'intensité définit son degré d'éveil, d'excitabilité.

3.1. Sélection des stimuli

Pour la réalisation de cette expérience de perception, les stimuli devaient répondre simultanément à deux conditions : avoir un faible taux d'intelligibilité (tâche de validation n°1) ainsi qu'un taux élevé d'identification émotionnelle (tâche de validation n°2).

Ainsi, pour la sélection finale des stimuli, Caterina Petrone et Francesca Carbone ont sélectionné uniquement les stimuli qui avaient à la fois un pourcentage d'intelligibilité inférieur à 30% et une moyenne d'identification émotionnelle supérieure à 70%.

Cependant, ces conditions de sélection n'ont pas permis d'obtenir des catégories équilibrées, c'est-à-dire d'obtenir autant de stimuli d'acteurs professionnels que d'acteurs non-professionnels avec, pour chacun, un stimulus représentant la colère et un stimulus

représentant l'état émotionnel neutre. Ainsi, il a donc fallu sélectionner 5 nouveaux stimuli avec un taux d'intelligibilité supérieur à 30% (3 stimuli d'acteurs professionnels et 2 d'acteurs non-professionnels). Afin de pallier, ce taux plus élevé d'intelligibilité de ces 5 stimuli, une modification du taux de morphing appliqué au moment de la délexicalisation a été réalisée sur ces derniers. Le taux de morphing initial de 0,5 a donc été abaissé à 0,3 pour ces 5 stimuli. Cette modification a permis de rendre les stimuli moins intelligibles, et a donc entraîné un amoindrissement de leurs caractéristiques spectrales.

Ainsi, au total, 20 stimuli ont été retenus pour l'épreuve de perception principale : 10 stimuli d'acteurs professionnels (dont 5 représentent la colère et 5 un état émotionnel neutre), et 10 stimuli d'acteurs non-professionnels (dont 5 représentent la colère et 5 un état émotionnel neutre). Ces stimuli sont repris dans le tableau suivant (tableau 8). Les stimuli en jaune sont ceux qui avaient obtenus un pourcentage d'intelligibilité supérieur à 30% et qui ont été rajoutés pour balancer la sélection des stimuli.

CODE	PHRASE	INTELLIGIBILITÉ (%)	IDENTIFICATION ÉMOTIONNELLE (%)
np_caneele_c_2	Tu comprends pas que c'est un cadeau que je te fais	30,30	95,00
np_caneele_n_1	Dis ce soir on va à la brasserie.	6,82	100,00
np_dequenne_c_4	Rends-le-moi.	21,21	95,00
np_dequenne_n_1	J'ai travaillé.	15,15	100,00
np_toure_c_2	J'te parle.	26,67	89,50
np_toure_n_2	Je veux pas faire de CAP.	14,00	89,50
np_karamoh_c_1	Tu tires dans le tunnel.	25,60	100,00
np_karamoh_n_1	On a qu'à faire un truc qui change.	1,78	94,70
np_ghesquere_c_2	C'est pas ce que j'attends d'une équipe.	84,55	95,00
np_ghesquere_n_1	Faut que je les revoie ces deux-là.	67,61	95,00
p_symon_c_1	Tu le connais ton fils.	58,40	94,70
p_symon_n_1	C'est vraiment pour te faire plaisir.	4,57	84,20
p_stael_c_2	Je suis pas dactylo.	58,00	94,70
p_stael_n_1	Je la récupère après le travail.	3,33	73,70
p_yann_c_1	Pousse-toi.	0,00	84,20
p_yann_n_1	Tu veux pas chercher quelque chose de plus intéressant, non.	52,80	78,90
p_carmet_c_1	J'y toucherai quand j'aurai envie d'y toucher.	0,00	90,00
p_carmet_n_1	J'aimerais m'entretenir quelques instants avec vous.	0,00	95,00
p_perrin_c_1	Je suis pas malade moi.	3,20	94,70
p_perrin_n_2	C'est fermé à clef.	12,80	94,70

Tableau 8 : Les 20 stimuli sélectionnés pour l'évaluation émotionnelle.

Tous les stimuli sont répertoriés selon un code bien précis (tableau 8) :

- La catégorie de l'acteur codée « p » pour professionnel et « np » pour non-professionnel ;
- Le nom de l'acteur ;
- L'émotion exprimée codée « c » pour colère et « n » pour l'état émotionnel neutre ;
- Le dernier chiffre correspond au numéro du stimulus dans la liste initiale des stimuli.

3.2. La création du test

L'objectif de ce test était de préciser les différences prosodiques qu'on peut retrouver entre les groupes d'acteurs professionnels et non-professionnels par le biais d'une évaluation de deux dimensions émotionnelles. La première concernait la valence émotionnelle (émotion perçue comme négative ou positive) ; la seconde concernait l'intensité (émotion perçue comme plus ou moins intense). Pour l'évaluation de ces variables, deux échelles continues, représentées par des barres glissantes, ont été définies : une échelle allant de -100 (pas du tout négative) à +100 (tout à fait négative) pour la valence, et une autre échelle de -100 (très faible) à +100 (très forte) pour l'intensité. L'utilisation d'une échelle continue plutôt que d'une échelle discrète se justifie d'une part par la nécessité de donner aux participants un large choix de réponses, et d'autre part par le souhait d'obtenir des résultats précis pour les analyses statistiques.

Tout comme les autres tests, en raison de la situation sanitaire, l'expérience principale a également été réalisée en ligne. Le test a été créé par Caterina Petrone et Francesca Carbone par l'intermédiaire de Qualtrics, logiciel qui permet de créer des questionnaires (<https://www.qualtrics.com>). Une fois les tests mis en place, nous avons envoyé un lien d'accès ainsi qu'une notice d'information aux personnes recrutées pour les tests (annexe 7). Avant de commencer l'épreuve, il était nécessaire que tous les participants aient lu et accepté le consentement.

Voici la consigne donnée aux participants : « *Ce questionnaire a pour but d'analyser l'expression des émotions par le ton de la voix. Vous allez écouter plusieurs enregistrements de quelques secondes. Il s'agit de phrases plus ou moins modifiées. Pour chaque enregistrement, il vous est demandé de concentrer votre attention uniquement sur le ton de la voix du locuteur. À la fin de chaque phrase, vous devrez évaluer sur deux barres glissantes :*

- *Si l'émotion n'est pas du tout négative ou tout à fait négative (« valence »)*
- *Si l'intensité du ressenti émotionnel véhiculé par le locuteur est très faible ou très forte (« intensité »)*

Vous pouvez positionner le curseur sur tout point de la barre – aux extrêmes, ou dans un point entre les deux extrêmes. Il n'y a pas de bonne ou de mauvaise réponse. Veuillez donc répondre le plus spontanément possible. »

Afin d'augmenter le pouvoir statistique et de minimiser les réponses données au hasard, nous avons fait écouter et évaluer les 20 stimuli de l'épreuve 3 fois ; ils ont ainsi été randomisés.

3.3. Les participants recrutés

Pour cette épreuve, nous avons recruté 90 participants. Une grande majorité des participants étaient des étudiants en orthophonie de l'école de Marseille ou de Paris. Les autres participants étaient des étudiants d'autres domaines ou des proches.

Ces derniers ne devaient n'avoir participé ni au test d'intelligibilité, ni au test d'identification émotionnelle. Les critères d'inclusion étaient les mêmes que pour les tests précédents, c'est-à-dire que les participants devaient avoir entre 18 et 55 ans, leur langue maternelle devait être le français, et ils ne devaient avoir aucun problème auditif diagnostiqué. Après avoir supprimé les tests inachevés et les participants à exclure selon les critères évoqués précédemment, nous avons retenu 57 participants pour l'analyse des résultats. Parmi ces 57 participants, 42 étaient des femmes, 14 étaient des hommes et une dernière personne a répondu « autre » à la question du sexe. La moyenne d'âge des participants était de 30,8 ans.

3.4. Résultats

Ainsi, nous avons relevé les résultats de l'ensemble des réponses des participants. Les stimuli étant écoutés trois fois, nous avons calculé une moyenne des trois résultats obtenus pour chaque variable (valence et intensité).

Les analyses statistiques ont été réalisées par Caterina Petrone. Elles se basent sur des modèles à effets mixtes qui considèrent à la fois les effets fixes (manipulés par l'expérimentateur) et les effets aléatoires (échantillonnés au hasard). Les effets fixes sont l'entraînement des acteurs (professionnels/non-professionnels), le sexe des acteurs (hommes/femmes), l'émotion exprimée (colère/neutre) et leurs interactions. Les effets aléatoires sont les acteurs sélectionnés.

Nous rappelons que l'ensemble des résultats obtenus s'inscrit sur une échelle d'évaluation allant de -100 (pas du tout négative) à +100 (tout à fait négative) pour la valence, et allant de -100 (très faible) à +100 (très forte) pour l'intensité.

Dans cette partie, nous nous attachons à étudier les 20 stimuli sélectionnés pour l'épreuve, dont 15 ont un taux de morphing à 0,5 et 5 ont un taux de morphing à 0,3.

3.4.1. Résultats comparatifs entre les deux groupes d'acteurs

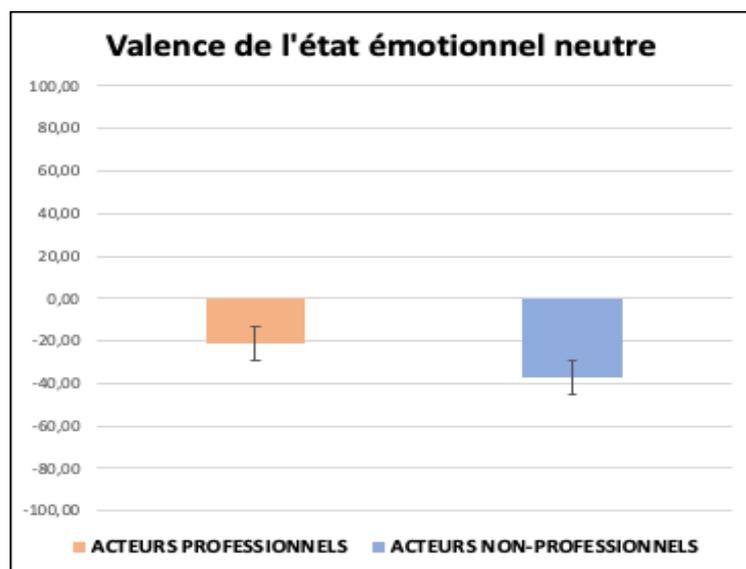
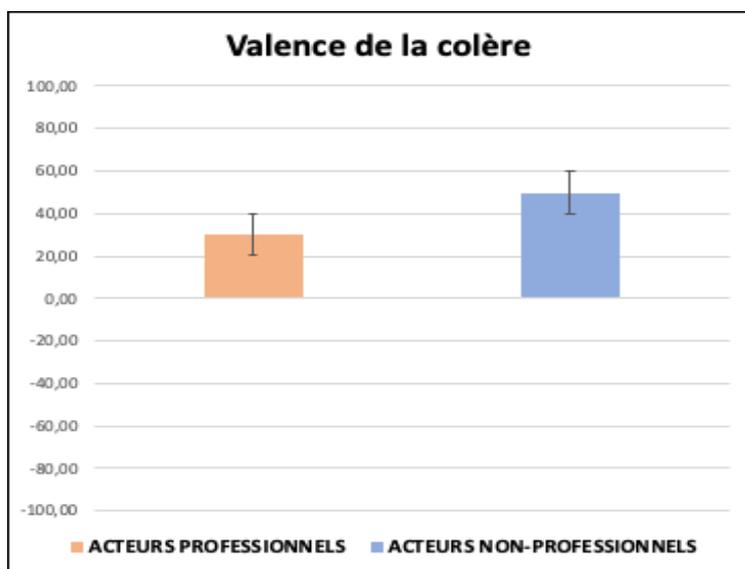
Pour la valence :

Concernant la valence de la colère et de l'état émotionnel neutre, nous avons repris les moyennes obtenues dans deux tableaux (tableaux 9 et 10) :

VALENCE DE LA COLÈRE				VALENCE DE L'ÉTAT ÉMOTIONNEL NEUTRE			
ACTEURS PROFESSIONNELS		ACTEURS NON-PROFESSIONNELS		ACTEURS PROFESSIONNELS		ACTEURS NON-PROFESSIONNELS	
p_symon_c_1	-23,37	np_caneele_c_2	57,61	p_symon_n_1	-16,86	np_caneele_n_1	-32,38
p_stael_c_2	-1,67	np_dequenne_c_4	49,77	p_stael_n_1	-22,77	np_dequenne_n_1	-40,73
p_yann_c_1	57,12	np_toure_c_2	57,12	p_yann_n_1	-25,52	np_toure_n_2	-42,97
p_carmet_c_1	67,68	np_karamoh_c_1	64,13	p_carmet_n_1	-24,31	np_karamoh_n_1	-34,20
p_perrin_c_1	51,16	np_ghesquere_c_2	20,44	p_perrin_n_2	-17,05	np_ghesquere_n_1	-36,48
MOYENNE TOTALE	30,18	MOYENNE TOTALE	49,82	MOYENNE TOTALE	-21,30	MOYENNE TOTALE	-37,35

Tableaux 9 et 10 : Moyennes obtenues pour la valence de la colère et de l'état émotionnel neutre.

Nous avons reporté ces moyennes dans les graphiques suivants (graphiques 1 et 2) :



Graphiques 1 et 2 : Graphiques représentant les moyennes obtenues pour la valence de la colère et de l'état émotionnel neutre.

Pour la valence, nous remarquons que les valeurs moyennes présentent des différences significatives entre les deux groupes d'acteurs.

En se référant à l'échelle d'évaluation de la valence allant de -100 (pas du tout négative) à +100 (tout à fait négative), nous constatons qu'en moyenne, les stimuli d'acteurs non-professionnels représentant la colère sont évalués avec une valence plus négative (49,82) que celle des acteurs professionnels (30,18) (graphique 1). De plus, en moyenne, les stimuli d'acteurs non-professionnels représentant l'état émotionnel neutre sont évalués avec une valence moins négative (-37,35) que celle des acteurs professionnels (-21,30) (graphique 2).

Pour l'intensité :

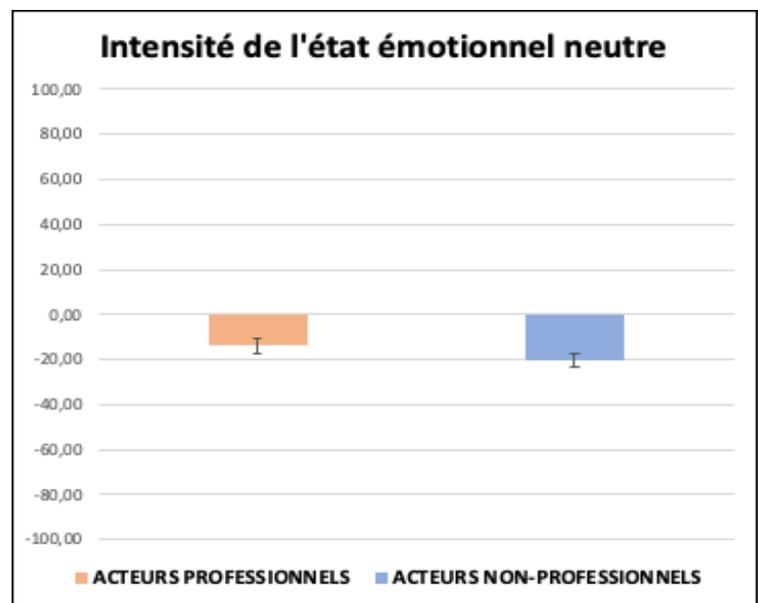
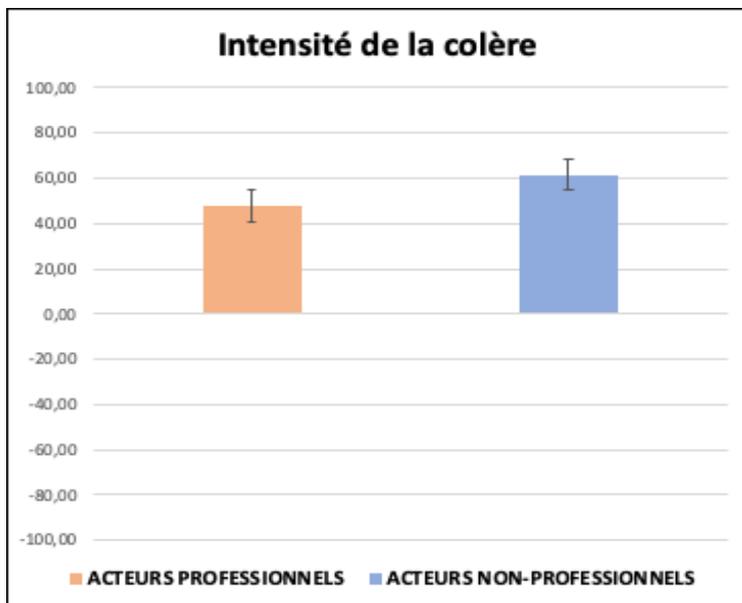
Concernant l'intensité de la colère et de l'état émotionnel neutre, nous avons repris les moyennes obtenues dans deux tableaux (tableaux 11 et 12) :

INTENSITÉ DE LA COLÈRE			
ACTEURS PROFESSIONNELS		ACTEURS NON-PROFESSIONNELS	
p_symon_c_1	25,27	np_caneele_c_2	67,46
p_stael_c_2	8,44	np_dequenne_c_4	69,25
p_yann_c_1	72,50	np_toure_c_2	69,24
p_carmet_c_1	74,77	np_karamoh_c_1	69,14
p_perrin_c_1	57,86	np_ghesquere_c_2	32,25
MOYENNE TOTALE	47,77	MOYENNE TOTALE	61,47

INTENSITÉ DE L'ÉTAT ÉMOTIONNEL NEUTRE			
ACTEURS PROFESSIONNELS		ACTEURS NON-PROFESSIONNELS	
p_symon_n_1	-7,56	np_caneele_n_1	-26,86
p_stael_n_1	-5,64	np_dequenne_n_1	-10,91
p_yann_n_1	-11,25	np_toure_n_2	-22,71
p_carmet_n_1	-20,03	np_karamoh_n_1	-21,44
p_perrin_n_2	-25,26	np_ghesquere_n_1	-19,53
MOYENNE TOTALE	-13,95	MOYENNE TOTALE	-20,29

Tableaux 11 et 12 : Moyennes obtenues pour l'intensité de la colère et de l'état émotionnel neutre.

Nous avons reporté ces moyennes dans les graphiques suivants (graphiques 3 et 4) :



Graphiques 3 et 4 : Graphiques représentant les moyennes obtenues pour l'intensité de la colère et de l'état émotionnel neutre.

Pour l'intensité, nous remarquons que les valeurs moyennes présentent des différences significatives entre les deux groupes d'acteurs.

En se référant à l'échelle d'évaluation de l'intensité allant de -100 (très faible) à +100 (très forte), nous constatons qu'en moyenne, les stimuli d'acteurs non-professionnels représentant la colère sont évalués avec une intensité plus forte (61,47) que celle des acteurs professionnels (47,77) (graphique 3). De plus, en moyenne, les stimuli d'acteurs non-professionnels représentant l'état émotionnel neutre sont évalués avec une intensité plus faible (-20,29) que celle des acteurs professionnels (-13,95) (graphique 4).

Ainsi, que ce soit pour la colère ou l'état émotionnel neutre, l'évaluation des dimensions émotionnelles (valence et intensité) a montré, dans un premier temps, des différences significatives entre les deux groupes d'acteurs. Comme nous l'avons noté précédemment, les acteurs non-professionnels ont obtenu des valeurs moyennes de valence et d'intensité plus élevées que les acteurs professionnels. De plus, nous constatons que cette différence est plus marquée pour les stimuli représentant la colère.

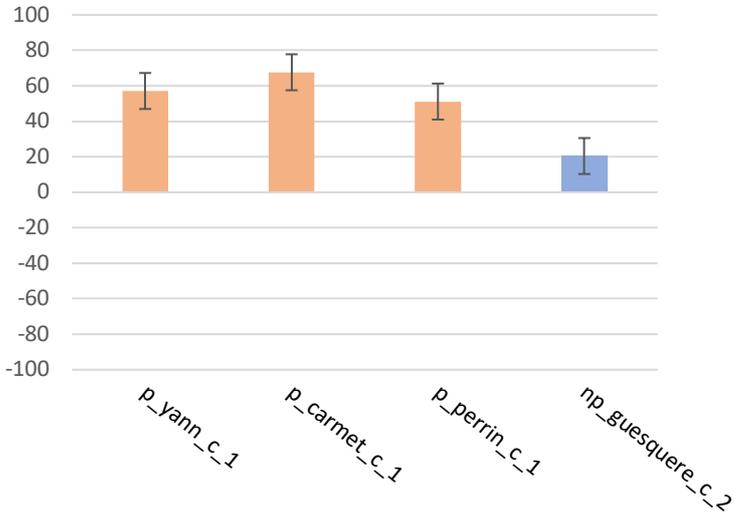
Dans un second temps, ces résultats nous ont également permis de mettre en exergue des différences significatives entre les deux états émotionnels (colère et état émotionnel neutre). De manière générale, la colère a une valence plus négative que l'état émotionnel neutre ($\beta = -91$, $SD = 5.21$, $t = -17.47$, $p < 0.001$) et une intensité plus forte que l'état émotionnel neutre ($\beta = 87.01$, $SD = 4.27$, $t = 20.36$, $p < 0.001$). Ces derniers résultats vont dans le sens de nos hypothèses.

3.4.2. Résultats comparatifs entre les sexes des acteurs

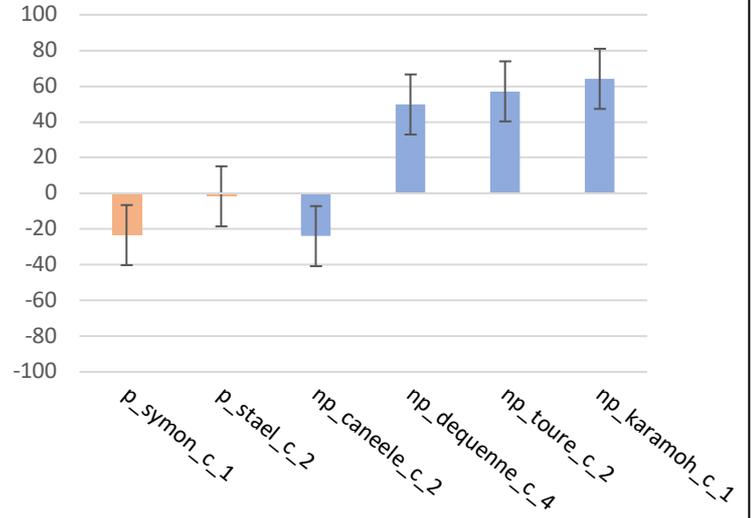
Après avoir étudié les différences entre les deux groupes d'acteurs et les deux états émotionnels, nous nous sommes attachées à la comparaison entre les hommes et les femmes.

Pour la valence :

Valence de la colère pour les hommes

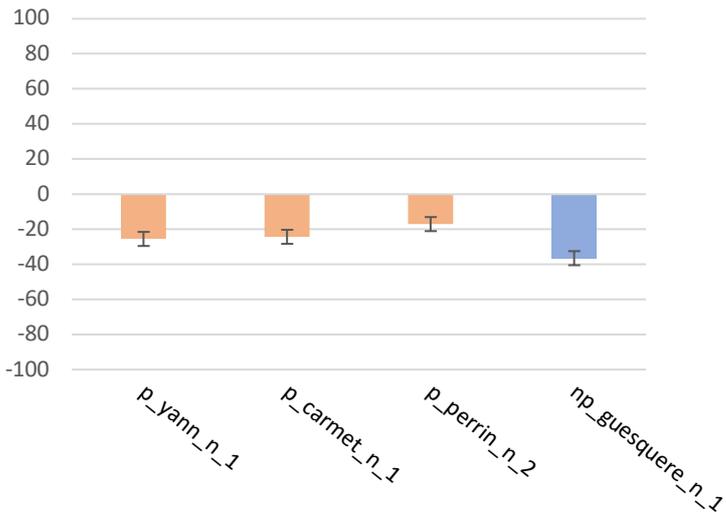


Valence de la colère pour les femmes

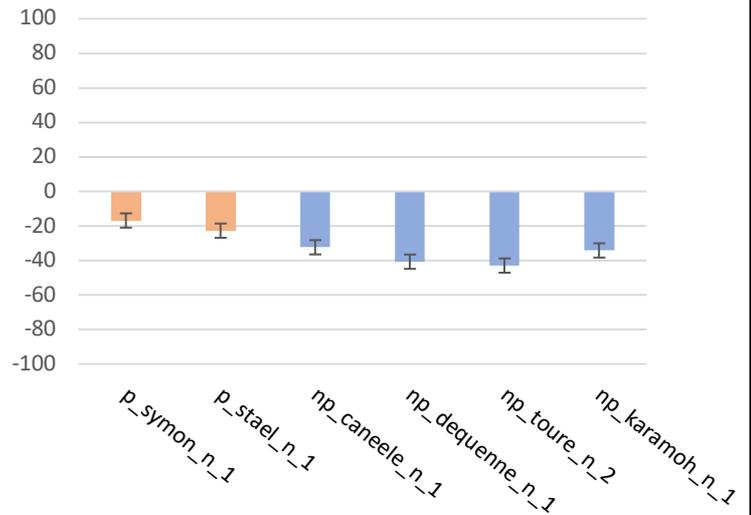


Graphiques 5 et 6 : Graphiques représentant la valence de la colère pour les hommes et pour les femmes.

Valence de l'état émotionnel neutre pour les hommes



Valence de l'état émotionnel neutre pour les femmes



Graphiques 7 et 8 : Graphiques représentant la valence de l'état émotionnel neutre pour les hommes et pour les femmes.

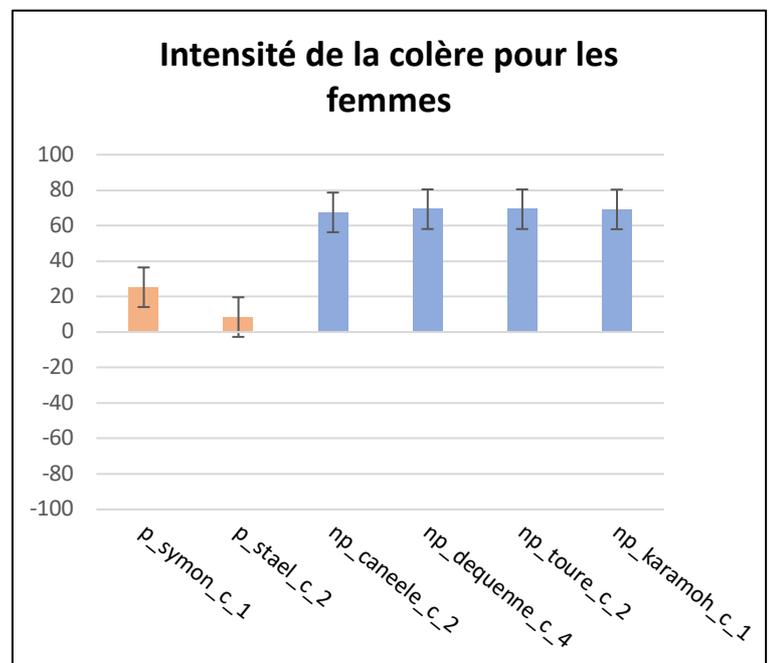
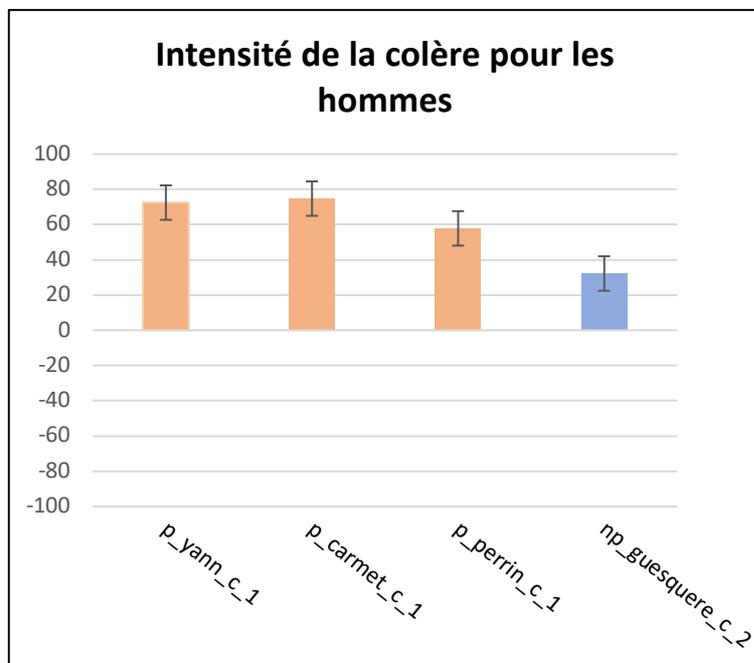
Nous avons vu précédemment que de manière générale, la colère a une valence négative plus importante que l'état émotionnel neutre ($\beta=-91$, $SD=5.21$, $t=-17.47$, $p<0.001$).

Cependant, à travers ces graphiques, nous remarquons que pour les valeurs de valence, on ne retrouve pas de différence significative entre l'évaluation de la colère et du neutre ($p>0.05$) pour les stimuli d'actrices professionnelles uniquement.

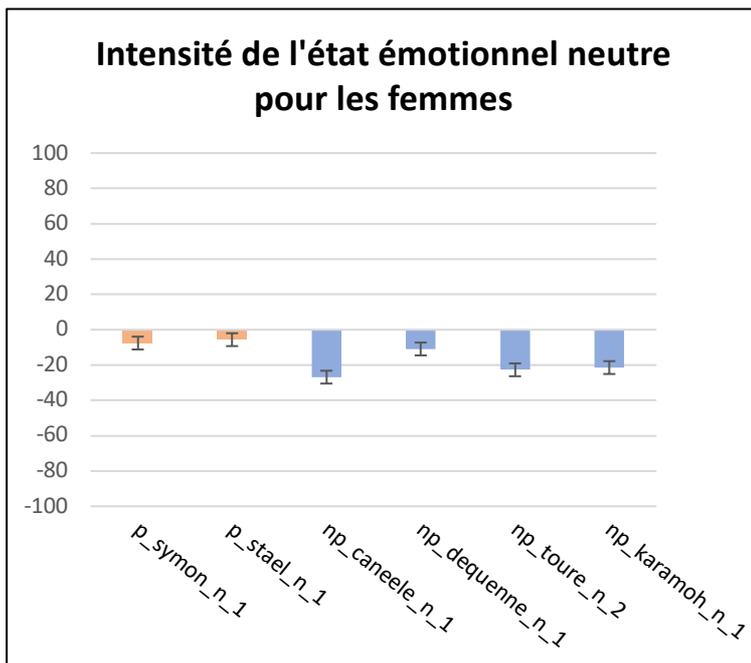
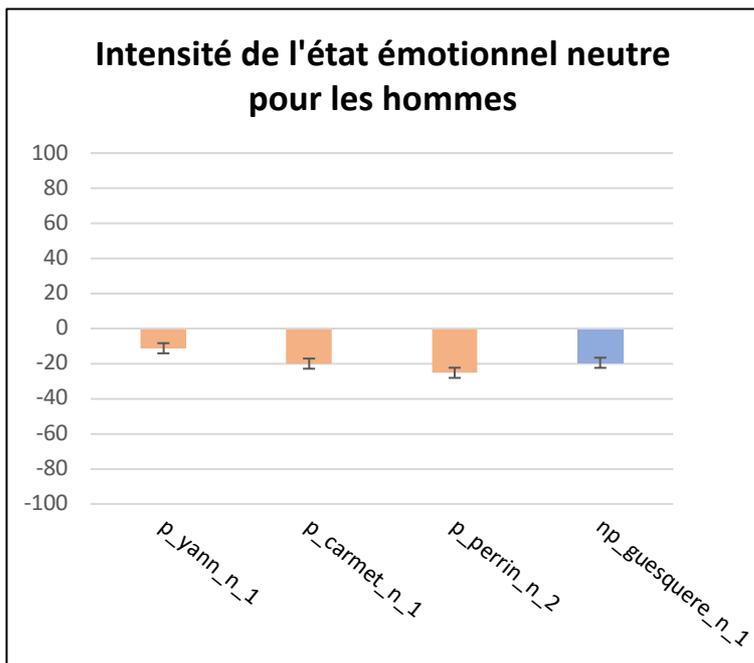
Pour la valence, nous constatons qu'il existe une interaction entre l'expérience cinématographique (professionnel ou non-professionnel) et le sexe de l'acteur. En effet, pour les hommes, l'écart entre colère et état émotionnel neutre est plus important pour les professionnels que pour les non-professionnels ($\beta=43.51$, $SD=6.99$, $t=6.22$, $p<0.001$).

Pour les femmes, c'est l'inverse qui se produit : l'écart entre colère et état émotionnel neutre est plus important pour les non-professionnels que pour les professionnels ($\beta=-69.68$, $SD=4.36$, $t=-15.97$, $p<0.001$).

Pour l'intensité :



Graphiques 9 et 10 : Graphiques représentant l'intensité de la colère pour les hommes et pour les femmes.



Graphiques 11 et 12 : Graphiques représentant l'intensité de l'état émotionnel neutre pour les hommes et pour les femmes.

Nous avons vu précédemment que de manière générale, la colère a une intensité plus importante que l'état émotionnel neutre ($\beta=87.01$, $SD=4.27$, $t=20.36$, $p<0.001$).

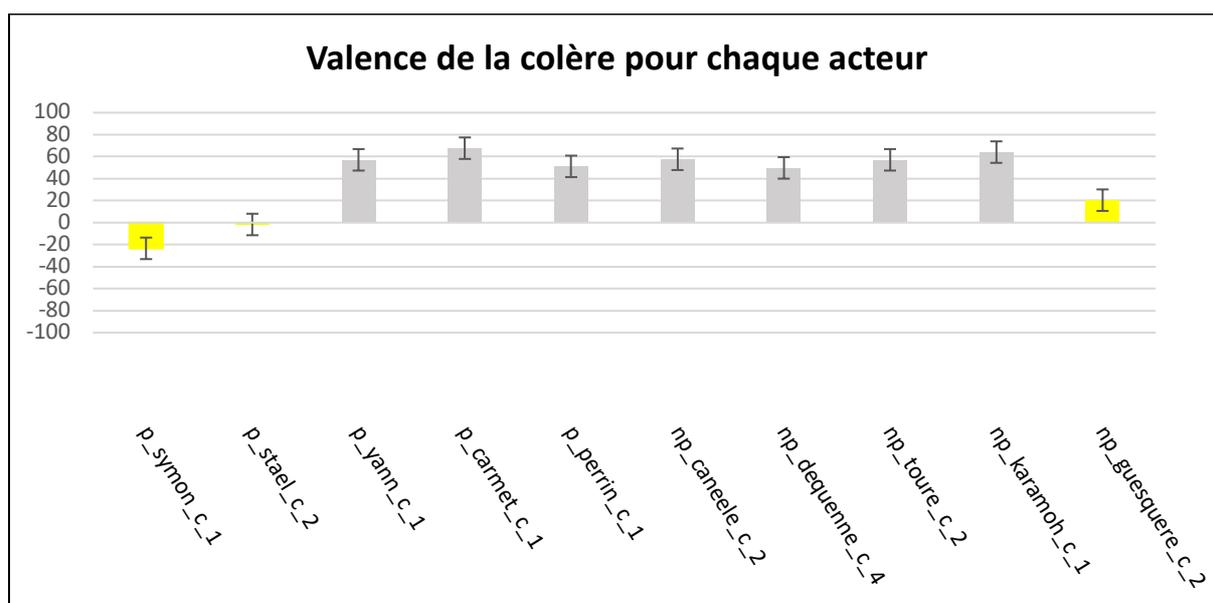
Pour l'intensité, nous constatons qu'il existe une interaction entre l'expérience cinématographique (professionnel ou non-professionnel) et le sexe de l'acteur. En effet, pour les hommes, l'écart entre colère et état émotionnel neutre est de nouveau plus important pour les professionnels que pour les non-professionnels ($\beta=35.44$, $SD=3.95$, $t=-8.97$, $p<0.001$). Pour les femmes, c'est l'inverse qui se produit : l'écart entre colère et état émotionnel neutre est plus important pour les non-professionnels que pour les professionnels ($\beta=-69.68$, $SD=4.36$, $t=-15.97$, $p<0.001$).

Les résultats obtenus dans l'étude des 20 stimuli (avec un taux de morphing à 0,5 ou à 0,3) ont permis de mettre en exergue l'impact de l'entraînement cinématographique, du sexe de l'acteur ou de l'état émotionnel exprimé sur l'évaluation des dimensions émotionnelles.

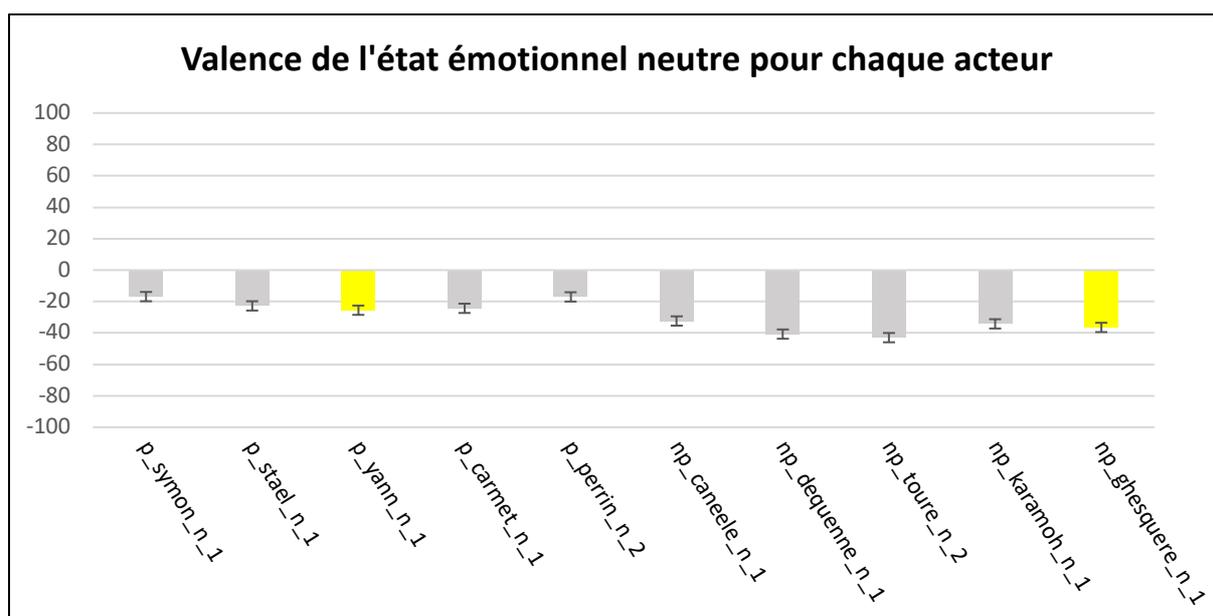
Cependant, il est important de noter la présence d'un facteur de confusion dû au taux de morphing différent pour les 5 stimuli ajoutés pour équilibrer la sélection des stimuli. De ce fait, nous avons décidé de comparer les résultats obtenus pour chaque acteur.

3.4.3. Résultats comparatifs entre tous les acteurs

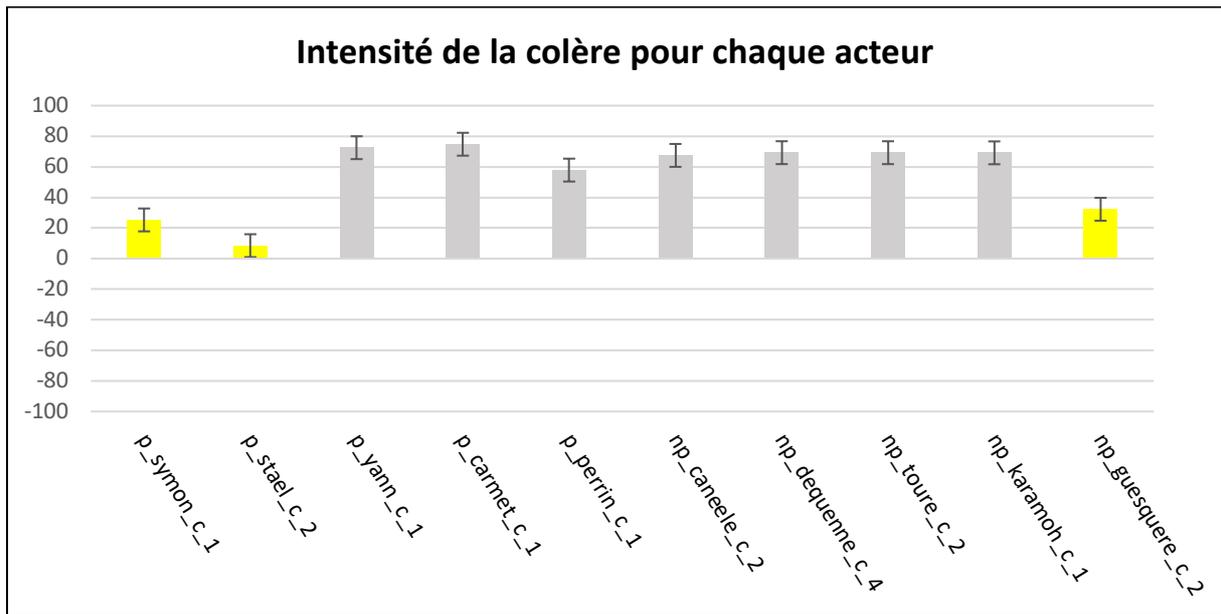
Nous avons reporté les résultats obtenus pour chaque acteur dans des graphiques (graphiques 13, 14, 15 et 16). Sur ces derniers, les 5 stimuli avec un taux de morphing à 0,3 sont représentés en jaune, et les 15 stimuli avec un taux de morphing à 0,5 sont représentés en gris.



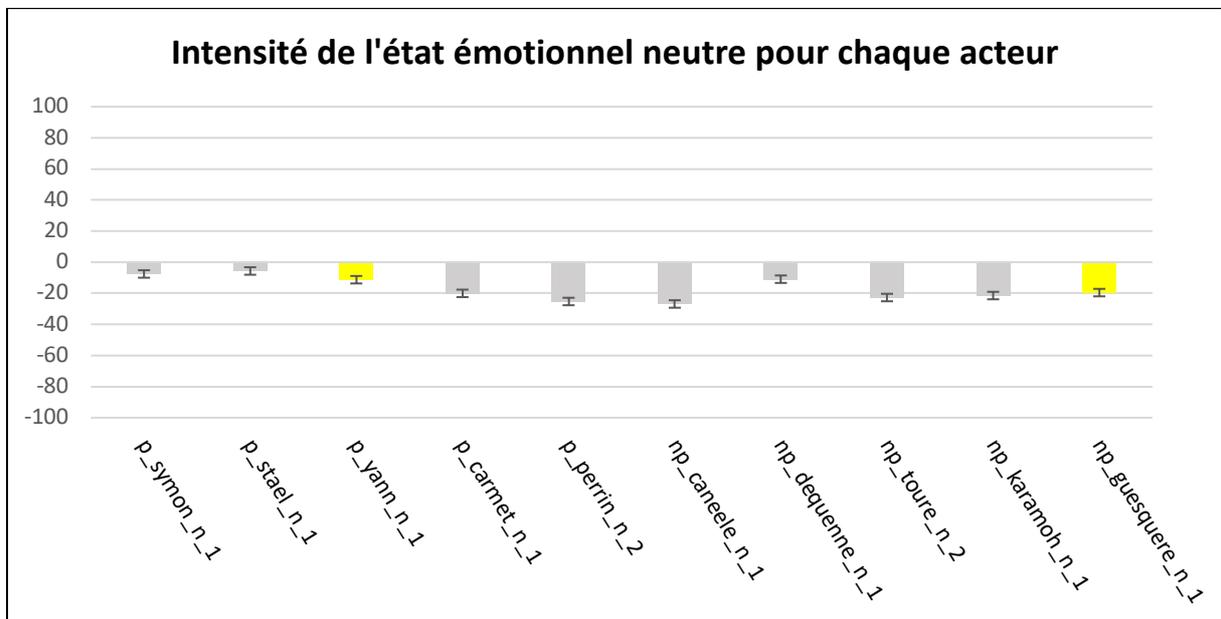
Graphique 13 : Graphique représentant les moyennes obtenues pour la valence de la colère et pour chaque acteur.



Graphique 14 : Graphique représentant les moyennes obtenues pour la valence de l'état émotionnel neutre et pour chaque acteur.



Graphique 15 : Graphique représentant les moyennes obtenues pour l'intensité de la colère et pour chaque acteur.



Graphique 16 : Graphique représentant les moyennes obtenues pour l'intensité de l'état émotionnel neutre et pour chaque acteur.

Ces graphiques attestent bien de la présence d'un facteur de confusion lié au taux de morphing différent (0,3) pour les 5 stimuli (graphiques 13 à 16). En effet, c'est principalement pour 3 de ces stimuli que nous relevons des résultats discordants pour les moyennes de valence et d'intensité ; ce sont les trois qui illustrent la colère.

En se référant à l'échelle d'évaluation de la valence allant de -100 (pas du tout négative) à +100 (tout à fait négative), nous constatons que, parmi ces trois stimuli, deux ont été évalués avec une valence nettement moins négative que tous les autres stimuli de l'épreuve. En effet, ce sont les seuls à avoir obtenu des valeurs négatives sur l'échelle d'évaluation : p_symon_c_1 (-23,37) et p_stael_c_2 (-1,67). Pour le troisième stimulus, la différence est moins nette, bien que ce stimulus ait tout de même été évalué avec une valence moins négative que tous les autres stimuli : np_ghesquere_c_2 (20,44).

En se référant à l'échelle d'évaluation de l'intensité allant de -100 (très faible) à +100 (très forte), nous constatons que les trois stimuli ont été évalués avec une intensité nettement inférieure à celle des autres stimuli de l'épreuve : p_symon_c_1 (25,27), p_stael_c_2 (8,44) et np_ghesquere_c_2 (32,25).

Ainsi, les résultats explicités précédemment suggèrent bien un biais lié au taux de morphing. De ce fait, une deuxième analyse a été réalisée par Caterina Petrone en considérant uniquement les 15 stimuli avec un taux de morphing à 0,5.

Ces nouveaux calculs statistiques permettent d'affirmer que, si l'on considère uniquement les stimuli avec un taux de morphing à 0,5, les différences explicitées précédemment entre les professionnels et les non-professionnels, ainsi qu'entre les hommes et les femmes, disparaissent ($p > 0,05$). En revanche, les différences entre l'évaluation dimensionnelle de la colère et celle de l'état émotionnel neutre persistent : la colère est toujours jugée plus négative ($\beta = 106.35$, $SD = 5.45$, $t = 19.48$, $p < 0.001$) et plus intense que l'état émotionnel neutre ($\beta = 95.29$, $SD = 5.17$, $t = -18.40$, $p < 0.001$).

Finalement, nous pouvons conclure que, si l'on considère uniquement les stimuli avec un taux de morphing à 0,5, on ne retrouve pas d'effet de l'entraînement et du sexe sur la perception des dimensions émotionnelles. Cela signifie que les premières analyses réalisées avec l'ensemble des stimuli sont artificielles. En revanche, seul l'effet des états émotionnels sur la perception des dimensions émotionnelles reste significatif.

4. CONCLUSION

4.1. Interprétation des résultats

Tout au long de notre étude, nous avons tenté de lever le voile sur l'existence potentielle d'un lien entre expérience professionnelle d'un acteur et sa manière d'exprimer une émotion. Plus indirectement, nous avons tenté de vérifier si l'entraînement vocal qu'induit l'expérience cinématographique riche d'un acteur peut influencer sa manière de jouer vocalement une émotion. Nous avons aussi cherché à savoir si cette expérience professionnelle pouvait impacter notre façon de percevoir la prosodie émotionnelle véhiculée par un acteur.

Les résultats globaux de notre étude ne vont pas en faveur de l'existence de ces liens.

En termes de caractéristiques acoustiques, les résultats de notre étude sont conformes avec ceux retrouvés dans la littérature (Scherer, 2003 ; Jürgens et al., 2015) : nous retrouvons une distinction entre les caractéristiques de la prosodie colérique et celles de l'état émotionnel neutre. En effet, nous relevons pour la prosodie de la colère des valeurs de fréquence fondamentale moyenne et d'intensité plus élevées. De plus, pour la colère, nous retrouvons des valeurs acoustiques plus dispersées que pour l'état émotionnel neutre, que ce soit pour les acteurs professionnels ou pour les non-professionnels. En effet les valeurs obtenues pour la déviation standard de la colère sont plus élevées que celles retrouvées pour l'état émotionnel neutre. Cela témoigne d'une palette plus large d'expression de la colère pour les deux groupes d'acteurs. Nous retrouvons notamment cet effet dans de nombreuses études, à travers lesquelles cette grande variabilité des contours acoustiques est liée aux émotions les plus éveillées comme la colère, tandis qu'une faible variabilité est relevée pour les expressions moins éveillées, comme l'état émotionnel neutre. Contrairement aux hypothèses de Jürgens, nos analyses acoustiques n'ont pas révélé l'existence de différences significatives entre acteurs professionnels et non-professionnels (Jürgens et al., 2015).

Par la suite, l'expérience d'évaluation des dimensions émotionnelles nous a permis de confirmer l'interprétation des analyses acoustiques. En effet, la colère est perçue avec une valence plus négative et une intensité plus importante que l'état émotionnel neutre, quel que soit le groupe d'acteur.

Les premières analyses ont mis en évidence, l'existence d'un effet de l'expérience cinématographique, du sexe de l'acteur, ou encore de l'état émotionnel exprimé sur l'évaluation des dimensions émotionnelles (valence et intensité). Cependant, la présence d'un facteur de confusion lié au taux de morphing abaissé à 0,3 a réfuté la validité de ces premières

observations. En effet, de nouvelles analyses ont permis d'affirmer qu'en considérant uniquement les 15 stimuli avec un taux de morphing à 0,5, les effets de l'expérience cinématographique et du sexe de l'acteur disparaissent. Seul l'effet des états émotionnels reste significatif. Ces dernières interprétations nous ont permis de faire le lien entre caractéristiques spectrales et prosodie colérique. En effet, nous avons retrouvé pour les 5 stimuli avec un taux de morphing à 0,3 des résultats s'éloignant de ceux des 15 autres stimuli avec un taux de morphing à 0,5 et du modèle prosodique de la colère. Ainsi, on peut faire le lien entre une reconstruction réduite des caractéristiques spectrales (entraînée par un abaissement du taux de morphing) et une évaluation faussée des dimensions émotionnelles. Ce lien de cause à effet nous permet d'affirmer qu'il est nécessaire de considérer les caractéristiques spectrales comme des indices importants pour décrire la prosodie de la colère ; constatation qui rejoint celle de Scherer en 1984 (Scherer et al., 1984).

Finalement, les résultats globaux de notre étude ne vont pas en faveur de l'hypothèse de Jürgens selon laquelle il serait favorable de recruter des acteurs non-professionnels pour les futures études (Jürgens et al., 2015). En effet, les deux groupes d'acteurs n'ont pour l'instant pas présenté de différences significatives qui nous permettraient d'attester d'une utilisation favorable de stimuli de parole jouée par l'un ou l'autre groupe d'acteurs.

4.2. Limites

Il est important de noter que notre étude souffre de certaines limites.

La première limite concerne les critères de sélection choisis pour les deux groupes d'acteurs (acteurs professionnels et non-professionnels). Nous avons classé les acteurs sans aucune expérience cinématographique dans le groupe des acteurs non-professionnels, et ceux avec au moins une expérience dans le groupe des acteurs professionnels. Ces critères ont permis une délimitation bien précise des deux groupes et une homogénéité du groupe d'acteurs non-professionnels (qui ont tous aucune expérience). Cependant, il est important de noter que le groupe d'acteurs professionnels n'est pas homogène ; en effet, les acteurs de ce groupe ont entre 1 et 157 expériences cinématographiques. De plus, on peut considérer que l'entièreté du groupe d'acteurs non-professionnels et les acteurs professionnels avec le moins d'expériences pourraient se confondre en raison de leur nombre quasi similaire d'expériences cinématographiques.

La deuxième limite concerne le choix de l'état émotionnel neutre comme l'une des deux émotions à étudier. En effet, définir cet état de neutralité est particulièrement difficile du fait de

sa subjectivité. Cependant, l'utilisation dans notre étude de l'état émotionnel neutre comme condition contrôle se justifie par le fait qu'il représente le paradigme le plus utilisé en recherche et répond ainsi à une nécessité de simplification expérimentale.

De plus, il est important de considérer comme limite l'effet de la distance entre le locuteur et le microphone sur les résultats obtenus pour les analyses acoustiques. En effet, le calcul de l'intensité des stimuli est peut-être biaisé par ce phénomène ; il aurait donc été pertinent de mesurer une intensité relative à la distance du microphone.

Un autre aspect également considéré comme une des limites de notre étude concerne l'échantillon de participants recrutés pour les différentes épreuves de perception. En effet, le fait que les participants soient principalement des étudiants, et majoritairement des femmes, tend à rendre l'échantillon trop homogène pour prétendre à une généralisation des résultats. Enfin, la dernière limite concerne les 5 stimuli ajoutés pour l'épreuve principale de perception afin d'équilibrer la sélection finale des 20 stimuli. Comme expliqué précédemment, contrairement aux 15 autres stimuli, ces 5 derniers avaient un taux d'intelligibilité supérieur à 30% et ont alors subi une modification du taux de morphing à 0,3. Ainsi, leurs résultats discordants nous permettent d'affirmer que le maintien d'un unique critère (d'intelligibilité et de taux de morphing) pour la sélection aurait été préférable pour une interprétation des résultats cohérente.

4.3. Ouvertures

Ainsi, notre travail n'a pas permis de mettre en évidence l'existence de l'effet de l'expérience professionnelle d'un acteur sur sa manière d'exprimer la colère et sur notre façon de percevoir la prosodie colérique véhiculée par ces acteurs. Pour les études futures, il serait pertinent de considérer les biais mis en avant dans la partie précédente pour tenter d'approfondir les points qui restent encore imprécis.

Nous avons tout de même tenté d'exploiter une des suggestions de Jürgens qui préconisait de s'intéresser aux caractéristiques individuelles des locuteurs (Jürgens et al., 2015). En effet, nous avons cherché à répondre à la question de l'effet du sexe de l'acteur sur la perception de la prosodie colérique. Cependant, nous n'avons pas pu mettre en avant l'existence de cet effet qui a finalement été faussée par le facteur de confusion lié à la modification du morphing. Pour les prochaines études, ce point pourra être approfondi en uniformisant le taux de morphing appliqué à l'ensemble des stimuli. De plus, il serait pertinent que les prochains travaux puissent s'intéresser davantage aux autres caractéristiques individuelles suggérées

par Jürgens : l'âge de l'acteur, l'utilisation et les particularités de sa voix ainsi que l'humeur de l'acteur.

Aussi, la combinaison des méthodes utilisées par Francesca Carbone pour la délexicalisation a montré des résultats encourageants. Cette méthode de synthèse vocale pourrait être réutilisée et approfondie dans les études à venir, en précisant le taux de morphing idéal pour l'obtention d'un taux correct d'inintelligibilité tout en préservant suffisamment les caractéristiques spectrales.

Par ailleurs, notre étude se concentre sur l'utilisation de l'émotion de la colère. Il pourrait être intéressant de reproduire la même étude avec une autre émotion comme la peur ou la tristesse ; cela nous permettrait de faire une analyse comparative entre les différentes émotions.

Enfin, il serait pertinent de se concentrer cette fois sur les différences interindividuelles des participants aux études de perception émotionnelle. Cela pourrait se réaliser par le biais d'une étude démographique et d'une étude de la personnalité des participants aux épreuves de perception. Ces travaux pourraient nous donner des indices sur la façon que chacun a de percevoir une prosodie émotionnelle selon son milieu socio-culturel, sa sensibilité, sa personnalité, ses habiletés de communication...

BIBLIOGRAPHIE

- Amir, N., Ron, S., Laor, N. (2000). Analysis of an Emotional Speech Corpus in Hebrew. *In Proceedings of ISCA WS on Speech and Emotion*, 29-33.
- Aubergé, V. (2002). A gestalt morphology of prosody directed by functions: the example of a step-by-step model developed at ICP. *1st International Conference Speech Prosody, Aix-en-Provence, France*, 151-155.
- Aubergé, V., Audibert, N., Rilliard, A. (2004). *E-Wiz: A Trapper Protocol for Hunting the Expressive Speech Corpora in Lab*, 179-182.
- Aubergé, V., Cathiard, M. (2003). Can we hear the prosody of smile? *Emotional Speech, Speech Communication Review*, 40, 87-97. <http://www.lrec-conf.org/proceedings/lrec2004/pdf/721.pdf>
- Audibert, N. (2008). Prosodie de la parole expressive : dimensionnalité d'énoncés méthodologiquement contrôlés authentiques et actés. *Institut National Polytechnique de Grenoble - INPG*.
- Bachorowski, J. A., Owren, M. J. (1995). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychol. Sci.* 6(4), 219-224.
- Banse, R., Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70(3), 614-636.
- Boersma, P., Weenink, D. (2021). Praat: doing phonetics by computer [Computer program]. *Version 6.1.48, retrieved 2 June 2021*. From: <http://www.praat.org/>
- Caelen-Haumont, G. (2005). Les états émotionnels et la Prosodie : paradigmes, modèles, paramètres. *Nguyen, Noël. Phonologie et phonétique : Forme et substance, Hermès*, 397-424. Hal- 00142950.
- Clément, J. (1999). Structure des représentations prosodiques. Développement normal et pathologique du traitement de la prosodie, *Thèse doctorale, Université de Paris V*.
- D'Alessandro, C. (2006). Voice source parameters and prosodic analysis. *Methods in Empirical Prosody Research. Berlin, New York: De Gruyter*, 63-87.

- Damasio, A. (1999). *The feeling of what happens: Body, emotion, and consciousness*. New York, Harcourt Brace.
- Darwin, C. (1872). *The Expression of Emotions in Man and Animals*. John Murray, London.
- Drolet, N., Schubotz, R., Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: behavioral and fMRI evidence. *Cognitive, Affective, & Behavioral Neuroscience*, 12, 140-150. <https://link.springer.com/article/10.3758/s13415-011-0069-3>
- Douglas-Cowie, E., Campbell, N., Cowie, R., Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*, 40, 33-60.
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., Vrecken, O. V. (1996). The MBROLA project: towards a set of high-quality speech synthesizers free of use for non-commercial purposes. *In ICSLP-1996*, 1393-1396.
- Ekman, P. (1971). Are There Basic Emotions? *Psychological Review*, 99(3), 550-553. DOI: 10.1037/0033-295X.99.3.550
- Ekman, P., Davidson R. J. (1994). *The nature of emotion: Fundamental questions*. New York: Oxford University Press.
- Ekman, P. (1999). Basic Emotion. In T. Dalgleish and T. Power (Eds.) *The Handbook of Cognition and Emotion*, Sussex, R.U.: John Wiley & Sons, Ltd., 45-60.
- Enos, F., Hirschberg, J. (2008). A framework for Eliciting Emotional Speech: Capitalizing on the Actor's Process, 6-10.
- Fletcher, H. (1953). *Speech and Hearing in Communication*. New York: Kreiger.
- Frijda, N. H. (1987). Emotions, Cognitive structures and Action tendency. *Cognition and Emotion*, 1, 115-143.
- Gobl, C., Chasaide, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication* (40), 189–212.

Granarolo, P. (2021). ÉMOTION (notions de base), *Encyclopædia Universalis*.

<https://www.universalis.fr/encyclopedie/emotion>

Hawk, S. T., Van Kleef, G. A., Fischer, A. H., Van Der Schalk, J. (2009). Worth a thousand words: Absolute and relative decoding of nonlinguistic affect vocalizations. *Emotion*, 9(3), 293–305. <http://doi.org/10.1037/a0015178>

James, W. (1884). What is an emotion? *Mind*, 9, 188-205.

Jürgens, R., Grass, A., Drolet, M., Fischer, J. (2015). Effect of Acting Experience on Emotion Expression and Recognition in Voice: Non-Actors Provide Better Stimuli than Expected. *Journal of Nonverbal Behavior*, 39, 195-214. <https://link.springer.com/article/10.1007/s10919-015-0209-5>

Karlsson, I., Bänziger, T., Dankovicova, J., Johnstone, T., Lindberg, J., Melin, H., Nolan, F., Scherer, K.R. (2000). Speaker verification with elicited speaking styles in the VeriVox project. *Speech Communication* 31 (2–3), 121-129.

Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoust. Sci. & Tech.* 27, 6.

Kracauer, S. (2005). Remarks on the actor. In K. Knopf (Ed.), *Theater and film: A Comparative anthology*, 323–333. New York: Yale University Press.

Kramer, E. (1964). Elimination of verbal cues in judgments of emotion from voice. *The Journal of Abnormal and Social Psychology*, 68(4), 390–396. <https://doi.org/10.1037/h0042473>

Lacheret, A., Beaugendre, F. (1999). La prosodie du français, *Paris : éditions du CNRS*, 11-12.

Ladd, D. R. (2008). Intonational phonology. *Functions of Language*, 17(2). DOI:10.1075/fo1.17.2.08ogr

Laver, J. (1980). The phonetic description of voice quality. *Cambridge University Press, New York*.

Le Breton, D. (1998). Les passions ordinaires. *Anthropologie des émotions, Armand Colin/Masson, Paris*, 31.

- LeDoux, J. (1996). Emotional networks and motor control: a fearful view. *Progress in Brain Research, 107*, 437-446.
- Leinonen, L., Hiltunen, M. L. (1997). Expression of emotional-motivational connotations with one-word utterance. *JASA*.
- Licklider, J. C. R., Miller, G. A. (1951). The perception of speech. In S. S. Stevens (Ed.), *Handbook of experimental psychology, Wiley*, 1040-1074.
- Murray, L. R., Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J. Acoust. Soc. Am, (2)*, 1097-1108.
- Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*.
- Ramus, F., Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America, 105(1)*, 512-521.
- Rottenberg, J., Ray, R., Gross, J. J. (2007). Emotion Elicitation Using Films. *Psychology Faculty Publications, 9-28*.
- Russell, J. A., Mehrabian, A. (1977). Evidence for a Tree-Factor Theory of Emotions. *Journal of Research in Personality*.
- Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology, 39(6)*, 1161.
- Sauter, D. A., Eisner, F., Ekman, P., Scott, S. K. (2010). Emotional Vocalizations Are Recognized Across Cultures Regardless of the Valence of Distractors. *Psychol Sci., 2015 Mar; 26(3)*, 354-356. DOI: 10.1177/0956797614560771
- Scheiner, E., Fischer, J. (2011). Emotion expression: The evolutionary heritage in the human voice. In W. Welsch (Ed.), *Interdisciplinary anthropology: The continuing evolution of man. Heidelberg, Germany: Springer*.
- Scherer, K. R., Ladd, R. D., Silverman, K. E. A. (1984). Vocal cues to speaker affect : testing two models. *Journal of the Acoustical Society of America 76*, 1346–1356.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin* 99, 143–165.

Scherer, K. R., Johnstone, T., Sangsue, J. (1998). L'état émotionnel du locuteur : facteur négligé mais non négligeable pour la technologie de la parole. *Actes des XXIIIèmes Journées d'Etudes sur la Parole, Martigny*.

Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Communication*, 40, 227-256.

Stanislavski, K. (1971). *An Actor Prepares*, Hapgood, E.R., trans., *Theater Art Books, New York*.

Starkweather, J. A. (1956). Content-free speech as a source of information about the speaker. *J. Pers. Soc. Psychol.* 35, 345-350.

Stemmler, G., Heldmann, M., Pauls, C. A., Scherer, T. (2001). Constraints for emotion specificity in fear and anger: The context counts. *Psychophysiology* 38(2), 275-291.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy*, 6, 473-482.

Williams, C. E., Stevens, K. N. (1972). Emotions and speech: some acoustical correlates. *JASA*, 52, 4(2), 1238-1250.

Wundt, W. (1874). *Grundzüge der physiologischen Psychologie, Fundamentals of physiological psychology, Leipzig*.

SITES CONSULTÉS :

Encyclopedia Universalis. (2021). Emotion, *Dictionnaire Cordial*. Consulté à l'adresse : https://www.universalis.fr/dictionnaire/?q=émotion&btn_search=Rechercher

ANNEXES

Annexe n°1 : Les 38 stimuli représentant la colère.

STIMULI REPRÉSENTANT LA COLÈRE		
FILM	ACTEUR	EXTRAIT
L'année dernière à Marienbad	Giorgio Albertazzi	« Ce n'est pas vrai »
L'année dernière à Marienbad	Giorgio Albertazzi	« Ce n'est pas vrai que nous ayons besoin de l'argent »
La Reine des pommes	Valérie Donzelli	« Tu me crois pas »
La Reine des pommes	Béatrice de Staël	« Du calme »
La Reine des pommes	Béatrice de Staël	« Je suis pas dactylo »
La Reine des pommes	Béatrice de Staël	« Et j'ai autre chose à faire »
Buffet froid	Jean Rougerie	« Des idées de quoi ? »
Buffet froid	Jean Carmet	« J'y toucherai quand j'aurai envie d'y toucher »
Buffet froid	Jean Carmet	« J'te parle »
Buffet froid	Jean Carmet	« Vous la laissez toute seule dans une tour vide »
Buffet froid	Bernard Blier	« Ça suffit »
Buffet froid	Bernard Blier	« J'ai pas fini »
Buffet froid	Bernard Blier	« Pas moyen de se reposer cinq minutes »
Buffet froid	Bernard Blier	« Exécution »
Buffet froid	Marco Perrin	« Je suis pas malade moi »
Buffet froid	Michel Fortin	« Mais je suis locataire ? »
Juste la fin du monde	Léa Seydoux	« Ce n'est pas la question »
Nous ne vieillirons pas ensemble	Jean Yanne	« Pousse-toi »

Rosetta	Émilie Dequenne	« Rends-le-moi »
Une étrange affaire	Nathalie Baye	« Mais c'est quoi la question ? »
Une étrange affaire	Gérard Lanvin	« Arrête »
La graine et le mulet	Farida Benkhetache	« On est tous au courant »
La graine et le mulet	Farida Benkhetache	« Vas-y arrête »
La graine et le mulet	Rosalie Symon	« Tu le connais ton fils »
La Drôlesse	Claude Hébert	« Je suis pas un voleur »
La Drôlesse	Claude Hébert	« Elle est pas à vous »
La Drôlesse	Claude Hébert	« Elle est à mon père »
Bande de filles	Karidja Touré	« Tu rentres »
Bande de filles	Karidja Touré	« J'te parle »
Bande de filles	Assa Sylla	« J'te parle là »
Bande de filles	Assa Sylla	« T'as peur de moi ou quoi là ? »
Bande de filles	Simina Soumaré	« Tu tires dans le tunnel »
Bande de filles	Simina Soumaré	« Y'a des règles »
L'Humanité	Ghislain Ghesquère	« Et y'a rien qui avance »
L'Humanité	Ghislain Ghesquère	« C'est pas ce que j'attends d'une équipe »
L'Humanité	Emmanuel Schotté	« Pas comme ça »
L'Humanité	Séverine Caneele	« C'est rien du tout »
L'Humanité	Séverine Caneele	« Tu comprends pas que c'est un cadeau que je te fais ? »
Shéhérazade	Dylan Robert	« Tu comprends quoi ? »

Annexe n°2 : Les 35 stimuli représentant un état émotionnel neutre.

STIMULI REPRÉSENTANT UN ÉTAT ÉMOTIONNEL NEUTRE		
FILM	ACTEUR	EXTRAIT
L'année dernière à Marienbad	Giorgio Albertazzi	« Et si c'était à vous de jouer le premier ? »
La Reine des pommes	Valérie Donzelli	« Je crois que je vais aller me coucher »
La Reine des pommes	Béatrice de Staël	« Je la récupère après le travail »
La Reine des pommes	Béatrice de Staël	« Tu n'as pas de téléphone portable »
Buffet froid	Jean Rougerie	« Pas vraiment »
Buffet froid	Jean Carmet	« J'aimerais m'entretenir quelques instants avec vous »
Buffet froid	Jean Carmet	« Vous boirez pas encore un petit verre ? »
Buffet froid	Jean Carmet	« 500 mètres »
Buffet froid	Bernard Blier	« Excusez-les monsieur »
Buffet froid	Bernard Blier	« En général j'arrive trop tard »
Buffet froid	Marco Perrin	« C'est fermé à clef »
Buffet froid	Marco Perrin	« Une grande artiste »
Buffet froid	Marco Perrin	« Elle jouait dans le trio des Marais »
Juste la fin du monde	Léa Seydoux	« Ça fait presque plus effet »
Nous ne vieillirons pas ensemble	Jean Yanne	« Tu veux pas chercher quelque chose de plus intéressant, non ? »
Rosetta	Émilie Dequenne	« J'ai travaillé »
Une étrange affaire	Nathalie Baye	« T'avais envie de le voir »
Une étrange affaire	Gérard Lanvin	« Tu veux pas que je descende Filo ? »
La graine et le mulet	Farida Benkhetache	« Tu manges avec nous ou quoi ? »

La graine et le mulet	Rosalie Symon	« C'est vraiment pour te faire plaisir »
La Drôlesse	Claude Hébert	« Tu viens on va faire une balade »
La Drôlesse	Claude Hébert	« C'est mon beau-père »
Bande de filles	Karidja Touré	« Laisse-moi au moins finir ma partie »
Bande de filles	Karidja Touré	« Je veux pas faire de CAP »
Bande de filles	Karidja Touré	« C'est une copine qui me l'a filé »
Bande de filles	Assa Sylla	« Faut que tu fasses ce que tu veux »
Bande de filles	Simina Soumaré	« On a qu'à faire un truc qui change »
L'Humanité	Ghislain Ghesquère	« Il faut que je les revoie ces deux-là »
L'Humanité	Ghislain Ghesquère	« J'irai les voir chez elles, ce sera moins dur »
L'Humanité	Emmanuel Schotté	« La porte, elle était ouverte »
L'Humanité	Séverine Caneele	« Dis, on sort on va à la brasserie »
L'Humanité	Séverine Caneele	« La grève elle est finie »
Shéhérazade	Dylan Robert	« T'es un musulman déjà »
Shéhérazade	Dylan Robert	« Y en a pour 180 euros là »
Shéhérazade	Dylan Robert	« Tu prends la moitié et je garde la moitié »

Annexe n°3 : Les analyses acoustiques des stimuli.

ACTEURS PROFESSIONNELS – colère.

CODE	f0 MINIMUM (Hz)	f0 MAXIMUM (Hz)	MOYENNE f0 (Hz)	f0 RANGE (Hz)	INTENSITÉ (dB)	DURÉE (sec)
p_albertazzi_c_1	117	157	131	40	78,17	0,636
p_albertazzi_c_2	80	126	102	46	74,81	1,491
p_baye_c_1	121	485	397	364	72,65	1,254
p_blier_c_1	208	449	334	241	77,00	0,672
p_blier_c_2	369	499	451	130	76,67	1,169
p_blier_c_3	167	455	311	288	76,35	1,227
p_blier_c_4	239	437	384	198	77,89	0,807
p_carmet_c_1	163	335	287	172	78,25	1,424
p_carmet_c_2	228	499	310	271	74,51	1,276
p_carmet_c_3	296	409	363	113	74,52	1,487
p_dequenne_c_4	386	604	466	218	73,70	0,591
p_donzelli_c_1	317	508	426	191	67,39	0,705
p_fortin_c_1	221	445	364	224	76,30	0,946
p_lanvin_c_1	163	460	389	297	73,82	1,136
p_perrin_c_1	204	477	291	273	76,38	0,915
p_rougerie_c_1	215	306	275	91	72,50	0,675
p_seydoux_c_1	248	453	352	205	63,29	0,824
p_stael_c_1	78	363	247	285	69,30	0,520
p_stael_c_2	190	447	303	257	67,03	0,885
p_stael_c_3	170	427	295	257	68,74	1,106
p_yann_c_1	150	350	306	200	69,59	0,617

ACTEURS PROFESSIONNELS – état émotionnel neutre.

CODE	f0 MINIMUM (Hz)	f0 MAXIMUM (Hz)	MOYENNE f0 (Hz)	f0 RANGE (Hz)	INTENSITÉ (dB)	DURÉE (sec)
p_albertazzi_n_1	79	180	106	101	71,99	1,754
p_baye_n_1	222	348	256	126	70,07	0,709
p_blier_n_1	75	106	89	31	66,01	1,136
p_blier_n_2	77	108	91	31	66,87	1,165
p_carmet_n_1	90	356	103	266	56,27	1,909
p_carmet_n_3	125	211	179	86	73,45	1,276
p_carmet_n_4	171	198	186	27	60,69	0,689
p_dequenne_n_1	189	279	222	90	64,91	0,594
p_donzelli_n_1	105	493	209	388	68,77	1,015
p_fortin_n_1	122	264	178	142	66,92	0,728
p_fortin_n_2	142	288	195	146	72,90	1,038
p_lanvin_n_1	120	197	154	77	67,31	1,003
p_perrin_n_2	99	121	112	22	76,86	0,893
p_rougerie_n_1	161	259	191	98	73,31	0,488
p_seydoux_n_1	167	228	193	61	53,95	0,738
p_stael_n_1	132	215	165	83	63,89	1,147
p_stael_n_2	143	197	171	54	64,76	1,383
p_yann_n_1	70	390	113	320	66,93	1,774

ACTEURS NON-PROFESSIONNELS – colère.

CODE	f0 MINIMUM (Hz)	f0 MAXIMUM (Hz)	MOYENNE f0 (Hz)	f0 RANGE (Hz)	INTENSITÉ (dB)	DURÉE (sec)
np_benkhetcha_c_1	276	350	304	74	63,41	0,697
np_benkhetcha_c_2	296	520	429	224	67,43	0,728
np_caneele_c_1	233	487	379	254	81,47	0,819
np_caneele_c_2	232	510	405	278	78,38	1,365
np_ghesquere_c_1	208	327	262	119	74,22	0,833
np_ghesquere_c_2	239	397	291	158	75,56	1,144
np_hebert_c_1	240	381	333	141	81,87	0,653
np_hebert_c_2	204	383	278	179	80,53	0,457
np_hebert_c_3	248	292	274	44	78,21	0,709
np_karamoh_c_1	164	473	336	309	69,65	0,800
np_karamoh_c_2	218	494	358	276	70,31	0,638
np_robert_c_1	163	259	202	96	78,77	0,554
np_schotte_c_1	188	288	245	100	74,59	0,776
np_sylla_c_1	140	404	268	264	68,93	0,553
np_sylla_c_2	179	457	299	278	71,67	0,893
np_symon_c_1	320	577	417	257	63,00	0,712
np_toure_c_1	280	312	296	32	66,55	0,509
np_toure_c_2	174	462	375	288	69,58	0,617

ACTEURS NON-PROFESSIONNELS – état émotionnel neutre.

CODE	f0 MINIMUM (Hz)	f0 MAXIMUM (Hz)	MOYENNE f0 (Hz)	f0 RANGE (Hz)	INTENSITÉ (dB)	DURÉE (sec)
np_benkhetcha_n_1	105	341	256	236	66,79	0,852
np_caneele_n_1	167	255	221	88	82,50	1,277
np_caneele_n_3	165	230	199	65	84,18	0,743
np_ghesquere_n_1	112	199	132	87	69,17	1,097
np_ghesquere_n_2	104	203	129	99	68,40	1,707
np_hebert_n_1	152	490	229	338	75,45	1,053
np_hebert_n_2	123	151	142	28	67,78	0,712
np_karamoh_n_1	183	269	218	86	65,52	0,833
np_robert_n_1	127	265	192	138	72,03	0,856
np_robert_n_2	89	153	130	64	70,75	1,284
np_robert_n_3	103	157	123	54	66,33	1,781
np_schotte_n_1	80	178	114	98	75,15	1,119
np_sylla_n_1	105	221	162	116	62,11	1,100
np_symon_n_1	155	342	233	187	61,34	1,001
np_toure_n_1	173	293	209	120	62,77	1,061
np_toure_n_2	158	235	194	77	61,66	0,896
np_toure_n_3	82	251	159	169	61,76	0,924

Notice d'information

Étude de la prosodie émotionnelle chez les acteurs

Madame, Monsieur,

L'investigatrice principale, **Petrone, Caterina**, vous propose de participer au protocole de recherche comportemental intitulé : « **Étude de la prosodie émotionnelle chez les acteurs** ».

Nous vous proposons de lire attentivement cette notice d'information qui a pour but de répondre aux questions que vous seriez susceptible de vous poser avant de prendre votre décision de participation.

Vous pourrez vous adresser à l'investigateur Francesca Carbone pour lui poser toutes les questions complémentaires (carbonefrancesca@hotmail.it)

Objectif de la recherche :

L'expérience à laquelle il vous est proposé de participer s'inscrit dans un projet de recherche du Laboratoire Parole et Langage d'Aix-en-Provence. Ce projet vise à étudier l'impact des expressions vocales des émotions produites par les acteurs sur les auditeurs. Les détails des hypothèses et des résultats attendus pourront vous être précisés après l'expérience, comme c'est l'usage dans les recherches en linguistique.

Qui peut participer ?

Les hommes et femmes adultes entre 18 et 55 ans et de langue maternelle française.

Quelle est la méthodologie et comment se déroule l'expérimentation ?

Votre participation consiste à entendre des extraits audio de courte durée et modifiés par l'expérimentateur. Après avoir écouté chaque extrait, vous allez retranscrire le contenu verbal de l'énoncé (la phrase entière, une partie ou un seul mot de la phrase) que vous avez entendu. Dans le cas où l'énoncé semble complètement inintelligible, vous laisserez le champ vide.

Cette tâche se déroulera en ligne. Vous devez être installé dans une pièce calme, assis, face à votre écran d'ordinateur.

Vous pourrez y lire les consignes et instructions, la question sur chaque audio, et vous pourrez répondre à cette question.

L'expérience dure environ 10-15 minutes.

Quelles sont les contraintes et désagréments ?

Cette expérience ne comporte pas de contraintes ou désagréments.

Votre participation à cette étude ne comporte aucun risque.

Quels sont vos droits en tant que participant(e) à cette recherche ?

Vous pouvez refuser de participer à cette recherche sans avoir à vous justifier. De même vous pouvez vous retirer à tout moment de l'essai sans justification, sans conséquence sur la suite de votre traitement ni la qualité des soins qui vous seront fournis.

Un traitement de vos données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de l'objectif de cette dernière qui vous a été présenté. Ces données seront identifiées par un numéro de code et vos initiales. Ces données pourront également, dans des conditions assurant leur confidentialité, être transmises aux autorités de santé françaises ou étrangères, à d'autres entités du CNRS. Conformément aux dispositions de la loi relative à l'informatique, aux fichiers et aux libertés et au Règlement Européen 2016/679 du 27 avril 2016 (Règlement Général sur la Protection des Données Personnelles ou « RGDP »), entré en vigueur le 25 mai 2018, vous disposez d'un droit d'accès et de rectification. Vous disposez également d'un droit d'opposition à la transmission des données couvertes par le secret professionnel susceptibles d'être utilisées dans le cadre de cette recherche et d'être traitées. Les données recueillies lors de cette recherche demeureront strictement confidentielles.

La confidentialité de votre dossier sera maintenue conformément à la loi. Seules les données de cette expérience pourront être utilisées dans le cadre d'une recherche et pourront être publiées. Votre nom n'apparaîtra jamais, ces données seront utilisées de manière totalement anonyme.

Vous allez accepter le traitement informatisé des données nominatives qui vous concernent en conformité avec les dispositions de la loi n°2004-801 du 6 août 2004 relative à la protection des personnes et la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Vous pouvez exercer votre droit d'accès et de rectification garanti par les articles 39 et 40 de cette loi en vous adressant auprès du responsable de l'étude.

Vous pouvez poser toutes les questions que vous voulez et vous pouvez disposer d'un temps de réflexion suffisant entre l'information et votre décision de participer à cette étude. Nous vous remercions de considérer notre demande de participation à cette étude.
Ce document vous appartient.

L'investigatrice principale de cette étude est la Chargée de Recherche CNRS Caterina Petrone. Cette étude est réalisée par le laboratoire Parole et Langage.

Annexe n°5 : Résultats du test d'intelligibilité.

CODE	PHRASE	Moyenne mots contenu (en %)	Moyenne mots totaux (en %)
np_benketache_c_1	On est tous au courant.	0,00	0,00
np_benketache_c_2	Vas-y arrête.	38,00	46,67
np_benketache_n_1	Tu manges avec nous toi ?	32,00	32,80
np_caneele_c_1	C'est rien du tout	X	45,45
np_caneele_c_2	Tu comprends pas que c'est un cadeau que je te fais.	24,24	30,30
np_caneele_n_1	Dis ce soir on va à la brasserie.	2,27	6,82
np_caneele_n_3	La grève elle est finie.	9,09	10,00
np_ghesquere_c_1	Y a rien qui avance.	9,09	1,82
np_ghesquere_c_2	C'est pas ce que j'attends d'une équipe.	79,55	84,55
np_ghesquere_n_1	Faut que je les revoie ces deux-là.	75,00	67,61
np_guesquere_n_2	J'irai les voir chez elles ce sera moins dur.	0,00	0,00
np_hebert_c_1	Pas un voleur.	22,73	9,09
np_hebert_c_2	Elle est pas à vous.	X	8,18
np_hebert_c_3	Elle est à mon père.	9,09	4,55
np_hebert_n_1	Viens on va faire une balade.	60,23	62,88
np_hebert_n_2	C'est mon beau père.	4,55	8,18
np_karamoh_c_1	Tu tires dans le tunnel.	22,00	25,60
np_karamoh_c_2	Y a des règles.	0,00	0,00
np_karamoh_n_1	On a qu'à faire un truc qui change.	1,33	1,78
np_robert_c_1	Tu comprends quoi ?	0,00	6,06
np_robert_n_1	T'es un musulman déjà	6,82	3,64
np_robert_n_2	Y en a pour cent-quatre vingts euros là.	2,73	1,52
np_robert_n_3	Tu prends la moitié et je garde la moitié.	0,00	0,00
np_schotte_c_1	Pas comme ça.	X	59,09
np_schotte_n_1	La porte elle était ouverte.	8,00	7,20
np_sylla_c_1	Je te parle là	68,00	64,00
np_sylla_c_2	T'as peur de moi ou quoi là.	18,00	13,50
np_sylla_n_1	Faut que tu fasses ce que tu veux.	10,67	8,50
p_symon_c_1	Tu le connais ton fils.	60,00	58,40
p_symon_n_1	C'est vraiment pour te faire plaisir.	0,00	4,57
np_toure_c_1	Tu rentres.	0,00	6,00
np_toure_c_2	J'te parle.	24,00	26,67
np_toure_n_1	Laisse-moi au moins finir ma partie.	38,00	32,57
np_toure_n_2	Je veux pas faire de CAP.	5,33	14,00
np_toure_n_3	C'est une copine qui me le file.	16,00	28,50
p_albertazzi_c_2	Ce n'est pas vrai que nous ayons besoin de l'argent.	36,36	43,18
p_albertazzi_n_1	Et si c'était à vous de jouer le premier.	75,00	80,00
p_baye_c_1	Mais c'est quoi la question.	100,00	97,73
p_baye_n_1	T'avais envie de le voir.	6,82	14,39
p_blier_c_1	Ça suffit.	22,73	27,27

p_blier_c_2	J'ai pas fini.	9,09	9,09
p_blier_c_3	Pas moyen de se reposer cinq minutes.	2,27	1,30
p_blier_c_4	Exécution.	36,36	36,36
p_blier_n_1	Excusez-les monsieur.	65,91	56,06
p_blier_n_2	En général j'arrive trop tard.	70,45	74,24
p_carmet_c_1	J'y toucherai quand j'aurai envie d'y toucher.	0,00	0,00
p_carmet_c_2	Prenez vos responsabilités.	22,73	21,21
p_carmet_c_3	Vous la laissez toute seule dans une tour vide.	29,55	34,34
p_carmet_n_1	J'aimerais m'entretenir quelques instants avec vous.	0,00	0,00
p_carmet_n_3	Boirait pas encore un petit verre ?	3,41	7,58
p_carmet_n_4	Cinq-cents mètres.	36,36	36,36
p_dequenne_c_4	Rends-le-moi.	4,55	21,21
p_dequenne_n_1	J'ai travaillé.	13,64	15,15
p_donzelli_c_1	Tu me crois pas.	68,18	71,59
p_donzelli_n_1	Je crois que je vais aller me coucher.	4,55	5,68
p_fortin_c_1	Mais suis locataire	4,00	13,33
p_fortin_n_1	Une grande artiste.	70,00	61,33
p_fortin_n_2	Elle jouait dans le trio des Marais.	0,00	0,00
p_lanvin_c_1	Arrête	20,00	20,00
p_lanvin_n_1	Tu veux pas je descende Filo ?	1,33	4,00
p_perrin_c_1	Je suis pas malade moi.	8,00	3,20
p_perrin_n_2	C'est fermé à clef.	14,00	12,80
p_rougerie_c_1	Des idées de quoi ?	12,00	9,00
p_rougerie_n_1	Pas vraiment.	0,00	2,00
p_seydoux_c_1	Ce n'est pas la question.	100,00	94,67
p_seydoux_n_1	Ça fait presque plus effet.	8,00	11,20
p_stael_c_1	Du calme.	88,00	72,00
p_stael_c_2	Je suis pas dactylo.	28,00	58,00
p_stael_c_3	Et j'ai autre chose à faire.	20,00	18,29
p_stael_n_1	Je la récupère après le travail.	2,00	3,33
p_stael_n_2	Tu n'as pas de téléphone portable.	96,00	88,57
p_yann_c_1	Pousse-toi.	0,00	0,00
p_yann_n_1	Tu veux pas chercher quelque chose de plus intéressant, non.	54,40	52,80

Annexe n°6 : Résultats du test d'identification émotionnelle.

CODE	PROSODIE	Pourcentage identification correcte
np_karamoh_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
np_karamoh_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	52,60
np_karamoh_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	94,70
np_benkhetache_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	21,10
np_benkhetache_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	68,40
np_benkhetache_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
np_schotte_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	94,70
np_sylla_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	89,50
np_sylla_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	42,10
np_sylla_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	89,50
np_symon_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	94,70
np_symon_n_1_intensity_combined.wav	Neutre	84,20
np_toure_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	26,30
np_toure_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	89,50
np_toure_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	84,20
np_toure_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	89,50
np_toure_n_3_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_fortin_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	84,20
p_fortin_n_1_intensity_combined.wav	Neutre	100,00
p_fortin_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	52,60
p_lanvin_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	89,50
p_lanvin_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	78,90
p_perrin_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	94,70
p_perrin_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	94,70
p_rougerie_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	73,70
p_rougerie_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_seydoux_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_seydoux_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_stael_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	47,40
p_stael_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	94,70
p_stael_c_3_delex_intensity_morphed_ta_log_0_5.wav	Colère	36,80
p_stael_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	73,70
p_stael_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	94,70
p_yann_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	84,20
p_yann_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	78,90
np_robert_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	50,00
np_robert_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	80,00
np_robert_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
np_robert_n_3_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
np_schotte_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	30,00

np_caneele_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	40,00
np_caneele_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	95,00
np_caneele_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
np_caneele_n_3_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
np_ghesquere_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	40,00
np_ghesquere_c_2_intensity_combined.wav	Colère	95,00
np_ghesquere_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	95,00
np_ghesquere_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	90,00
np_hebert_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	85,00
np_hebert_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	45,00
np_hebert_c_3_delex_intensity_morphed_ta_log_0_5.wav	Colère	50,00
np_hebert_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	95,00
np_hebert_n_2_delex_intensity_morphed_ta_log_0.5.wav	Neutre	100,00
p_albertazzi_c_2_intensity_combined.wav	Colère	10,00
p_albertazzi_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_baye_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_baye_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_blier_c_1_intensity_combined.wav	Colère	75,00
p_blier_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_blier_c_3_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_blier_c_4_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_blier_n_1_intensity_combined.wav	Neutre	100,00
p_blier_n_2_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_carmet_c_1_delex_intensity_morphed_ta_log_0_5.wav	Colère	90,00
p_carmet_c_2_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_carmet_c_3_delex_intensity_morphed_ta_log_0_5.wav	Colère	100,00
p_carmet_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	95,00
p_carmet_n_3_delex_intensity_morphed_ta_log_0_5.wav	Neutre	95,00
p_carmet_n_4_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_dequenne_c_4_delex_intensity_morphed_ta_log_0_5.wav	Colère	95,00
p_dequenne_n_1_delex_intensity_morphed_ta_log_0_5.wav	Neutre	100,00
p_donzelli_c_1_delex_intensity_morphed_ta_log_0.5.wav	Colère	95,00
p_donzelli_n_1_intensity_combined.wav	Neutre	100,00

Notice d'information

Étude de la prosodie émotionnelle produite par les acteurs

Madame, Monsieur,

L'investigatrice principale, **Petrone, Caterina**, vous a proposé de participer au protocole de recherche comportementale intitulé : « **Études de la prosodie émotionnelle produite par les acteurs** ».

Nous vous proposons de lire attentivement cette notice d'information qui a pour but de répondre aux questions que vous seriez susceptible de vous poser avant de prendre votre décision de participation.

Vous pourrez durant l'essai vous adresser à l'investigateur Francesca Carbone (carbonefrancesca@hotmail.it) pour lui poser toutes les questions complémentaires.

Objectif de la recherche :

L'expérience à laquelle il vous est proposé de participer s'inscrit dans un projet de recherche du Laboratoire Parole et Langage d'Aix-en-Provence. Ce projet vise à étudier l'impact des expressions vocales des émotions produites par les acteurs sur les auditeurs.

La confidentialité de votre dossier sera maintenue conformément à la loi. Seules les données de cette expérience pourront être utilisées dans le cadre d'une recherche et pourront être publiées. Votre nom n'apparaîtra jamais, ces données seront utilisées de manière totalement anonyme.

Les détails des hypothèses et des résultats attendus pourront vous être précisés après l'expérience, comme c'est l'usage dans les recherches en linguistique.

Qui peut participer ?

Les hommes et femmes adultes entre 18 et 55 ans et de langue maternelle française.

Quelle est la méthodologie et comment se déroule l'expérimentation ?

Votre participation consiste à entendre une série d'extraits audio de courte durée et sans contenu verbal, dans lesquels les acteurs expriment des émotions uniquement par le ton de la voix. Après avoir écouté chaque extrait, vous allez évaluer (1) la valence de chaque son sur une barre glissante allant de « pas du tout négative » à « tout à fait négative » ; (2) l'intensité émotionnelle (arousal) de chacun son sur une barre glissante allant de « très faible » à « très forte ».

Cette tâche se déroulera en ligne. Vous devez être installé dans une pièce calme, assis, face à votre écran d'ordinateur.

Vous pourrez y lire les consignes et instructions, les questions sur chaque audio, et vous pourrez répondre à ces questions.

Votre participation à cette étude ne comporte aucun risque.

Quelles sont les contraintes et désagréments ?

Cette expérience ne comporte pas de contraintes ou désagréments.
Votre participation à cette étude ne comporte aucun risque.

Quels sont vos droits en tant que participant(e) à cette recherche ?

Vous pouvez refuser de participer à cette recherche sans avoir à vous justifier. De même vous pouvez vous retirer à tout moment de l'essai sans justification, sans conséquence sur la suite de votre traitement ni la qualité des soins qui vous seront fournis.

Un traitement de vos données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de l'objectif de cette dernière qui vous a été présenté. Ces données seront identifiées par un numéro de code et vos initiales. Ces données pourront également, dans des conditions assurant leur confidentialité, être transmises aux autorités de santé françaises ou étrangères, à d'autres entités du CNRS. Conformément aux dispositions de la loi relative à l'informatique, aux fichiers et aux libertés et au Règlement Européen 2016/679 du 27 avril 2016 (Règlement Général sur la Protection des Données Personnelles ou « RGDP »), entré en vigueur le 25 mai 2018, vous disposez d'un droit d'accès et de rectification. Vous disposez également d'un droit d'opposition à la transmission des données couvertes par le secret professionnel susceptibles d'être utilisées dans le cadre de cette recherche et d'être traitées. Les données recueillies lors de cette recherche demeureront strictement confidentielles.

La confidentialité de votre dossier sera maintenue conformément à la loi. Seules les données de cette expérience pourront être utilisées dans le cadre d'une recherche et pourront être publiées. Votre nom n'apparaîtra jamais, ces données seront utilisées de manière totalement anonyme.

Vous allez accepter le traitement informatisé des données nominatives qui vous concernent en conformité avec les dispositions de la loi n°2004-801 du 6 août 2004 relative à la protection des personnes et la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Vous pouvez exercer votre droit d'accès et de rectification garanti par les articles 39 et 40 de cette loi en vous adressant auprès du responsable de l'étude.

Vous pouvez poser toutes les questions que vous voulez et vous pouvez disposer d'un temps de réflexion suffisant entre l'information et votre décision de participer à cette étude. Nous vous remercions de considérer notre demande de participation à cette étude.
Ce document vous appartient.

**L'investigatrice principale de cette étude est la Chargée de Recherche CNRS Caterina Petrone.
Cette étude est réalisée par le laboratoire Parole et Langage.**

RÉSUMÉ

Ce mémoire s'inscrit dans la recherche scientifique sur la prosodie émotionnelle. Il entre dans le cadre du projet COMEDIA mené par Caterina Petrone et Maud Champagne-Lavau, et impliquant plusieurs chercheurs et étudiants de différents domaines. À l'origine, les études scientifiques portant sur les émotions, qui ont cherché à comparer la parole authentique et la parole jouée, ont rencontré un succès mitigé en raison de nombreuses limites éthiques. Par la suite, les études cliniques réalisées en laboratoire se sont concentrées sur les émotions produites par les acteurs. Les scientifiques ont alors tenté de comparer le jeu des acteurs et des non-acteurs et ont soulevé le questionnement suivant : existe-t-il un lien entre l'expérience professionnelle d'un acteur et sa manière d'exprimer une émotion ?

C'est sur la base de ces observations que nous avons décidé d'utiliser, pour notre projet, des groupes d'acteurs professionnels ayant une grande expérience cinématographique, et des acteurs non-professionnels débutant dans le domaine du cinéma. Au travers de cette étude, nous tentons d'explicitier et de quantifier les différences entre les deux groupes d'acteurs à travers des extraits de films exprimant deux états émotionnels bien distincts : la colère (qui est une des six émotions de base) et l'état émotionnel neutre (qui est notre condition contrôle). Pour cela, après avoir réalisé un état des lieux de la littérature scientifique, un corpus de 73 stimuli a été constitué, dont 14 représentent les voix d'acteurs professionnels et 9 d'acteurs non-professionnels. Cet échantillon nous a servi de support pour la réalisation d'analyses acoustiques et l'élaboration d'une tâche de validation de l'intelligibilité, avant de procéder à une épreuve d'évaluation des dimensions émotionnelles. En corrélant l'ensemble de nos résultats obtenus, nous avons cherché à explorer et à élucider la question de l'impact de l'entraînement vocal sur l'expression et la perception de la prosodie émotionnelle.

MOTS-CLÉS :

Orthophonie, recherche, émotions, prosodie, voix, acteurs professionnels, acteurs non-professionnels, colère, état émotionnel neutre, valence, intensité, expression émotionnelle, perception émotionnelle, analyses acoustiques, épreuves de perception.