

Université de Montréal

**Est-ce que la reconnaissance de la prosodie émotionnelle dans la langue française
est modulée par les accents régionaux français et québécois ?**

Par

Amélie Rolinat

Département de psychologie

Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de Maîtrise ès sciences (M. Sc.) en
psychologie

Août 2023

© Amélie Rolinat, 2023

Université de Montréal

Ce mémoire intitulé

**Est-ce que la reconnaissance de la prosodie émotionnelle dans la langue française
est modulée par les accents régionaux français et québécois ?**

Présenté par

Amélie Rolinat

A été évalué(e) par un jury composé des personnes suivantes

Karim Jerbi

Président

Simon Rigoulot

Directeur de recherche

Alexandru Hanganu.

Membre du jury

Résumé

Contexte : La prosodie de la parole, c'est-à-dire les variations du ton de la voix lorsque l'on parle, joue un rôle clé dans les interactions sociales en apportant entre autres des informations importantes liées à l'identité, l'état émotionnel ou encore l'origine géographique. La prosodie est modifiée par les accents d'une personne, en particulier si elle parle une langue étrangère. Ces accents ont un impact important sur la façon dont la parole est reconnue, avec des conséquences significatives sur la façon dont le locuteur est perçu socialement, comme une baisse d'empathie ou encore une moins grande confiance. Cependant, il est moins clair si cet impact, généralement négatif, persiste dans le contexte des accents régionaux qui constituent des variations plus subtiles du signal vocal. **Objectif et hypothèse** : L'objectif de ce présent mémoire est de comprendre comment des individus francophones de différentes régions (France, Québec) expriment et reconnaissent des phrases émotionnelles prononcées par des personnes originaires de la même région ou non. Plusieurs études suggèrent un avantage de groupe, qui renvoie à l'idée que même si les émotions pourraient être reconnues de manière universelle, nous reconnaissons mieux les productions émotionnelles de personnes de notre propre groupe culturel que de personnes extérieures à ce groupe. Est-ce que cet avantage persiste dans le cas des accents régionaux, pour lesquels deux populations partagent la même langue ? Cette question reste très peu étudiée et ne l'a jamais été avec la langue française. Nous souhaitons 1) créer et valider une banque de phrases émotionnelles prononcées en français avec des accents de France et du Québec ; 2) caractériser les profils acoustiques de ces productions émotionnelles. Sur la base de données de la littérature (e.g., Mauchand et Pell, 2020), nous nous attendons à ce que les québécois (Qc) montrent une prosodie émotionnelle plus expressive que les Français (Fr). **Méthode** : Nous avons créé de courtes phrases émotionnelles dans 5 émotions (Joie, Tristesse, Colère, Fierté, Honte), prononcées par des acteurs quebecois.es et français. Cette banque de stimuli a été validée avec une étude en ligne par des françaises et québécoises.

Avec un modèle général mixte, nous avons analysé les paramètres vocaux: moyenne et l'écart type de la fréquence fondamentale, l'écart-type et la moyenne de l'intensité, Shimmer moyen, Jitter, HNR, l'indice Hammarberg, pente spectrale et durée des phrases. **Résultats** : Les paramètres de la fréquence fondamentale moyenne (F0M), d'intensité, de durée, de pente spectrale et d'indice Hammarberg sont significativement différents selon les émotions et entre les origines, Nous avons aussi noté une interaction entre les sexes des locuteurs et leurs origines. De manière globale, sur les cinq émotions considérées, les Fr parlent avec une F0M plus élevée, sauf pour la tristesse. Les Qc parlent eux, pour toutes les émotions, avec une plus grande intensité et une plus longue durée. Au final, nous pouvons considérer que les Qc expriment de manière plus prononcée les émotions que les Fr, sauf au niveau de la colère. En ce qui concerne les différences liées au sexe des locuteurs, nous avons remarqué que les hommes Qc ont une prosodie émotionnelle plus forte que les hommes Fr. Des différences entre les femmes Qc et Fr ont été seulement observées dans les émotions de honte et de fierté, des émotions plus sociales. **Conclusion** : Nous avons pu caractériser l'expression vocale émotionnelle des Fr et Qc qui, malgré leur langue commune, s'expriment de manière très distincte pour transmettre leurs émotions. Ces résultats ouvrent des perspectives intéressantes sur les interactions interculturelles d'une même langue mais de régions différentes et confirment la prosodie de langage, en particulier émotionnelle, comme un véritable marqueur identitaire.

Mots clés : Prosodie émotionnelle, accents régionaux, paramètres vocaux

Abstract

Context: The prosody of speech, i.e. the variations in tone of voice when speaking, plays a key role in social interactions by providing important information linked to identity, emotional state and geographical origin. Foreign accents have a major impact on the way speech is recognized, with significant consequences for social evaluations, such as reduced empathy towards the speaker. However, it is less clear whether this impact persists in the context of regional accents, which are more subtle variations of the speech signal. **Objective and hypothesis:** The aim of this dissertation is to understand how French-speaking individuals from different regions (France, Quebec) express and recognize emotional phrases spoken by people from the same or different regions. Given the hypothesis of group advantage, we believe that expression and perception differ according to culture. Based on Mauchand and Pell's (2020) study, we expect Quebecers (Qc) to show a more expressive emotional prosody than French people (Fr). **Method:** We created short emotional sentences in 5 emotions (Joy, Sadness, Anger, Pride, Shame), spoken by Quebecois and French actors. This bank of stimuli was validated with an online study by French and Quebecers. Using a general mixed model, we analyzed the vocal parameters: mean and standard deviation of the fundamental frequency, standard deviation and mean of the intensity, mean Shimmer, Jitter, HNR, Hammarberg index, Spectral Slope, Duration. **Results:** The parameters of fundamental frequency (F0M), intensity, duration, spectral slope and Hammarberg index differed significantly between the emotions, origins and sexes. For example, of the five emotions, Frs spoke with a higher F0M except for sadness, but Qc spoke with greater intensity and longer duration. The Qc expressed the emotions in a more pronounced way than the Fr, except for anger. Also, many significant differences show that Qc men have a stronger emotional prosody than Fr men. Finally, only differences between Qc and Fr women were observed in the emotions of shame and pride, emotions that are not in the 6 primary emotions, but which

would be more cultural emotions. **Conclusion:** We were able to characterize the emotional vocal expression of Fr and Qc who, despite their common language, express themselves in very distinct ways to convey their emotions. These results open up interesting perspectives on intercultural interactions in the same language but in different regions and show that speech prosody, particularly emotional prosody, could be considered as a marker of regional identity.

Key words: Emotional prosody, regional accents, vocal parameters

Table des matières

Résumé	3
Abstract	5
Remerciements	11
Listes des tableaux.....	12
Liste des figures	13
Liste des sigles et abréviations.....	14
Position du problème.....	15
Chapitre I. Introduction générale.....	17
1. Prosodie émotionnelle	17
1.1. Prosodie.....	17
1.2. Prosodie émotionnelle	18
2. Prosodie émotionnelle : Modèle général neuroscientifique.....	19
2.1 Aspect structurel.....	19
2.2. Aspect fonctionnel.....	20
2.3. Aspect temporel.....	22
3. Paramètres vocaux importants	25
3.1. Paramètres de base	25
3.2. Paramètres de qualité de la voix.....	25
3.3. Variations des paramètres acoustiques dans la transmission d'informations émotionnelles	27
4. Prosodie émotionnelle des accents	28

4.1. Les émotions au travers des cultures	28
4.2. Les accents	31
4.3. Reconnaissance de la prosodie émotionnelle des accents régionaux	32
4.4. Marqueur cérébral : Reconnaissance de la prosodie émotionnelle des accents	33
4.5. Communication interculturelle Français - Québécois	33
5. Objectif et hypothèses	34
Chapitre II. Article de mémoire: <i>Acoustic profile of french emotional sentences uttered with French and Quebecois accents.</i>	36
Abstract	37
Introduction	39
1. Prosody	39
1. 1. Emotional prosody.....	39
1.2. Communicating emotions according to culture.....	40
2. Emotional prosody of regional accents	41
2.1. Prosody of regional accents, in neutral and emotional speech	41
Materials and Method	44
Statistical analyses	46
Results	47
Discussion	57
1. Cross-cultural emotional prosody	57
1.1. Expression and recognition of cross-cultural emotional prosody.....	57

2. Cultural comparisons: France-Québec	58
2.1. Hofstede's model.....	58
2.2. Shame and pride	59
3. Gender comparisons	60
Conclusion.....	61
Chapitre III. Discussion générale	62
1. Synthèse des résultats.....	62
2. Expression cross-culturelle de la prosodie émotionnelle	63
2.1. Méthodologie des banques de stimuli	63
2.1.1. Choix du matériel	63
2.1.2. Construction des phrases	64
2.1.3. Choix d'enregistrement	65
2.2. Validation des stimuli : reconnaissance des émotions prosodiques	66
2.2.1. Paramètres vocaux: profils interculturels.....	67
2.2.2. Profils acoustiques des émotions.....	67
3. Cultural aspect of prosody Québec-France	68
3.1. Les échelles de Hofstede	68
3.2. Education.....	70
3.3. Différences de prosodie selon le sexe	71
4. Fierté et honte : Des émotions culturelles ?	72
4.1. Profils acoustiques de la honte et de la fierté.....	72
4.2. Expression et reconnaissance de la fierté et de la honte.....	73
4.3. Émotions de conscience de soi.....	73
4.4. Fierté et honte : Différences entre les sexes	74

4.5. Contexte politique.....	75
5. Limites et futures recherches	75
5.1. Etude EEG en cours.....	76
Conclusion.....	78
Références	79
Annexes	97

Remerciement

Je tiens à remercier toutes les personnes qui m'ont accompagné durant ce beau projet. Tout d'abord je remercie mon superviseur Dr. Simon Rigoulot qui m'a donné l'opportunité de venir faire mon master au Canada à ses côtés dans un laboratoire d'une bienveillance tout à fait remarquable. Je tiens aussi à remercier tous les professeurs, techniciens et élèves de mon laboratoire, le Brams, pour l'expérience humaine et professionnelle que vous m'avez apportée. Pour l'entraide, leurs soutiens et leurs bons conseils qui m'ont aidé à progresser et a mené à bien ce projet. Merci à mes deux collègues de master Beatriz Oliveira et Samaneh Zanjani pour avoir relu le mémoire ainsi que d'avoir été présentes durant ces deux années comme soutien amical et professionnel.

Merci à mes parents et ma famille de m'avoir encouragé, guidé et accompagné durant toutes cette expérience. Vous me soutenez dans chaque étape, qu'elles soient dures ou merveilleuses. Je suis reconnaissante à mes amis de France de rester toujours aussi fidèle et de toujours me soutenir dans tous mes projets même si cela engendre d'être loin de vous, ainsi qu'à mes ami.e.s de Montréal qui ont su faire de cette aventure de beaux souvenirs et un nouveau foyer avec plein de joie de vie.

On ne peut pas finir sans remercier le chat Marcel qui m'a tenu compagnie durant la rédaction de ce mémoire.

Liste des tableaux

Dans l'article « Acoustic profile of french emotional sentences uttered with French and Quebecois accents »

Table 1: Vocal parameters studied in prosody	44
Table 2: Comparison Qc-Fr of vocal parameters per emotions	55
Table 3: Comparison of men Qc-Fr of vocal parameters per emotions	56
Table 4: Comparison of women Qc-Fr of vocal parameters per emotions.....	56

Liste des figures

Dans l'introduction du mémoire

Figure 1 : Réseaux neuronaux du traitement des informations émotionnelles prosodiques...22

Figure 2: Modèle intégratif du traitement cérébral structurel et temporel des informations émotionnelles prosodiques24

Dans l'article « Acoustic profile of french emotional sentences uttered with French and Quebecois accents »

Figure 1: FOM for emotions and gender between Qc and Fr 50

Figure 2: IntM for emotions and gender between Qc and Fr 51

Figure 3: Duration for emotions and gender between Qc and Fr 52

Figure 4: Spectral Slope for emotions and gender between Qc and Fr 53

Figure 5: Hammarberg index for emotions and gender between Qc and Fr 54

Figure 6: Jitter for emotions and gender between Qc and Fr 56

Liste des sigles et abreviations

ANOVA : Analysis of variance

EEG : Electroencéphalogramme

ERP : Event-related potential

Emo : Emotions

F0 : Fréquence fondamentale

Fr : Français

GLM : General Linear Model

HNR : Harmonics-to-Noise Ratio

Int : Intensité

M : Mean

OFC : Zones orbitofrontales

Qc : Québécois

SD : Standard Deviation

TVA : Zones temporelles de la voix

Position du problème

Lorsque nous communiquons avec les autres, l'interprétation du message de notre interlocuteur va au-delà de la compréhension des mots échangés. En effet, à travers la parole, nous échangeons non seulement des idées liées aux mots qu'on utilise et leur sémantique (ce qu'on dit) mais aussi liées à la façon et au contexte dans lequel ces mots sont dits (comment on le dit). Parmi les différents facteurs contextuels de notre parole, la prosodie, qui renvoie aux variations du ton de la voix, joue un rôle important aussi bien pour aider à la compréhension du message délivré que dans la communication sociale. En effet, la prosodie fournit une multitude d'informations sur le locuteur comme son sexe (les femmes ont une voix plus aigüe), son âge, son milieu social et géographique et son état émotionnel (Ladd, 2008 ; Gobl, 2003, Scherer, 1986). Elle influence ainsi la compréhension du message, l'interprétation des intentions de l'orateur et même la qualité de notre interaction sociale (Mauchand et al., 2020; Krüger, 2018). Par exemple, une phrase comme 'il a laissé la porte ouverte' ne transmettra pas le même message lorsqu'elle est exprimée avec une intonation de tristesse (signifiant un départ non voulu) ou de colère (dans le cas d'un oubli). La prosodie est aussi un marqueur de l'origine géographique voire culturelle d'une personne via son accent, c'est-à-dire que l'on peut distinguer à travers la prononciation des mots, l'origine d'une personne. Lorsqu'on apprend une langue, il faut être conscient des variations allant de la prononciation des mots à la prosodie, et des personnes apprenant une langue étrangère font des erreurs de production et de compréhension faute d'utiliser la prosodie adéquate. Plusieurs études ont montré que les locuteurs natifs d'une langue sont plus sensibles (ont une meilleure reconnaissance) à la prosodie des locuteurs de leur langue maternelle qu'à celle de locuteurs parlant des langues étrangères (Kjellin, 1999 ; Lengeris, 2012 ; Bhatara et al., 2016 ; Chu et al., 2016). De plus, les variations d'accent peuvent avoir un impact important sur la communication sociale. Souvent, les gens attribuent certains traits de personnalité ou des

stéréotypes sociaux, généralement négatifs, aux personnes en fonction de leur accent (Labov, 2006 ; Lippi-Green, 1997 ; Giles, 1970). Par exemple, les personnes avec un fort accent du nord de la France peuvent être perçues comme plus franches et directes découlant des généralisations sur les stéréotypes du caractère pragmatique des habitants du nord. Au final, la prosodie et l'accent d'une personne lorsqu'elle s'exprime jouent un rôle important dans la communication sociale et la manière dont les gens perçoivent les autres. Il est donc indispensable de comprendre comment les accents modifient la prosodie émotionnelle et la reconnaissance du message, d'autant plus dans une ville comme Montréal considérée comme une Métropole multiculturelle où interagissent quotidiennement des personnes d'origines différentes, et plus particulièrement les Français avec les Québécois (Lavoie et Denis, 2008). Cependant, aucune banque de stimuli contenant des phrases émotionnelles prononcées en français et avec différents accents n'est disponible pour atteindre les objectifs de ce projet. Notre objectif est de créer une telle banque de stimuli afin de caractériser les signaux prosodiques émotionnelles prononcés selon l'origine culturelle française (de France) et québécoise. Ensuite, nous voulons comprendre la différence d'expressions de la prosodique émotionnelle des québécois et des Français en faisant des analyses des paramètres vocaux. Ce mémoire de maîtrise s'inscrit dans un projet plus large qui permettra l'utilisation de ces stimuli et de ces données pour de futures investigations en neuroimagerie, afin d'identifier les corrélats neuronaux des mécanismes de perception et de reconnaissance de la prosodie émotionnelle en fonction de l'origine culturelle d'une personne.

Chapitre I. Introduction générale

1. Prosodie émotionnelle

1.1. Prosodie

Deux niveaux d'analyse entrent en jeu lorsque deux personnes ou plus communiquent: le niveau sémantique (segmental) c'est-à-dire la signification littérale d'un énoncé vocal, ainsi qu'un niveau extra-sémantique (supra-segmental) permettant l'échange d'informations au-delà des mots (Grandjean, 2021). Ces informations peuvent passer par le langage corporel, les émotions faciales, l'intensité du regard, mais également par différentes variations du ton de la voix ou prosodie. La prosodie, ou variation du ton de la voix, correspond aux changements dynamiques de la voix, que l'on peut mesurer à l'aide de plusieurs paramètres acoustiques, comme la variation de la hauteur (fréquence fondamentale (F0M) de la voix, la longueur des sons (durée), le niveau sonore (intensité) ainsi que le timbre de la voix. Ces variations vont servir plusieurs fonctions langagières comme fournir des informations grammaticales en distinguant une affirmation d'une question (c'est-à-dire que la voix va monter en tonalité pour une question et baisser pour une affirmation ; Pierrehumbert et Hirschberg, 1990). Ce faisant, la prosodie fournit de nombreuses autres informations, comme l'identité du locuteur (Hirschfeld et Trouvain, 2008), son sexe biologique (les femmes ont en moyenne une fréquence plus élevée que les hommes), son âge (les tout-petits ont une prosodie qui est souvent monotone et peu expressive, les jeunes adultes ont souvent une prosodie plus dynamique, et les personnes âgées peuvent avoir une prosodie plus plate en raison de problèmes de santé ou de vieillissement), mais aussi dans la perception de la personnalité de notre interlocuteur. Par exemple, une personne qui parle très rapidement avec une intonation monotone peut être perçue comme ennuyeuse ou peu enthousiaste. D'autre part, une personne qui parle lentement avec une intonation fluctuante peut être perçue comme plus expressive et émotionnelle (Cohen et al., 2001). L'analyse de la prosodie aide aussi à déterminer les intentions d'un locuteur. Celui-ci

peut masquer ce qu'il veut vraiment dire en utilisant des figures de style comme l'ironie, les métaphores. L'ironie est souvent marquée par une prosodie particulière, caractérisée par des inflexions de voix inhabituelles, des pauses et des accentuations inattendues (Bryant et Fox Tree, 2005). Plusieurs études ont montré que ces variations de prosodie aident à signaler l'ironie et à transmettre une interprétation différente du sens littéral des mots. Par exemple, un locuteur peut dire "Super!" avec un ton ironique pour indiquer qu'il est effectivement très déçu ou contrarié, malgré les apparences (Bryant et Fox Tree, 2005 ; Regel et al., 2011 ; Mauchand et al., 2021). La prosodie est aussi utilisée pour évaluer l'humeur et l'état émotionnel d'une personne. Par exemple, une personne qui parle rapidement avec une intonation haute et excitée peut être considérée comme heureuse ou excitée, tandis qu'une personne qui parle lentement avec une intonation basse et monotone peut être considérée comme triste ou déprimée (Juslin et Laukka, 2003).

Ainsi, la prosodie joue un rôle important dans nos interactions sociales, en fournissant une multitude de renseignements sur le locuteur telle que ses intentions ou encore ses états émotionnels (Mauchand et al., 2020; Krüger, 2018).

1.2. Prosodie émotionnelle

Les émotions se définissent comme des phénomènes psychophysiologiques complexes impliquant de nombreux facteurs neurologiques, sociologiques et biochimiques entraînant une réaction globale, intense et brève (LeDoux, 2000 ; Levenson, 1999 ; Panksepp, 1998 ; Ekman, 1992). Historiquement, les émotions ont surtout été étudiées dans la modalité visuelle. Depuis un peu plus d'une dizaine d'années, un intérêt s'est porté vers les sons émotionnels, qu'ils soient environnementaux, musicaux ou langagiers.

Plusieurs travaux ont montré un impact important, sur notre comportement et notre cognition, de la prosodie émotionnelle, c'est-à-dire les variations du ton de notre voix lorsque

notre message est prononcé de manière émotionnelle, (Juslin et Västfjäll, 2008 ; Johnstone et al., 2007 ; Schirmer, 2005 ; Scherer, 2003). Ces travaux ont montré par exemple que la prosodie émotionnelle influence la mémorisation et la reconnaissance des mots (Kotz et Paulmann, 2011 ; Schirmer et Kotz, 2002) et qu'on se souviendrait davantage des mots joyeux que des mots neutres (Siegel et al., 2013). En outre, la prosodie émotionnelle peut également avoir un impact significatif sur la communication interpersonnelle et la prise de décision. En effet, les médecins qui utilisent une prosodie émotionnelle appropriée sont perçus comme plus compétents et empathiques. Les patients sont alors plus enclins à leur faire confiance et à suivre leurs recommandations (Dóra et Csíkos, 2019 ; Yee et Müller 2013).

Des recherches en neuroimagerie ont aussi été menées pour identifier les activités cérébrales responsables du traitement de la prosodie émotionnelle et de son impact sur notre cognition, et ont ainsi mis en évidence l'existence d'un ensemble de structures impliquées à différentes étapes de traitement.

2. Prosodie émotionnelle : Modèle général neuroscientifique

L'étude de la prosodie émotionnelle permet de mieux comprendre les mécanismes neuronaux sous-jacents à cette communication. La perception de la prosodie émotionnelle activerait des régions spécifiques du cerveau. Kotz et Paulmann (2011) présentent un modèle intégrant des données structurelles, fonctionnelles et temporelles des mécanismes cérébraux lors de l'écoute de phrases avec différentes prosodies émotionnelles (peur, colère, joie, surprise, agréable) (**figure 1**).

2.1. Aspect structurel

Sur un plan structurel, les auteurs confirment des activations au niveau du gyrus temporal moyen et supérieur bilatéral, du gyrus frontal inférieur gauche et des zones orbitofrontales (OFC) médiane et latérales. On sait maintenant que le cortex préfrontal, non

mentionné dans le modèle, est aussi impliqué dans la régulation émotionnelle et la prise de décision (Craig, 2009 ; Singer et al., 2009). Au niveau sous-cortical, on retrouve le thalamus et le complexe amygdalien qui traitent les informations émotionnelles et déclenchent des réponses émotionnelles automatiques (Phelps et LeDoux, 2005, Lichev et al., 2015). Le complexe amygdalien aurait une fonction plus large dans la cognition sociale, en particulier comme "détecteur de pertinence" car il orienterait l'attention sur des informations jugées pertinentes et ainsi biaiser les informations objectivement transmises (Sander et al., 2003). Pour finir, une structure non mentionnée dans ce modèle, dont le rôle est maintenant bien établi dans la littérature, sont les ganglions de la base. Son rôle dans les émotions est dû au fait que la partie ventrale du striatum reçoit des projections neuronales du système limbique (Weydenmyer et Gallman, 2007). Son altération, notamment lors dans la maladie de Parkinson, a montré des dérèglements émotionnelles (Péron et al., 2012). Son activité serait modulée par les indices acoustiques temporels tels que les variations du rythme, de la hauteur et de l'intensité du signal auditif (Paulmann al., 2008; Frühholz et al., 2011). Des lésions focales des ganglions de la base gauche altèrent le processus exécutif qui sous-tend la tâche de reconnaissance durant le traitement prosodique émotionnel. En fait, ces lésions n'altèrent pas le traitement prosodique émotionnel implicite mais les processus évaluatifs (Paulmann et al., 2008).

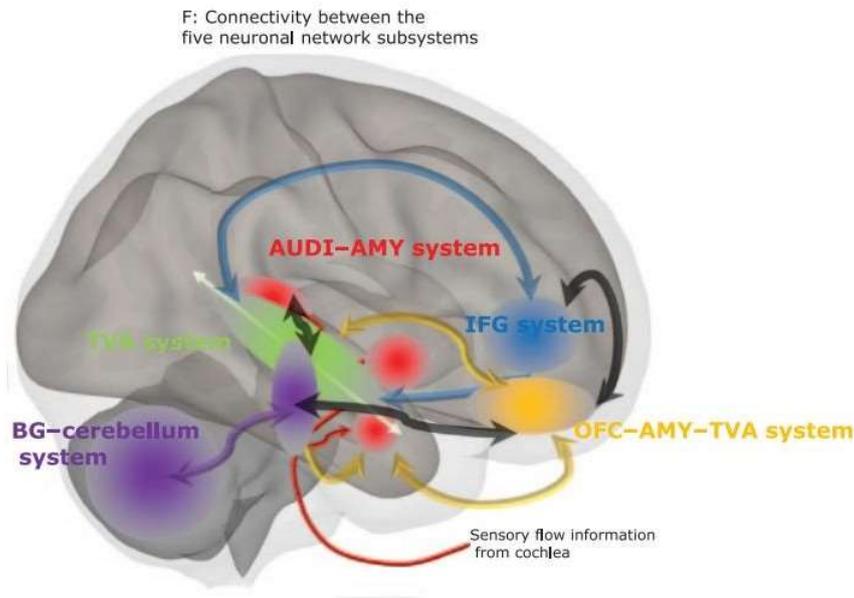
2.2. Aspect fonctionnel

D'un point de vue fonctionnel, une communication / connectivité efficace entre les régions corticales frontales inférieures et les zones temporeles de la voix (TVA) serait indispensable pour catégoriser et discriminer les émotions d'un discours vocal (Frühholz & Grandjean, 2012). De plus, une bonne connectivité serait nécessaire entre les régions TVA avec les parties médianes de l'OFC et les amygdales afin de coder les stimuli sociaux auditifs et de les contextualiser (Grandjean, 2021).

De manière plus approfondie, un modèle intégratif fonctionnel de cinq réseaux neuronaux pour traiter la prosodie émotionnelle a été proposé plus récemment (Grandjean, 2021). Après les relais des informations par le tronc cérébral et le mésencéphale, un *premier* réseau sous-cortical représentant le flux ascendant composé de thalamus auditif des régions auditives primaires et l'amygdale. Sa fonction serait principalement liée à l'encodage brut du stimulus auditif émotionnel. Les amygdales modèleraient aussi un deuxième système, dit réseau cortical temporal, qui comprend les TVAs situées dans les gyri et sulci temporaux moyens et supérieurs. Il y a un gradient de complexité sonore de la partie médiane vers les parties antérieures de ces régions, ainsi qu'un gradient d'intégration multimodale vers l'arrière du système comprenant, la branche ascendante du gyrus temporal inférieur. Ce système serait impliqué dans la perception intégrée de l'objet auditif et ses relations avec les autres informations comme celles visuelles. Un *troisième* système frontal étroitement lié au système TVA est composé de gyri et de sulci frontaux inférieurs impliqués dans le traitement implicite et explicite de la catégorisation et la discrimination des informations émotionnelles auditives. Ce système serait aussi important dans sa capacité à moduler l'utilisation des informations extraites du signal acoustique pour réaliser une tâche spécifique et serait ensuite impliqué dans le flux d'informations en retour sur les réseaux précédents. Le *quatrième* système important dans le flux d'informations en retour comprend l'OFC avec sa partie médiane étroitement liée au système sous-cortical TVA, essentiel dans l'intégration des aspects contextuels pour l'évaluation d'un objet sonore spécifique. Enfin, le *cinquième* réseau est composé de noyaux sous-corticaux, des ganglions de la base (Ceravolo et al., 2021), et du cervelet permettant l'organisation du traitement des sons, ainsi que dans la formation des habitudes dans le contexte de la perception auditive des émotions et de ses relations avec les tendances à l'action et les comportements. En outre, il joue un rôle important dans l'inhibition des réponses prédominant dans le contexte du traitement des informations auditives.

Figure 1

Réseaux neuronaux du traitement des informations émotionnelles prosodiques



Note: Figure reprise de “Brain Networks of Emotional Prosody Processing” par Grandjean, D., 2021, *Emotion Review*. Connectivité entre les cinq réseaux neuronaux anatomiquement et fonctionnellement distincts (flèches noires) entre les réseaux OFC-AMY-TVAs et IFG-TVAs, les réseaux OFC-AMY-TVAs et BG-cerebellum, et les réseaux AUDI-AMY et TVA. AMY = amygdale et IFG-TVAs. AMY = amygdale ; AUDI : auditif ; IFG : gyrus frontal inférieur ; OFC : aire orbitofrontale ; TVA : aire vocale temporale.

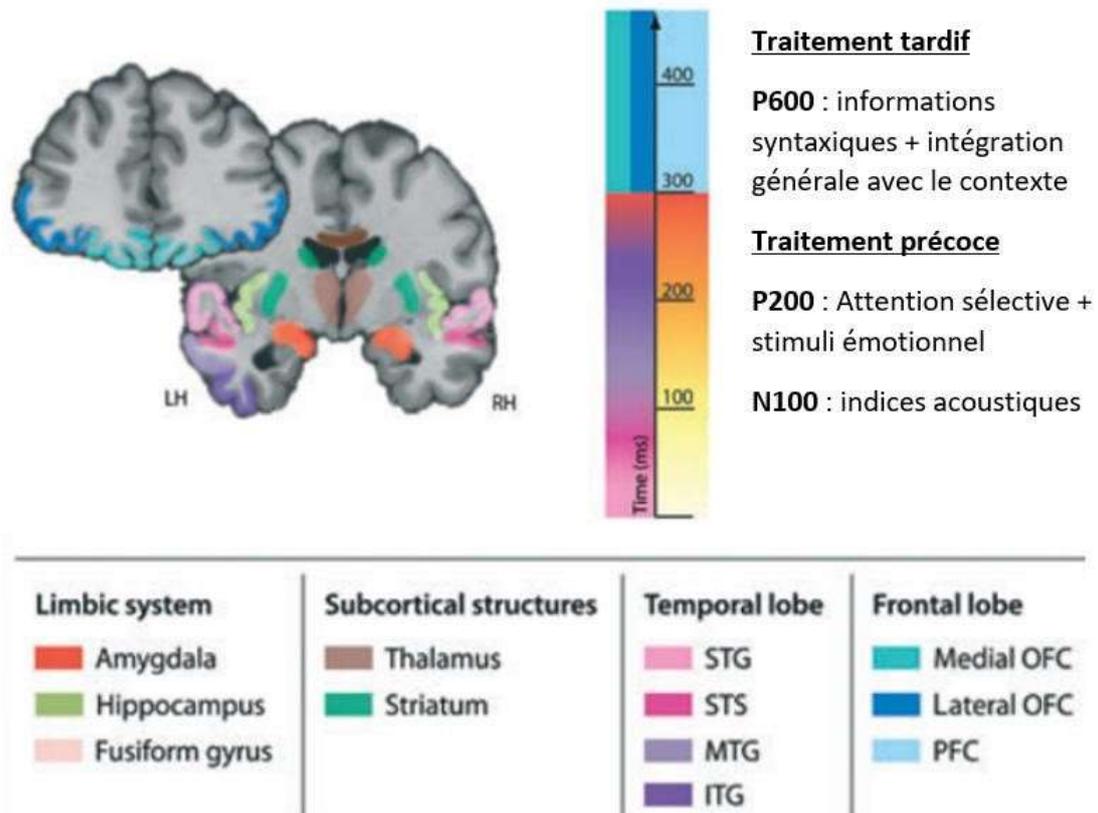
2.3. Aspect temporel

Sur le plan temporel, le modèle de Paulmann et Kotz renvoie l'idée d'un traitement précoce (rapide) de l'activité cérébrale, comme le montre des études portant sur les composantes N100 et P200, (N100 et P200: réponse neuronale à 100 et 200 ms après le stimulus sensoriel) (Paulmann al., 2008), mais aussi de traitements plus tardifs, plus complexes, avec des effets sur la composante P600 (Voeten et Levelt, 2019). De manière fonctionnelle, la composante précoce N100 semble coder pour les indices acoustiques liés à l'identité des

signaux vocaux (fréquence, intensité). Par la suite, la composante P200 coderait pour les processus d'attention sélective avec un traitement préférentiel précoce des stimuli à forte signification motivationnelle. Elle serait également modulée par les stimuli émotionnels entendus (Paulmann et al., 2006). De façon tardive, la composante P600 jouerait dans le traitement des informations syntaxiques et serait également corrélée avec un processus d'intégration des informations déjà encodées avec le contexte prosodique (Voeten & Levelt, 2019). Ainsi, ce modèle général montre l'importance du séquençage temporel du flux d'informations et une bonne connectivité entre structures pertinentes activées lors de l'écoute de la prosodie émotionnelle.

Figure 2

Modèle intégratif du traitement cérébral structurel et temporel des informations émotionnelles prosodiques



Note: Adaptée de “Emotion, Language, and the Brain” par Kotz, S.A. & Paulmann, S., 2011, *Language and Linguistics Compass*, de la figure 1, p. 119. *STG*, gyrus temporal supérieur ; *STS*, sillon temporal supérieur ; *MTG*, gyrus temporal moyen ; *ITG*, gyrus temporal inférieur ; *OFC*, cortex orbito-frontal ; *PFC*, gyrus préfrontal ; *LH*, hémisphère gauche ; *RH*, hémisphère droit.

Le modèle montre le rôle important d’une première étape dans le traitement des informations émotionnelles. Cette étape correspond au décodage « brut » des informations acoustiques. Ces variations acoustiques sont liées aux changements physiques au niveau du

système de l'appareil de phonation (production de la parole) (Schirmer et Kotz, 2006) avec des différences de tensions du larynx et de la pression sous-glottale qui vont moduler le signal acoustique du son, caractérisé par une multitude de paramètres vocaux.

3. Paramètres vocaux importants

La communication à travers la voix se caractérise par des variations de paramètres vocaux, permettant ainsi une communication efficace comme la transmission des nuances et des émotions. Il en existe une multitude dont je présente, ici, les plus importants dans la transmission de la prosodie émotionnelle.

3.1. Paramètres de base

La **Fréquence fondamentale (F0)** d'un son complexe renvoie à sa hauteur, grave ou aigue. Physiquement, elle correspond à la fréquence d'ouverture et de fermeture des cordes vocales. Chaque individu possède une fréquence fondamentale de base qui lui est propre (Titze, 1994 ; Laver, 1980). Cette fréquence est modifiée selon un tas de facteurs comme la variabilité de l'état émotionnel de la personne. **L'intensité** renvoie à l'énergie d'un son (à la force sonore de la voix) et peut aller de faible à forte. Elle est différente du volume sonore, qui est l'amplitude du signal. **La durée** réfère à la longueur de la voix parlée. **Le débit de la voix** est la vitesse à laquelle on parle c'est-à-dire le rythme de parole à peut-être, lent ou rapide.

3.2. Paramètres de qualité de la voix

Le jitter et le shimmer sont des mesures communément utilisées pour évaluer la qualité de la voix. Le *jitter* réfère aux variations de fréquence fondamentale (F0) de la voix, tandis que le *shimmer* renvoie aux variations d'amplitude de la voix. Le *jitter* est mesuré en calculant l'écart-type des différences de F0 entre les cycles vocaux successifs. Le *shimmer* est mesuré en calculant l'écart-type de l'amplitude du signal de parole sur une période donnée. Il existe un

certain nombre de méthodes pour mesurer le *jitter* et le *shimmer*, notamment l'analyse de la forme d'onde, l'analyse de la période fondamentale et l'analyse de la source de bruit (Hirano ; 1981; Titze, 1989; Dejonckere et al., 2001). Ces paramètres sont utiles dans la reconnaissance du locuteur avec la sensation et la perception de bruit dans le signal vocal dû à une irrégularité de l'onde glottique (Farrus, 2007). **Les harmoniques** sont des fréquences multiples de la F0 et jouent un rôle important dans la structure sonore de la voix. La présence ou l'absence d'harmoniques peut donner un aperçu de la qualité de la voix, et renseigner sur l'âge, le sexe et les anomalies vocales du locuteur ou dysphonie (altération des qualités acoustiques de la voix comme la dysprosodie ; Fredouille, 2019). **Le HNR (Harmonic to noise ratio)** se définit par le rapport entre les composantes périodiques (harmonique avec les vibrations des cordes vocales) et non périodiques (bruit glottal) d'un segment de voix (Aswathy et Shetty, 2022; Murphy et Akande, 2005; Teixeira, 2013). Ce paramètre joue sur les caractéristiques de perception de la voix rude ou enrouée ou à l'inverse essoufflée (Ikuma et al., 2022). L'indice **Hammarberg** mesure la quantité de bruit dans le signal vocal et son impact sur la qualité de la voix. Plus précisément, il mesure le rapport entre l'énergie sonore dans la bande de fréquence des vibrations des cordes vocales et l'énergie sonore dans les fréquences où le bruit est produit (Hammarberg, 1977). Les voix avec un indice élevé sont jugées comme moins qualitatives que celles avec un indice plus bas (Titze et al., 1995).

Enfin, **la pente spectrale** est un indice important de la qualité de la voix prenant en compte les différentes harmoniques et faisant le rapport entre l'amplitude fondamentale par rapport aux composantes de hautes fréquences et des amplitudes des harmoniques les plus près des 3 premiers formants (maximas de pics d'intensité). Une pente plate se caractériserait par une voix plus grinçante alors qu'une pente raide renverrait à une voie plus aérée. De plus, cet indice jouerait comme indicateur phonétique dans la prononciation d'une langue (Kuang et Libermann, 2015).

Ainsi la prosodie se caractérise par une multitude de paramètres vocaux du signal sonore, qui peuvent varier en fonction de l'état émotionnel du locuteur.

3.3. Variations des paramètres acoustiques dans la transmission d'informations émotionnelles :

Les variations complexes de l'ensemble de ces paramètres peuvent avoir un impact significatif sur la perception des émotions dans la parole. Les paramètres vocaux formeraient des patrons acoustiques permettant la perception et la discrimination des émotions (Brunswik 1956; Banse et Scherer, 1996 ; Hammerschmidt et Jurgens, 2007; Mauchand et al., 2020). En effet, les liens entre le discours émotionnel et les caractéristiques prosodiques sont communément établis depuis plusieurs années. Ekberg et al., (2023) ont analysé des catégories des paramètres vocaux liées à la fréquence, la temporalité, l'énergie (par et l'équilibre dynamique dans quatre émotions : la colère, peur, joie, tristesse et surprise chez des locuteurs suédois. Ils observent notamment que la joie se caractérise par une F0M plus élevée que la colère, dont la F0M est plus élevée que celle de la tristesse. Le Jitter montre lui un niveau plus élevé pour la joie, un peu moins pour la tristesse et encore moins pour la colère. L'intensité vocale est la plus forte pour la colère et la plus faible pour la tristesse. Aussi, la joie ainsi que la colère seraient prononcées plus rapidement que la tristesse. Enfin, l'indice Hammarberg montre des valeurs plus petites pour la colère suivi de la tristesse puis la joie (Ekberg et al., 2023; Laukka et al., 2016). Quant à la pente spectrale, celle-ci est plus grande pour la tristesse et plus plate pour la joie et la colère (Scherer et Banse, 1991 ; Guzman et al., 2013; Ekberg et al., 2023).

Ainsi, les paramètres vocaux agiraient ainsi comme des codeurs d'informations en fournissant des indications sur l'intensité, la qualité et les valences émotionnelles (Banse et Scherer, 1996). Ces informations seraient ensuite décodées en tant qu'éléments émotionnels par

l'auditeur (Brunswik, 1956). Rijn et Larrouy-Maestri (2023) expliquent que cette correspondance est étudiée soit en modélisant la relation entre les paramètres acoustiques et les expressions émotionnelle d'un locuteur, soit en analysant le taux de reconnaissance humaine (perception d'un interlocuteur). Les auteurs avancent que ces approches comportent une limite fondamentale car elles ne prennent pas en compte des informations importantes rendant compte de la diversité de la population testée. Ils ont analysé un corpus de plus de 3000 enregistrements émotionnels et ont cartographié les liens entre paramètres vocaux et prosodie émotionnelle. Ils ont utilisé plusieurs modèles et ont remarqué que même si un modèle général permettait d'expliquer de manière fiable ces liens, les autres modèles prenant en compte les différences entre les individus, les différences de sexes, de langue et de pays étaient encore plus robustes. Ainsi, la prosodie émotionnelle se caractérise par des modulations complexes de nombreux paramètres vocaux. Cette communication émotionnelle par la voix dépasse les modèles simplistes des émotions (six émotions basiques) et compose avec de nombreux autres facteurs comme le sexe ou l'aspect évolutif de la culture et l'éthologie car une langue et le vécu des émotions sont modelés par son environnement culturel et linguistique.

4. Prosodie émotionnelle des accents

4.1. Les émotions au travers des cultures

Paul Ekman en 1972 propose au travers d'études sur les expressions faciales émotionnelles chez des populations isolées de Nouvelle-Guinée qu'il existe six émotions basiques, qui seraient universellement reconnues : la joie, la tristesse, la colère, la peur, le dégoût et la surprise. Des résultats semblables ont été trouvés plus récemment avec la voix et la prosodie. Les émotions prosodiques sont reconnues même lorsqu'elles sont exprimées dans une langue étrangère malgré un temps de reconnaissance plus long (Skorup et Pell, 2008 ; Pell et al., 2009 ; Sauter et al., 2010; Laukka et al., 2016). Dans ce cadre, Pell et al. (2009) ont voulu comprendre comment les émotions sont reconnues et différenciés acoustiquement selon

quatre langues étrangères : l'Anglais, l'Allemand, l'Hindi et l'Arabe. Pour ce faire, ils ont créé une banque de phrases pseudo-phrases (des phrases reprenant l'organisation grammaticale d'une langue avec des mots n'ayant pas de sens, par exemple « *Someone migged the pazzing* » en anglais) en exprimant la colère, le dégoût, la peur, la tristesse, le bonheur, la surprise et l'agréable. Après avoir validé la banque de phrases par des participants natifs de chaque langue, des analyses des caractéristiques acoustiques ainsi que la capacité des participants à reconnaître l'émotion de ces pseudo-phrases prononcées dans leur propre langue ou dans des langues étrangères. Chaque émotion a été reconnue par les participants de toutes les origines mais à des taux de reconnaissances variés. De manière importante, l'indice vocal prédominant dans cette discrimination d'émotions est la fréquence fondamentale. Ces données suggèrent un traitement universel pour certaines émotions mais avec un traitement sous-jacent différents selon la langue innée ou étrangère.

Des données plus précises affinent cette vision. L'expression des émotions (faciales et vocales au moins) est modelée par notre culture et notre vécu personnel (Matsumoto et al., 2005; Jack et al., 2012). En effet, chaque région a des pratiques, des croyances et des interprétations différentes. Le modèle d'Hofstede (Hofstede, 1980) permet de mesurer la distance culturelle entre différents pays en se basant sur le rapport individualisme/collectivisme. Les cultures individualistes, comme les États-Unis, valorisent les émotions positives et la colère exprimée de manière socialement acceptable, tandis que les cultures collectivistes, comme la Chine, ont une attitude moins favorable envers les émotions positives et la colère, les considérant comme une violation de l'autorité et de l'harmonie social (Eid & Diener, 2001). Les émotions sont donc perçues différemment selon la culture. Par exemple, dans la culture occidentale, se vanter de ses réalisations est souvent considéré comme une forme d'arrogance ou de prétention, tandis que dans certaines cultures asiatiques, cela peut être considéré comme une forme de modestie et d'humilité. Par conséquent, une personne occidentale pourrait

éprouver de l'inconfort ou de la gêne lorsqu'elle est confrontée à quelqu'un qui se vante de ses réalisations, tandis qu'une personne d'une culture asiatique pourrait éprouver de l'admiration et du respect envers cette même personne. Cette différence interculturelle dans l'interprétation d'une situation objective peut donc entraîner des émotions différentes (Weisz et Chen, 2018 ; Markus et Kitayama, 2010 ; Morling et al., 2002).

Ces différences sont d'autant plus importantes à comprendre qu'elles sont régies par des règles de société. Les émotions ressenties sont ainsi soumises à un processus de régulation allant de l'inhibition à l'expression forte de celles-ci qui serait déterminée par l'expérience personnelle ainsi que l'expérience socioculturelle (Mesquita, 2003), ce qu'on appelle parfois les « *display rules* », étant des normes ou des conventions culturelles qui régissent la manière dont les émotions doivent être exprimées ou dissimulées dans un contexte social donné.

Par ailleurs, les différences émotionnelles culturelles se font voir aussi à travers la voix, à travers la prosodie émotionnelle. Les études sur la reconnaissance de la prosodie émotionnelle selon les langues étrangères s'accordent que les émotions sont universellement reconnues mais qu'il y a une meilleure reconnaissance des émotions venant d'un locuteur du même groupe culturel (*in-group*) comparé à un locuteur venant d'un groupe culturel différent (*out-group*) (eg. Laukka et Elfenbein, 2021 ; Paulmann et Uskul, 2014 ; Golestani et al., 2011) (Elfenbein & Ambady, 2002). Afin de pousser ces résultats, Laukka et al. (2016) ont étudié les indices acoustiques principaux de phrases émotionnelles pour vérifier l'existence au niveau vocal d'un 'avantage du propre groupe'. Un nombre de 550 phrases émotionnelles ont été sélectionnées de la banque de stimuli VENEC avec des locuteurs natifs de cinq pays différents (Australie, Inde, Kenya, Singapour et les Etats-Unis) dans 11 états émotionnels (ex : peur, colère, fierté et honte). Grâce aux analyses physiques de la voix de chaque locuteur ainsi qu'une étude de reconnaissance des émotions entre chaque groupe, les auteurs remarquent que les indices acoustiques (ex: F0M, IntM, HNR, indice Hammarberg) sont utilisés de manière relativement similaires par tous les

groupes pour produire et reconnaître les émotions, mais, il y a tout de même des différences subtiles notables, soit une meilleure validité et utilisation des indices pour les membres du propre groupe.

Les émotions (exprimées et perçues par des canaux sensibles comme le traitement d'expressions faciale et la voix) seraient donc universelles mais sujettes à des variations culturelles. Mais qu'en est-il des personnes parlant la même langue mais ayant un accent régional différent?

4.2. Les accents

Les accents font référence aux variations du signal vocal observées suivant l'origine géographique du locuteur. Selon « l'hypothèse de la distance perceptuelle », les accents sont classés sur une échelle perspective en fonction de la distance acoustique par rapport à la langue maternelle. Les accents étrangers se situent à l'extrémité de l'échelle alors que les accents régionaux se situent le plus proche de la langue maternelle. Goslin et al. (2012) expliquent que les accents régionaux seraient une « déviation cohérente » du niveau phonologique, phonétique et prosodique alors que pour les accents étrangers, il y aurait une variation significativement différente avec la langue maternelle. Ces accents ne sont pas sans effet au niveau cognitif, par exemple en termes de reconnaissance de mots (Floccia et al., 2006 ; Adank et al., 2014). Floccia et al. (2006), ont réalisé une série de 6 expériences visant à étudier le temps de traitement impliqué lors du traitement de la parole avec un accent régional Français (Parisien, Lille et Toulouse) et un accent étranger anglais (Londres). Les participants français avaient pour tâche de prendre des décisions lexicales sur les phrases entendues. En manipulant la durée des phrases et les accents, les auteurs observent une identification plus longue des mots lorsqu'ils sont prononcés avec un accent régional (30ms) et encore plus longue lorsqu'ils sont produits avec un accent étranger (100ms). Cette étude conclut la présence d'un mécanisme d'ajustement à

court terme variant selon le type d'accent. Ainsi, il est montré un effet de reconnaissance dû aux accents étrangers et régionaux lors d'écoutes de stimuli sonores tels que des phrases.

4.3. Reconnaissance de la prosodie émotionnelle des accents régionaux

Afin d'étudier la prosodie émotionnelle des accents régionaux, Mauchand et Pell (2020) ont étudié la production et la perception de la plainte avec des accents français et québécois. Ils ont tout d'abord créé une base de stimuli étant 320 courtes phrases de plainte avec des acteurs Français et Québécois. Afin d'étudier la production prosodique, une analyse acoustique a été effectuée sur plusieurs paramètres vocaux dont F0M, HNR, Jitter, Shimmer, et la durée de fin de mots. Les auteurs ont observé que le paramètre permettant le plus de discriminer la plainte est F0M. De plus, même si associé à de faibles effets, les locuteurs québécois (Qc) exprimaient de manière plus forte la plainte que les Français (Fr). Plus précisément, les Québécois avaient une plus grande variabilité de F0 (F0range) et un patron rythmique distinct. En termes de qualité de la voix, les Québécois avaient un HNR plus petit donnant une perception de la voix plus sévère.

En ce qui concerne la perception de la plainte, des participants des deux régions (Qc et Fr) ont écouté les phrases des différents acteurs (Qc et Fr). Ils avaient pour tâche d'évaluer sur une échelle de Likert à 7 points 'A quel point la personne est en train de se plaindre?'. Les auteurs ont constaté un effet d'avantage du propre groupe, c'est-à-dire que les québécois reconnaissaient mieux la plainte lorsqu'elle était produite par un locuteur Québécois et inversement pour les Français. Ils ont également noté que les Québécois étaient perçus comme plus en colère et plus surpris, alors que les Français plus tristes.

Cette étude suggère une différence d'expression des émotions à travers la voix, prosodique, ainsi qu'une différence de perception selon des accents régionaux différents ici en français. D'autres données au niveau cérébral viennent appuyer cette idée.

4.4. Marqueur cérébral : Reconnaissance de la prosodie émotionnelle des accents

Afin de comprendre au niveau cérébral ce qui pourrait expliquer la différence de reconnaissance des informations prosodiques régionales, Jiang et al. (2020) ont étudié le rôle de la prosodie dans l'expression de la confiance et de l'impact d'un accent sur l'impression social lors d'échange interculturelle. Des participants Anglais Canadiens ont entendu des phrases en anglais exprimées avec de la confiance ou du doute par des locuteurs parlant leur propre dialecte (condition dite *in-group*), par des locuteurs parlant un dialecte régional différents (Australien) et un dialecte étranger (Québécois). Les participants devaient évaluer le degré de crédibilité de l'orateur. Les réponses cérébrales étaient enregistrées avec un casque EEG. Les résultats montrent que lors de l'écoute du propre groupe (*in-group*) avec une prosodie de confiance, les participants avaient un signal ERP précoce (N100 et P200) significativement différent du groupe étranger (*out-group/foreign*) et inexistant lors de l'écoute du groupe étranger régional (*out-group/regional*). Par contre dans ce dernier cas, les auteurs ont observé une augmentation du signal tardif N400, ce qu'ils interprètent comme le reflet d'un traitement plus accru du contexte et une réanalyse de la notion non native de la crédibilité. En somme, les locuteurs *in group* étaient perçus comme plus confiants tout comme les locuteurs *out-group regional* et les locuteurs étrangers étaient perçus comme les moins confiants.

Cette étude apporte des arguments sur le fonctionnement du système neuronal des effets lors du traitement des informations acoustiques de la prosodie émotionnelle et des accents étrangers et régionaux. Ils apportent l'idée que les accents peuvent moduler le traitement cérébral d'une émotion, un phénomène qui serait subtil mais bien présent avec des accents régionaux.

L'étude de ce phénomène au niveau cérébral est récent et peu fourni.

4.5. Communication interculturelle Français - Québécois

C'est donc un petit corpus d'études émergent sur le sujet qui montrent des données intéressantes. Il serait alors pertinent d'approfondir au niveau cérébral les différences du traitement prosodique selon des accents régionaux. Une place favorable pour ce genre d'étude est Montréal, ville cosmopolitaine où français.e.s de France et québécois.e.s interagissent quotidiennement du fait de la forte immigration de Français à Montréal (et pas seulement au Plateau).

Les Français et les Québécois sont issus d'une même culture lointaine mais les québécois se sont ensuite appropriés une nouvelle culture couplée avec les Anglo-Saxons et le système américain. Les accents distincts en sont les marqueurs de ces dérivations culturelles. Montréal est une ville avec une politique de forte immigration francophone, ce qui fait que les interactions entre Français (de France) et Québécois sont quotidiennes. Des incompréhensions, des stéréotypes et des discriminations sont possibles et pourraient prendre origine, en moins en partie, à une distorsion de la reconnaissance du langage en général et de la prosodie émotionnelle en particulier (Buchanan et al., 2018). Il est donc important de comprendre ces enjeux afin d'apporter une bonne communication interculturelle.

Afin d'étudier la reconnaissance de la prosodie émotionnelle entre les Français et les québécois, il est nécessaire d'utiliser des phrases validées et standardisées en français mais il n'existe pas de telles banques disponibles à la communauté scientifique.

Il n'existe cependant pas, à ce jour, de banque de stimuli comprenant des accents français et québécois avec un large éventail d'émotions comprenant des émotions primaires et culturelles afin d'étudier la reconnaissance de la prosodie émotionnelle des accents régionaux.

5. Objectif et hypothèses

Dans un premier temps, ce qui fera l'objet de notre premier article, nous avons pour objectif de créer et valider une banque de phrases émotionnelles auditives avec des accents québécois et français. Dans un second temps, nous souhaitons analyser les paramètres vocaux :

F0 moyen (M), F0 écart-type (SD), IntM, IntSD, Duration, Spectral Slope, Jitter, Shimmer, Hammarberg) des stimuli validées dans notre banque afin d'étudier les profils acoustiques des québécois et des Français lors de la transmission d'émotions de Colère, Tristesse, Joie, Honte et Fierté. Nous pensons que les Québécois ont une prosodie plus expressive dans chaque émotion. Plus concrètement, comparés aux français ils auraient une F0M plus haute et plus variable, une plus grande intensité et une durée plus courte, un petit indice Hammarberg et une pente spectrale plus plate ainsi qu'un HNR réduit pour la joie, la colère, la fierté et inversement la tristesse et la honte.

Chapitre II. Article en vue de soumission dans le journal *Emotion*

Acoustic profile of french emotional sentences uttered with French and Quebecois accents

Amélie Rolinat, Simon Rigoulot

Laboratoire International du Cerveau de le Musique et du Son (BRAMS), Centre de Recherche sur
le Cerveau le Langage et la Musique (CRBLM)

Département de Psychologie, Université de Montréal

Author Note

Amélie Rolinat: <https://brams.org/members/amelie-rolinat/>

Simon Rigoulot: <https://simonrigoulot.wordpress.com/team/>

We have no conflict of interest to disclose. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. This study's design and its analysis were not pre-registered. Materials and analysis code for this study are available upon request. The data have not been used in prior published or in press manuscripts.

Correspondence concerning this article should be addressed to Simon Rigoulot, Laboratoire International du Cerveau de le Musique et du Son (BRAMS), Centre de Recherche sur le Cerveau le Langage et la Musique (CRBLM).

Email: simon.rigoulot@umontreal.ca

Abstract

Context: The prosody of speech, i.e. the variations in tone of voice when speaking, plays a key role in social interactions by providing important information linked to identity, emotional state and geographical origin. Foreign accents have a major impact on the way speech is recognized, with significant consequences for social evaluations, such as reduced empathy towards the speaker. However, it is less clear whether this impact persists in the context of regional accents, which are more subtle variations of the speech signal. **Objectif and hypothesis:** The aim of this dissertation is to understand how French-speaking individuals from different regions (France, Quebec) express and recognize emotional phrases spoken by people from the same or different regions. Given the hypothesis of group advantage, we believe that expression and perception differ according to culture. Based on Mauchand and Pell's (2020) study, we expect Quebecers (Qc) to show a more expressive emotional prosody than French people (Fr). **Method:** We created short emotional sentences in 5 emotions (Joy, Sadness, Anger, Pride, Shame), spoken by Quebecois and French actors. This bank of stimuli was validated with an online study by French and Quebecers. Using a general mixed model, we analysed the vocal parameters: mean and standard deviation of the fundamental frequency, standard deviation and mean of the intensity, mean Shimmer, Jitter, HNR, Hammarberg index, Spectral Slope, Duration. **Results:** The parameters of fundamental frequency (F0M), intensity, duration, spectral slope and Hammarberg index differed significantly between the emotions, origins and sexes. For example, of the five emotions, Frs spoke with a higher F0M except for sadness, but Qc spoke with greater intensity and longer duration. The Qc expressed the emotions in a more pronounced way than the Fr, except for anger. Also, many significant differences show that Qc men have a stronger emotional prosody than Fr men. Finally, only differences between Qc and Fr women were observed in the emotions of shame and pride, emotions that are not in the 6 primary emotions, but which

would be more cultural emotions. **Conclusion:** We were able to characterize the emotional vocal expression of Fr and Qc who, despite their common language, express themselves in very distinct ways to convey their emotions. These results open up interesting perspectives on cross-cultural interactions in the same language but from different regions. Prosody is seen as a genuine marker of identity.

Key words: Emotional prosody, regional accents, vocal parameters

Introduction

Globalization has led to an increase in intercultural encounters, which means that intercultural interactions have become more commonplace. However, this cultural diversity poses challenges in terms of communication. Linguistic markers vary from one culture to another. One important aspect of intercultural communication concerns the way in which we express our emotions through the melody of our voice, known as emotional prosody, and how it can be affected by individual characteristics such as their origin.

1. Prosody

When we communicate with our voice, there is what we say through words and sentences (semantic aspect) and how we say it (supra-semantic aspect), which is transmitted at least partly through the tone of our voice (i.e, speech prosody). Prosody is the melody of our voice characterized by variations in a multitude of sound signal parameters such as pitch (F0), intensity, or duration. Thanks to the acoustic structure of the sound produced by a speaker, prosody provides information such as gender (e.g. women have a higher F0M than men), social level, geographical origin (with accent) or emotional state (Ladd, 2008; Cutler et al., 1997; Scherer, 1986). Emotional prosody is the transmission of emotions through the voice.

1. 1. Emotional prosody

Emotions are complex psychophysiological reactions that involve physical and psychological changes, such as bodily sensations, subjective experiences, behaviours, facial and verbal expressions (Gross and Barrett, 2011; Keltner et al., 2003; Ekman 1999; Leventhal and Scherer, 1987). Emotional prosody is the communication of emotion through variations in voice pitch. According to Brunswik lens model (1956), when a speaker expresses emotion through their voice, this is characterized by distinct patterns of acoustic cues for each primary emotion, which are then decoded as an emotional message by the listener (Laukka et al., 2016;

Mauchand and Pell, 2020). For example, a speaker increasing the pitch of his fundamental frequency, duration and intensity may be interpreted as joyful by a listener (e.g., Banse and Scherer, 1996; Scherer and Bänziger, 2004).

The tone of voice will change according to our emotions, marked by distinct acoustic patterns (**see table 1** to see most studied acoustical parameters in emotional prosody research). Anger is expressed by short pauses with a rapid rate of speech (Banse and Scherer, 1996), syllables are often pronounced with great intensity with extreme accentuation often present on the last words of the sentence (Banse and Scherer, 1996) as well as frequent variations in the pitch and rhythm of the voice (Ohala, 1994). On the opposite, sadness is associated with lower frequency variations, longer pauses and less intense tones, and an increase in vowel duration (Thompson et Balkill, 2006). This is how, emotion is communicated through the voice, but some of variations according to the culture has been shown.

1.2. Communicating emotions according to culture

Emotions play a crucial role in social interactions, enabling us to establish links with others, understand their needs and attitudes, communicate our own feelings and cooperate to achieve common goals (Mauss et al., 2007; Ekman, 2003; Lazarus and Lazarus, 2000; Hatfield et al., 1994). Whereas historically, emotions have been thought to be innate and universally recognized (Ekman, 1969), more and more observations suggest that in fact emotions are shaped by many contextual factors, including social norms, cultural values and life experiences (Mesquita Frijdeta, 1992; Kitayama et al., 2006). For example, Japanese culture values emotional restraint, whereas Italian culture encourages emotional expression (Kitayama et al., 2000).

The cultural aspect is an essential factor in communication because it influences the way in which individuals understand and interpret messages. According to Hofstede (1980), cultural values such as individualism or collectivism, hierarchy or equality, distance from power or

tolerance of uncertainty have a significant impact on communicative behavior. Overall, these cultural differences can lead to misunderstandings, misinterpretations, and conflicts between people from different cultures (Gudykunst, 1998). One of the vocal markers of a person's culture and origin is their speech accent. An accent is a set of phonetic, prosodic and rhythmic characteristics that are used by a specific speaker or linguistic community to pronounce the words of a particular language. Accents can vary according to region, ethnic origin, education and other socio-cultural factors (L'Harmattan, 2010; Riches and Foddy, 1989). Accents can be a barrier to understanding speech between two people causing poorer and delayed recognition of verbal information (Floccia et al., 2009; Floccia et al. 2006; Ayako and Hansen, 2007).

2. Emotional prosody of regional accents

2.1. Prosody of regional accents, in neutral and emotional speech

The prosody of foreign accents is often different from that of the mother tongue (Mareuil and Vieru-Dimulescu, 2006). This phenomenon seems happening even for regional accents, which mean having the same mother tongue but with different accent. Gholipour et al (2012) investigated the role of prosodic features in the identification of different Persian accents. The analysis of speech parameters (rhythm-related features and statistics on pitch contour, energy contour and their derivatives) of different regional accents (e.g., Turkish, Kordish, Tehrani, Isfahah) showed that pitch contour drift was the most discriminating cue when identifying a regional accent. This study shows the existence of prosodic differences between regional accents. Prosody also provides emotional information about the speaker, and would also differ according to culture and accent (Rao and Koolagudj, 2013).

Research shows that people listening to speakers of another language (out-groups) recognize emotions through their voice but less accurately. This is known as the 'in-group advantage hypothesis (Elfenbein and Ambady, 2002). There would therefore be cultural and universal

principles in the recognition of vocal emotions (Paulmann and Uskul, 2013; Pell et al., 2009). Also, the emotional acoustic characteristics F0m, F0 range and speech rate are important for recognizing emotions through speech. They are similar but their agency together are not identical across countries (Liu and Pell, 2014). Studies on the subject often have focused on the six universal emotions (Ekman, 1969): joy, anger, surprise, sadness, fear and disgust (Paulmann et Uskul, 2014; Jiang, 2017; Sauter et al., 2019; Rijn et Larrouy-Maestri, 2023). However, there are some exceptions. Mauchand and Pell (2020) studied the production and perception of complaints in Quebec and French accents. They created a sentence bank with 8 actors from these two regions pronouncing short sentences with complaint or neutral. The authors observed that, when they complained, the Quebecers used greater pitch variability, a reduced HNR which the voice is perceived harsher and a different speech rhythm than the French. Quebecers were perceived as angrier, while the French were perceived as sadder. To study the perception of complaining, French and Québécois participants were asked to listen to the audios of the Québécois and French actors and to rate on a 7-point Likert scale 'How much is this person complaining? The results are in line with the in-group advantage hypothesis (Elfenbein & Ambady, 2002), i.e. Quebecers were more likely to recognize a complaint when they heard a Quebec speaker than when they heard a French speaker.

This study shows us that expressions of emotion according to culture involve different variations in subtle vocal parameters. There would also be an advantage in perceiving emotions for one's own group during cross-cultural interaction, even across similar cultures with different regional accents, such as the French and the Quebecois.

French is a pluricentric language, i.e. it involves several cultural systems (Schneider and Barron, 2008) such as French and Quebecois. There are, to our knowledge, no studies looking at the expression of emotions through the voice between the French and the Quebecois, using banks of phrases in french as stimuli. In nutshell, the transmission of emotions through the

voice, or emotional prosody, varies according to accents, even more subtle ones such as regional accents, and influences the foreign listener's perception of the message. Montreal is a perfect example of a cosmopolitan city, with cultural communication issues between French immigrants and French-speaking Quebecers who interact on a daily basis. A bank of emotional sentences with French and Quebec accents should be built in order to study the production and perception of emotional prosody in regional accents.

To the best of our knowledge, there is currently no research on the vocal differences in French and for a large spectrum of emotions including pride, shame, , despite the fact there are commonly used in everyday life. The aim of this study is to 1) create an emotional phrase bank for the emotions of joy, anger, sadness, pride and shame, in French and Quebec accents and 2) analyze the vocal parameters in order to observe acoustic profiles (F0M, IntM, Duration, Spectral Slope, Jitter, Shimmer, Hammarberg index) according to emotions and accents. We predict that Quebecers have a more expressive acoustic profile than French speakers. More precisely, they would have a higher and more variable F0M, greater intensity and shorter duration, a flatter Hammarberg index and spectral slope, and a reduced HNR for joy, anger and pride and, conversely, sadness and shame.

Table 1

Vocal parameters studied in prosody:

- **Fundamental frequency (F0):** pitch, in Hz
- **Intensity:** energy of the sound, in dB
- **Duration:** length of the spoken voice, in seconds
- **Jitter:** Variation in the period of a periodic signal, as a percentage of the sampling period.
- **Shimmer:** Variation in the amplitude of a periodic signal, in dB
- **HNR (Harmonic-to-Noise Ratio):** Measures the ratio between the harmonic (tonal) part and the noisy (non-tonal) part, in Hz.
- **Spectral Slope:** Slope of the curve that relates the amplitude of a given frequency to that of another frequency in a higher band (distribution of acoustic energy in the frequency spectrum of the sound signal), in dB/octave.
- **Hammarberg index:** Represent the spectral distribution of the acoustic intensity of a complex sound, in dB/octave.

MATERIALS AND METHOD

Participants for recording database

We recruited 12 actors equally divided between gender and origin. We selected actors in the studio to achieve good sound quality, crucial for voice analysis, as well as for controlling the desired emotions. They are all between 18 and 35 years old, have French as their mother tongue, and have experience in oral expression (comedy, theater, radio, etc.). Six actors were Quebecers (with Montréal accent) who have not lived abroad for more than 1 year and other six actors were French people (with Parisian accent) who have lived in Quebec for less than 2 years (in order to avoid habituation to the accent). Participants were recruited via posters strategically distributed in theater schools and associations and radio groups. Digital posters on

Facebook groups of actors' clubs, improvisation and comedy were also posted. In a second step, people who responded to the ad were asked to fill out a socio-demographic questionnaire sent by email for a duration of 5 minutes. Finally, the persons corresponding to the criteria were called to evaluate the quality of their accent. The telephone interview was held as follows: An explanation of the study project and their task; questions about their acting and recording experiences; and then they were invited to ask their questions. The session was paid 50 dollars for a duration of 2 to 3 hours.

Stimuli

A total of 280 short sentences were created, 40 sentences in 5 emotions (e.g. *Quelle belle journée / What a beautiful day*) with seven emotions (Joy, Anger, Sadness, , Shame, Pride, (see annex A). In order to control potential grammatical and syntactic effects, a verification of the sentence bank was made considering: the type of personal pronoun used by emotions, the type of sentences (declarative, exclamative, interrogative and imperative), the forms of the sentences (active and passive) and the type of the sentence (personal and impersonal). The number of words per sentence as well as the number of syllables per sentence were checked and balanced.

Procedure

Recording

We followed standardized procedure in order to ensure the highest quality of the stimuli (Mauchand and Pell, 2021). The actors were recorded in a soundproof of the International Brain Laboratory of Music and Sound (BRAMS) at Université de Montréal. Their task was to interpret the sentences in different emotions: Joy, Anger, Sadness, Shame, Pride: (e.g. in Joy: 'Quelle belle journée' (What a beautiful day!); in shame: 'Je ne sais pas nager' "I don't know

how to swim). No personalised linking sentences/words were allowed. For each sentence, two to five tries were made to ensure the quality of the recordings. Author Amélie Rolinat was in charge of the recording and cleaning of data with Reaper software. The files were sorted to keep only what seemed the most expressive version of each sentence. Each audio file was then cut to keep only the sound signal of the sentence, cleaning of parasite sounds (ex: pop of the microphone) which means that the recording duration starts precisely when the speaker begins the sentence and ends when they finish it. In the end, all files were saved and normalized with the following parameters: sampling rate 44.1 KHz, 24-bit PCM, mono, wav format and a maximum peak intensity of 70 dB.

Validation study

We conducted an online validation study of the sentences. The study was performed with JsPsych and uploaded onto the BramsOTP platform: <https://brams.org/category/online-testing-platform/>. The duration of the experiment around 30 minutes, and we built four versions of the experiment, each of them with only eight sentences, uttered by the six actors of one group. In total, each version had 240 sentences (8 sentences x 6 actors x 5 emotions) that were submitted for validation.

Originally, a total of 195 participants of female (due to a too low males' participation we excluded them from the validation study) was aimed to validate the full set of stimuli. They were recruited in France and Quebec online via the newsletters of research centers such as RISC (France) or CRBLM (Quebec) as well as social networks such as Facebook. The remuneration was a draw of 16 vouchers of 50 dollars each. Firstly, we administered the HADS questionnaire (Zigmond and Snaith, 1983) (annex B). People who obtained a total score or greater than 11 in either the anxiety or depression category were excluded from the study. Then, the participants of each group (French and Quebecois) were asked to assess the emotional utterance from the actor or their own group (Quebecois assess Quebecois, French assess French) in order to avoid

accent bias. Participants were asked to 1) choose the emotion they were listening to from 9 choices (Anger, Jealousy, Sadness, Joy, Fear, Surprise, Shame, Fear, Other) and 2) answer to the question “How strong is the emotion expressed?” with a visual analogous scale ranging from “None” to “Maximum (annex C). The utterances’ quality was validated when recognized at a minimum threshold of 70%, which is largely above the chance level (1/9 possibilities) (Castro and Lima, 2010).

Acoustic Parameters

In order to understand the differences in the emotional prosody of Quebec and French speakers, we analyzed their acoustic parameters. For this purpose, we selected the sentences that were recognized accurately by more than 70 % of the participants in the validation study. Analysis was performed on Praat software (Boersma & Weenink, 2014) on relevant parameters chosen based on the parameters used in the literature. Thus, we analyzed the mean standard deviation and range of the fundamental frequency (F0), mean and standard deviation of the Intensity, mean Shimmer, Jitter, HNR, Hammarberg index, Spectral Slope, Duration.

Statistical analyses

We performed a Linear Mixed Effect Model (Gelman, et al., 2020) created with R-studio and packages *lme4*, primarily used for mixed-effects linear modeling and mixed-effects generalized linear modeling (Bates et al., 2015) and *nlme*, Used for modeling longitudinal data and mixed-effects linear modeling (Kuznetsova et al., 2017) followed by an ANOVA (Type II Wald chisquare tests) with the packages *car*, used for visualization, linearity tests, homoscedasticity tests, and regression diagnostics, and *emmeans*, used to estimate marginal means after fitting a model, to further investigate the inter-group effects. The model was run with each voice parameter as a continuous dependent variable. Emotions (Joy/Sadness/Anger/Shame/Pride), origin (Quebecer/French) and gender (Male/Female) were

entered as predictors (model <- lmer(F0M ~ Actor*Emo*Origin + (1|Actor) + (1|Emo)). We report here models that explain more than or equal to 65% of the variance ($r^2 \geq 0.65$).

RESULTS

On the eleven dependent variables that were analyzed, we found triple effects of the variables Emotions, Origin and Gender on the six following parameters: F0M ($X^2(4, N=12) = 58.8676, p < .001$); IntM ($X^2(4, N=12) = 62.5852, p < .001$); Duration ($X^2(4, N=12) = 65.2793, p < .001$); Spectral Slope ($X^2(4, N=12) = 99.9645, p < .001$); Jitter ($X^2(4, N=12) = 25.6681, p < .001$); Hammarberg index ($X^2(4, N=12) = 30.8974, p < .001$). These triple interactions allowed us to investigate the differences of gender and origin for each emotion. All the results presented below are displayed in **tables 1, 2 and 3**. Finally, we conducted permutation tests across participants of different origins, resulting in an absence of origin effects. These results confirm the influence of origin on our findings.

F0m ($r^2=0.8$) (Figure 1)

Regardless of gender, fundamental frequency (F0M) was significantly higher for French when expressing anger, ($\beta = 30.72, se = 11.5, t(8.74) = 2.678, p = 0.02$). Interactions between gender and emotions were also found. When expressing joy and pride, French men have a higher F0M than Quebec men (for joy, $\beta = -0.1571, se = 0.0616, t(14.9) = -2.553, p = 0.02$; for pride: $\beta = 0.1890, se = 0.0601, t(13.6) = 3.142, p < 0.01$) whereas Quebec men had a higher F0M than French men when expressing sadness ($\beta = -0.1668, se = 0.0643, t(17.7) = -2.595, p = 0.0185$).

As for women, French women use a significantly higher F0M than the Quebec women when expressing pride ($\beta = 0.1890, se = 0.0601, t(13.6) = 3.142, p < 0.01$) and shame ($\beta = 0.0316, se = 0.113, t(5.55) = 0.281, p = 0.7890$). No significant differences were found for other emotions (see statistics in tables 2, 3 and 4).

IntM $r^2=0.7$ (Figure 2)

Overall, Quebecers spoke with a higher mean Intensity (IntM) compared to French when expressing pride ($\beta = -3.81$, $se = 1.72$, $t(8.71) = -2.222$, $p=0.05$) and shame ($\beta = -5.86$, $se=1.74$, $t(9.32) = -3.358$, $p<0.01$). Interactions between gender and emotions were also found. When expressing joy, Quebec men expressed themselves with greater IntM than French men ($\beta = -6.423$, $se= 2.10$, $t(4.54) = -3.057$, $p= 0.03$), pride ($\beta = -5.723$, $se= 2.10$, $t(4.49) = -2.731$, $p=0.05$) and shame ($\beta = -5.536$, $se= 2.14$, $t(4.91) = -2.583$, $p=0.05$). No other differences were found for other emotions (see statistics in tables 2, 3 and 4).

Duration $r^2=0.68$ (Figure 3)

Overall, Quebecers took longer than French to express sadness ($\beta = -0.2370$, $se= 0.0693$, $t(17.4) = -3.420$, $p < 0.01$) and shame ($\beta = -0.2370$, $se= 0.0693$, $t(17.4) = -3.420$, $p < 0.01$). Interactions between gender and emotions were also found. Quebec men have a longer duration in the expression of joy ($\beta = -0.1571$, $se= 0.0615$, $t(14.9) = -2.553$, $p= 0.02$), and sadness ($\beta = -0.1668$, $se=0.0643$, $t(17.7) = -2.595$, $p= 0.02$). On the other hand, French men took longer than Quebec men to express pride ($\beta = 0.1890$, $se= 0.0601$, $t(13.7) = 3.143$, $p < 0.01$). Quebec women had a longer duration than French women in the expression of pride ($\beta = -0.3309$, $se= 0.112$, $t(5.36) = -2.967$, $p=0.03$) and shame ($\beta = -0.3656$, $se= 0.118$, $t(6.79) = -3.088$, $p= 0.02$).

Spectral Slope $r^2=0.8$ (Figure 4)

Analyses demonstrated an effect of origin as the Spectral Slope was higher for Quebecers than for French when expressing joy ($\beta = -3.677$, $se=1.66$, $t(9.12) = -2.214$, $p= 0.05$), sadness ($\beta = 4.153$, $se=1.67$, $t(9.40) = -2.482$, $p= 0.03$), pride ($\beta = -4.889$, $se= 1.66$, $t(9.00) = -2.954$, $p= 0.01$) and shame ($\beta = -6.052$, $se=1.70$, $t(9.89) = -3.570$, $p< 0.01$). There were also interactions between emotions and gender as Quebec men have a larger Spectral Slope than French men in the expression of joy ($\beta = -9.488$, $se= 2.31$, $t(4.55) = -4.106$, $p= 0.01$), sadness ($\beta = -6.381$, $se= 2.32$, $t(4.65) = -2.746$, $p= 0.04$), pride ($\beta = -8.711$, $se= 2.30$, $t(4.50) = -3.780$, $p=0.0157$) and shame (β

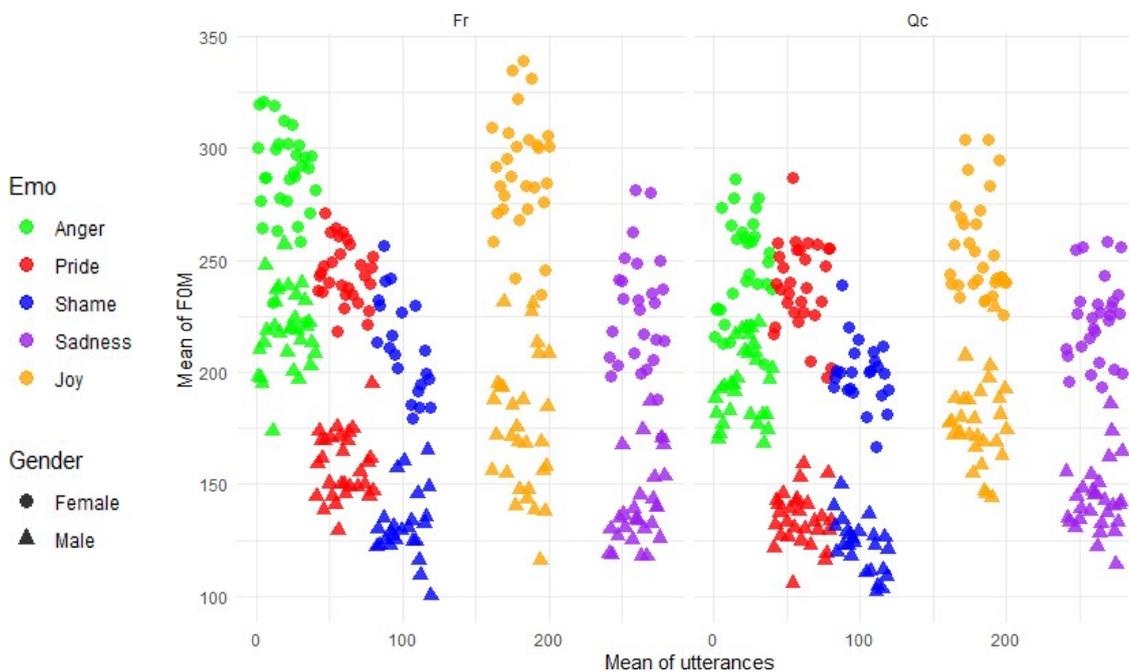
=-5.699 se=2.36, $t(4.94)=-2.416$, $p=0.0610$). When expressing shame, Quebec women shows a greater Spectral Slope than French women ($\beta = -6.334$, se= 2.45, $t(5.01)= -2.588$, $p= 0.05$).

Hammarberg index $r^2=0.7$ (Figure 5)

Quebec men have a higher Hammarberg index than French men when expressing anger ($\beta = 4.36$, se= 1.39, $t(5.71)= 3.131$, $p= 0.02$), joy ($\beta = 4.53$, se= 1.42, $t(6.21)= 3.191$, $p= 0.02$), sadness ($\beta = 4.67$, se= 1.45, $t(6.65)= 3.233$, $p= 0.01$) pride ($\beta = 5.69$, se= 1.41, $t(6.00)= 4.037$, $p < 0.01$). On the other hand, French men have a higher Hammarberg index in the expression of shame ($\beta = -3.31$, se= 1.51, $t(7.96)= -2.185$), $p=0.0605$.

Figure 1:

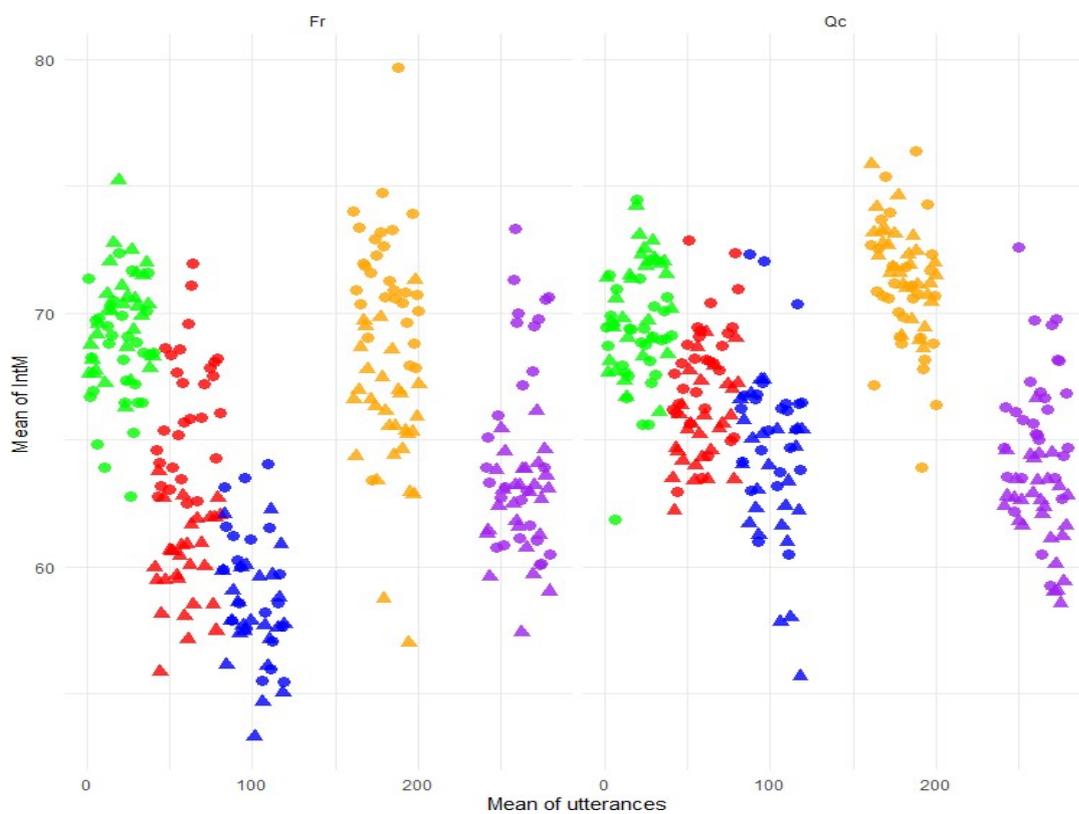
F0M for emotions and gender between Qc and Fr



Note: Mean of fundamental frequency (F0M) as a function of emotions (anger, pride, shame, sadness and joy) and gender (female and male) and Fr (left panel) and Qc (right pane) of the speakers.

Figure 2

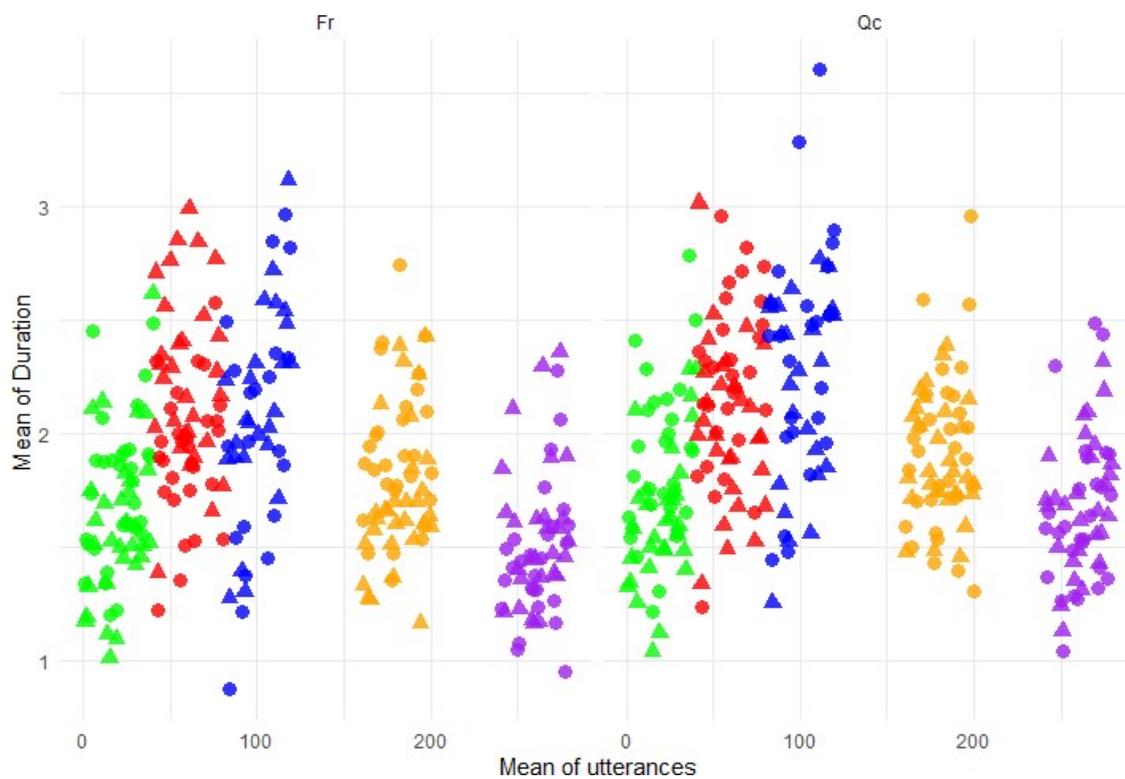
F0M for emotions and gender between Qc and Fr



Note: Mean of intensity (IntM) as a function of emotions (anger, pride, shame, sadness and joy) and gender (female and male) and Fr (left panel) and Qc (right pane) of the speakers.

Figure 3

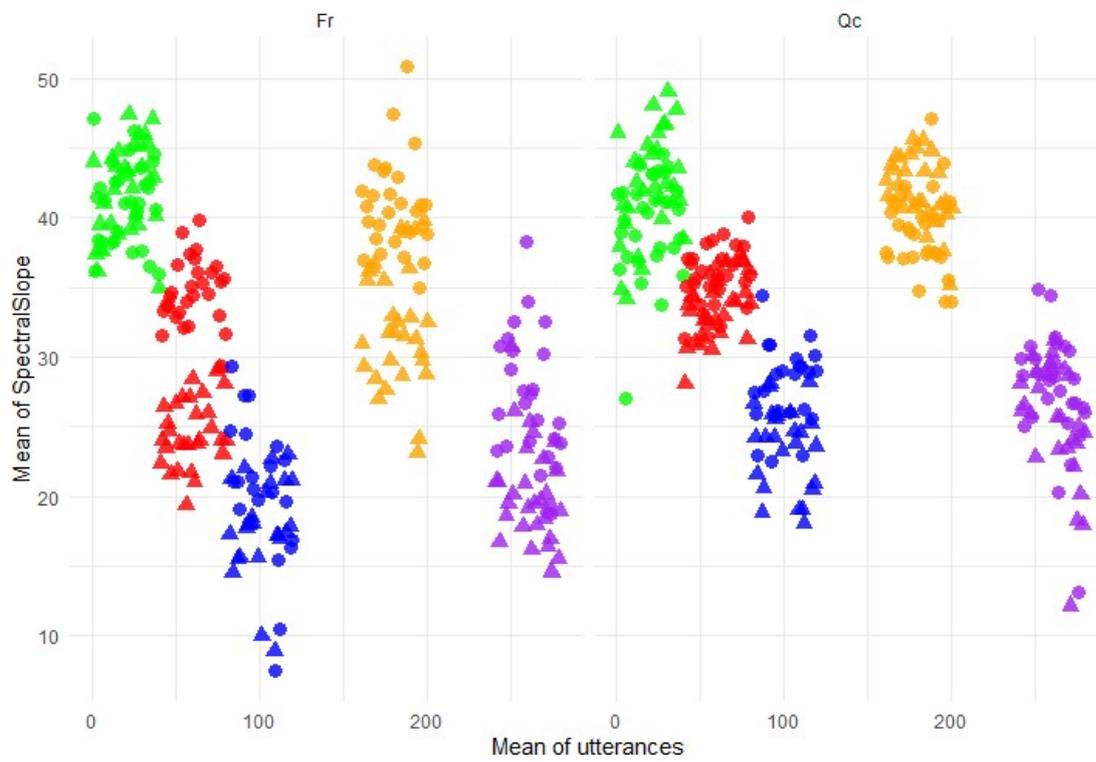
Duration for emotions and gender between Qc and Fr



Note: Mean of Duration as a function of emotions (anger, pride, shame, sadness and joy) and gender (female and male) and Fr (left panel) and Qc (right pane) of the speakers.

Figure 4

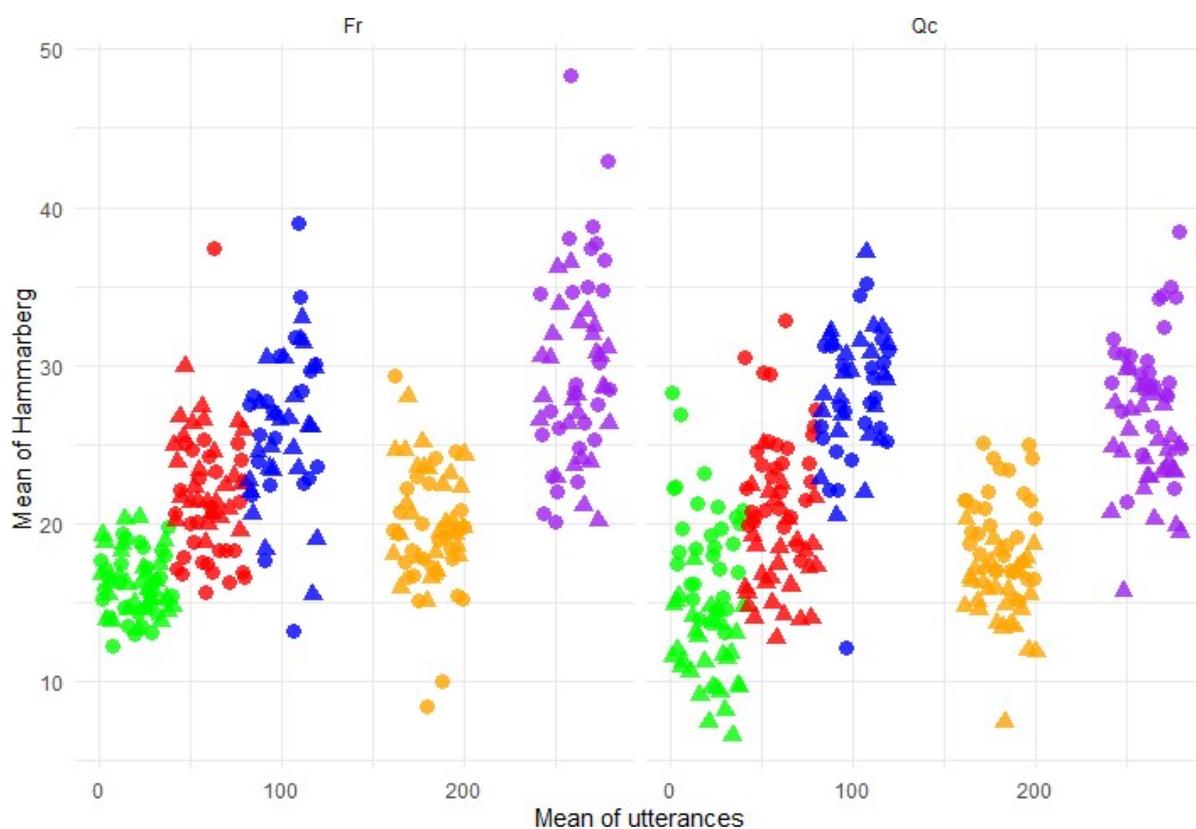
Spectral Slope for emotions and gender between Qc and Fr



Note: Mean of Spectral Slope as a function of emotions (anger, pride, shame, sadness and joy) and gender (female and male) and Fr (left panel) and Qc (right pane) of the speakers.

Figure 5

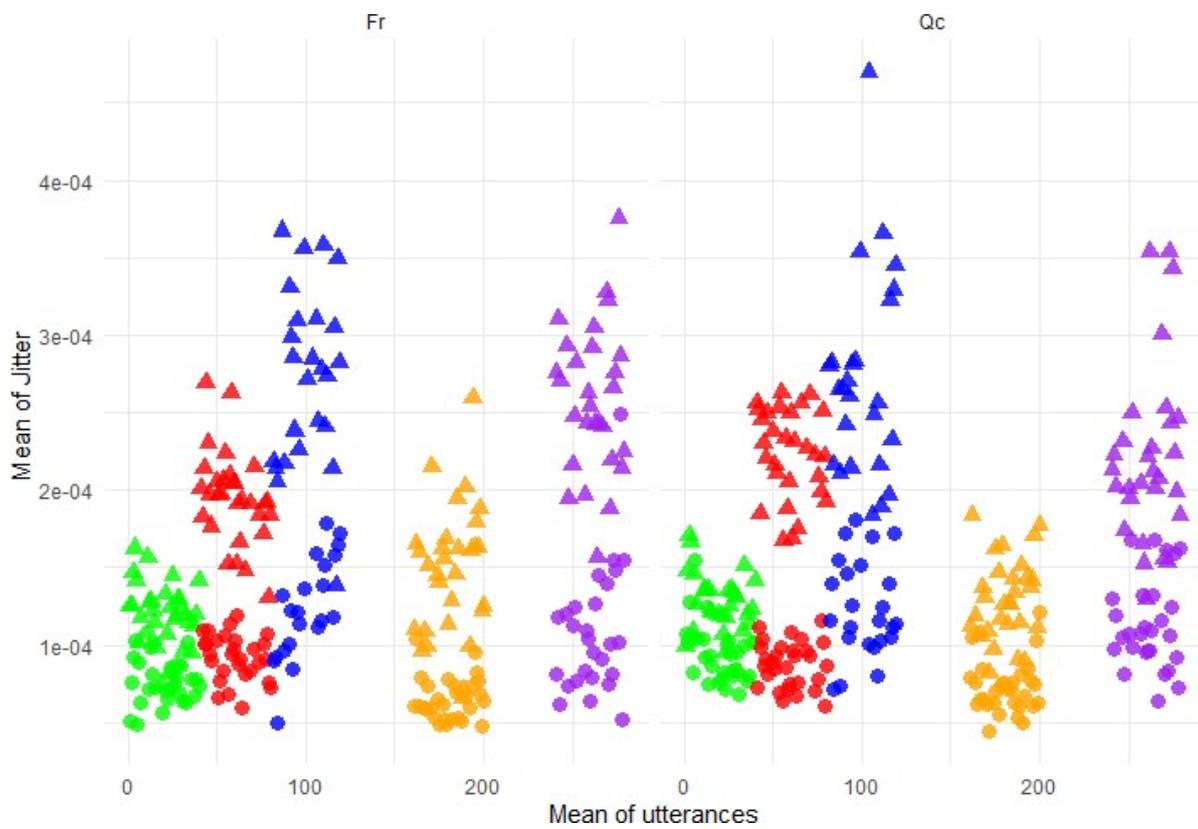
Hammarberg index for emotions and gender between Qc and Fr



Note: Mean of Hammarberg index as a function of emotions (anger, pride, shame, sadness and joy) and gender (female and male) and Fr (left panel) and Qc (right pane) of the speakers.

Figure 6

Jitter for emotions and gender between Qc and Fr



Note: Mean of Jitter as a function of emotions (anger, pride, shame, sadness and joy) and gender (female and male) and Fr (left panel) and Qc (right pane) of the speakers.

Table 2

Comparison Qc-Fr of vocal parameters per emotionsF

	Joy	Pride	Anger	Sadness	Shame
F0M			p=0.02		
IntM		p=0.05			p<0.01
Duration				p<0.01	p<0.01
SpectralSlope	p=0.05	p=0.01		p=0.03	p<0.01
Hammarberg					

Table 3

Comparison of men Qc-Fr of vocal parameters per emotions

	Men				
	Joy	Pride	Anger	Sadness	Shame
F0M	p=0.02	p< 0.01		p=0.0185	
IntM	p=0.03	p=0.05			p=0.05
Duration	p=0.02	p<0.01		p=0.02	
SpectralSlope	p=0.01	p=0.0157		p=0.04	p=0.0610
Hammarberg	p=0.02	p<0.01	p= 0.02	p=0.01	p=0.0605

Table 4

Comparaison of women Qc-Fr of vocal parameters per emotions

	Women				
	Joy	Pride	Anger	Sadness	Shame
F0M		p<0.01			p=0.7890
IntM					
Duration		p=0.03		p=0.02	
SpectralSlope					p=0.05
Hammarberg					

Discussion

The aim of this study was to create and validate a bank of emotional phrases in French, pronounced with regional French and Quebec standard accents. In addition, we wanted to characterize the emotional acoustic profiles between Quebecers and French. We found that the parameters classically considered most relevant for emotion discrimination (e.g., Sobin and Alpert, 1999; Scherer et al., 2015; Scharenborg et al., 2018; Ekbert et al., 2023), i.e., F0m, intensity, tempo and spectral balance (including Hammarberg index and Spectral Slope) were modulated by the French and Quebec cultures, as well as between the genders within these two cultures. We validated the hypothesis that, overall, Quebecers have a more expressive emotional prosodic profile than the French. However, this trend is reversed for anger. We also demonstrated an important influence of gender, with Québécois men generally being more expressive in their voices than French men, whereas we found, specifically for women, acoustic differences for shame and pride. Our results are the first to establish acoustic profiles of a range of emotions between different regional accents in French, including new data by integrating

two more social and cultural emotions: shame and pride. We will discuss the acoustic profiles of these emotions and interpret the influence of culture and gender in the following sections.

1. Cross-cultural emotional prosody

1.1. Expression and recognition of cross-cultural emotional prosody

In order to validate our stimulus bank, we administered an in-group validation study in which participants had to choose the emotion they were listening to from among 9 choices. We obtained a validation percentage of 95.8% for anger, 93.5% for sadness, 88.7% for pride, 87.3% for joy and 60.7% for shame. These observations are consistent with the study by Laukka et al. (2016) and Cowen et al. (2019). In the first study, the authors created and validated a stimulus database in order to study the production and perception of intercultural prosodic emotions. This database consisted of sentences spoken in English by actors from five different countries: Australia, India, Kenya, US and Singapore in 11 emotions, including basic emotions (such as joy, sadness and anger) and more social emotions (such as pride and shame). The sentences were validated by listeners in each country (in-group + out-group). Anger was the best emotion recognized, followed by sadness (3rd) and joy (7th). Pride (10th) and shame (11th) were the least recognized emotions. In Cowen et al, (2019), the authors asked a large number of Americans and Indians (2,345) to judge 53 emotions (e.g., surprise, dominance, distress) produced by a hundred actors from the same five different countries as in Laukka et al., 2016. The authors then ranked the emotions best recognized by the participants. Sadness was the first emotion best recognized (1st/53), followed by anger (3rd/53), but pride (35th/53) and especially shame (52nd/53) were among the least recognized.

These results therefore suggest that there are cross-cultural and intra-cultural differences in the recognition of prosody according to different emotions. The results of our study are therefore in line with the literature, which is still fairly poor on cross-cultural prosodic differences, considering a wide range of emotions. Taken together, we confirm differences in emotion

production and recognition across cultures and even at regional level, as we show between the French and the Quebecois. To explore this phenomenon, acoustic analyses were carried out.

2. Cultural comparaisons: France-Québec

2.1. Hofstede's model

Our data suggest that Quebecers have more expressive emotional prosody than the French (higher IntM, longer duration, steeper spectral slope and higher Hammarberg index for men) except for anger. These results could be explained by four dimensions of Hofstede's model (1980). Scores are measured using survey. First, the Power Distance Index (54 Qc -68Fr) measures the acceptance of power hierarchy and authority. France scored higher than Quebec, indicating a culture that values emotional discretion in public. Second, the Uncertainty Avoidance Index (Qc:60-Fr:86) measures the extent to which a culture feels threatened by unknown situations. The French score higher, leading to more intense expression of anger through prosody. Third, the Individualism or Collectivism scale (74 Qc - 71 Fr) shows a similar score for both cultures, indicating individualistic cultures that prioritize self-fulfillment. That would explain why Quebecers express pride more strongly. Last, the Indulgence scale shows that Canada (rate 68) has a more indulgent culture than France (rate 48), leading to more intense expression of positive emotions by the Quebecers. Altogether, emotions are therefore not fixed concepts but are part of social and cultural contexts. An indicator of these emotional changes can be found in prosody. Another factor at the origin of the differences between Fr and Qc could be pre-school education, which is thought to have an impact on how children learn to interact with others and express their emotions, depending on the culture (Brougère, 2010, Audrin, 2020). In Quebec, for example, children are encouraged to develop social skills from an early age, which can lead to greater openness and greater ease in expressing their emotions and opinions. In France, where the emphasis is more on academic preparation in primary school, children may focus more on intellectual learning rather than social skills.

So, even through the voice, emotions are more likely to be sculpted by culture. Emotions such as pride and shame are more likely to emerge from the cultural experience of the country.

2.2. Shame and pride

Shame and pride are very little studied. In our present research, these are the two emotions to show the most significant differences in prosodic profile between the French and the Quebecers, which is striking since they are the two most social and cultural emotions in our study. More specifically, using sound parameters, we show that differences are particularly marked for shame and pride, even within the same language and similar cultures. These observations are supported in terms of perception by Cowen et al, 2019 and Laukka et al (2016) showing that pride and shame are the least well recognised emotions. Also in the Laukka et al (2019) study, the emotions of shame and pride were the only two that were more often misclassified than correctly classified. Pride was often confused with neutrality and interest, while shame was confused with sadness and neutrality. Interestingly, these errors only occurred when evaluating people from another culture, which was not the case for pride, which was also confused by individuals from their own culture. These results support our study, which highlights the fact that there are considerable cultural differences in the emotional prosodic schema of certain emotions, which would appear to be more sculpted by culture. These observations could be explained with Tangney et al. (2012) theory of self-conscious emotions. Self-conscious emotions develop throughout our lives in response to societal influences and vary across cultures. They play a role in guiding moral and socially appropriate behavior. Pride and shame according to the authors are self-conscious emotions shaped by self-representation in relation to society. Since Quebecers and French do have major cultural differences, this would explain why we found the largest differences in these emotions. Importantly, we found differences only between French and Quebec women in these two emotions.

3. Gender comparisons

A large proportion of our results for all emotions show differences between French men and Quebec men. In general, we found a greater average intensity (IntM) (Joy, Pride and Shame) and a longer duration (Joy, Sadness) as well as a flatter spectral slope (Joy, Sadness; Pride and Shame). It seems that Quebec men express emotions more through prosody. Note that French men have a lower fundamental frequency than Quebec men for sadness, whereas the opposite is observed for the other four emotions. The intercultural differences observed may be influenced by cultural expectations and gender stereotypes.

Indeed, Quebec and France do not have the same approach to emotions. For example, in France, there is a more traditional social norm, according to which men should not express their emotions excessively (Anxo et al., 2011), which would lead them to adopt a more reserved behaviour and therefore repress emotions. As for Quebec, the expression of emotions is considered natural and valued, regardless of gender (Roy et al. 2014), which would explain why we have no differences between women except for the duration (Shame and Pride) and the spectral slope (Shame). Quebecers would feel more comfortable expressing their emotions, especially negative emotions such as sadness or shame.

Conclusion

In conclusion, our study is the first to show results on the physical parameters of sound on the differences in the expression of emotional prosody between the neighboring cultures French and Québécois. First of all, Québécois and French do not have the same emotional acoustic profiles. In general, the French have a higher F0M while the Québécois speak with greater intensity. We also observed that Quebecers express more emotions through their prosodies, except for anger. There were also a number of significant differences in pride and shame, which were the only emotions to differ between Quebec and French women. Our results therefore suggest that the emotional expression of French and Quebec men differs significantly,

highlighting the influence of cultural gender norms. Overall, these initial results point to interesting findings that need to be explored with more actors and in ecological contexts in order to be able to generalise these conclusions. Our results confirm, in line with other studies, that prosody, especially with emotional content, could be a strong cultural marker governed by the historical, mores, educational and political aspects of the societies in which people live.

Chapitre III. Discussion générale

1. Synthèse des résultats

La présente étude visait à créer une banque de stimuli afin de comprendre la production de la prosodie émotionnelle selon les accents français et québécois. Sur la base des propositions montrant un avantage de la reconnaissance des émotions chez des participants du même groupe ethnique (Elfenbein et Ambady, 2002) nous pensions que la prosodie émotionnelle des Français et Québécois présenteraient des profils acoustiques similaires mais avec des différences notables. En référence à la recherche de Mauchand et Pell (2020) nous avons émis l'hypothèse que les québécois présenteraient des indices acoustiques plus intenses (des plus grandes valeurs) que les Français.

Dans un premier temps nous avons créé une banque de stimuli étant de courtes phrases dans la langue française prononcées dans des émotions basiques: joie, tristesse et colère, ainsi que des émotions plus complexes : la fierté, et la honte. Elles ont été enregistrées par 12 comédiens professionnelles Québécois et Français (également répartis entre les sexes). Puis, une étude de validation en ligne a été menée afin de sélectionner les meilleures phrases pour notre banque de stimuli par des participants intra-groupe. Dans un second temps, nous avons analysé les paramètres vocaux des acteurs des phrases validées afin de comprendre s'il existe des différences en termes d'indices acoustiques sur la prosodie émotionnelle entre accents

régionaux différents. Nous avons donc analysé 8 paramètres vocaux (F0, Intensité, Durée, Jitter, Shimmer, HNR, Pente spectrale, indice Hammarberg).

Nos résultats montrent que les Français et les Québécois semblent exprimer les émotions avec des patrons acoustiques similaires (indices acoustiques émotionnels universels), mais des différences sont tout de même notables. Nos données appuient les affirmations de la littérature sur ce sujet (Laukka et al., 2016). Plus particulièrement, nous avons trouvé des différences saillantes dans les cinq émotions étudiées. De manière générale, nous constatons que les valeurs des paramètres IntM, durée, pente spectrale des Québécois sont plus élevées que les Français. À l'inverse, les Français montrent des valeurs de F0M plus élevées sauf au niveau de la tristesse, l'indice Hammarberg est aussi plus élevé sauf au niveau de la honte en comparant les hommes des deux cultures. Pour finir, les émotions de fierté et honte sont celles qui ont révélé le plus de résultats significatifs, et notamment entre les femmes Québécoises et Françaises.

2. Expression cross-culturelle de la prosodie émotionnelle

2.1. Méthodologie des banques de stimuli

Afin d'étudier la prosodie émotionnelle, différentes banques de stimuli ont vu le jour au fil des années dans différents langages tels que l'anglais avec la banque RAVDESS (Livingstone et Russo, 2018), l'Allemand (EMO-DB (Burkhardt et al., 2005), le Turc (TurEV-DB (Canpolat, 2020), ou encore le Portugais (Castro et Lima, 2010). La création d'un corpus de textes prosodiques émotionnels est un processus qui prend du temps et nécessite plusieurs choix méthodologiques importants.

2.1.1. Choix du matériel. En ce qui concerne la création des phrases, certains utilisent des pseudo-phrases afin de ne pas avoir le sens qui interfère, c'est à dire minimiser le biais liés aux phrases elle-même et d'examiner les caractéristiques acoustiques et perceptuelles de la prosodie sans autre influence. (voir : Navas et al., 2006; Paulmann et al., 2013 ; Gong et al., 2023). Par

exemple Brennan et al. (2011), utilisent des unités de langage (**speech units**) telles que des phonotactiques commun, ou bien Castro et Lima (2010) utilisent des pseudo-phrases étant des suites de mots n'ayant aucun sens mais qui suivent la syntaxe, la morphologie et les phonotactiques de leur langue (le Portugais). D'autres, utilisent des phrases avec une sémantique (ex : Rao et al., 2013 ; Vicsi et Szaszak, 2010 ; Mauchand et Pell, 2020) collant avec l'émotion voulue (ex: des scénarios ou des situations) (comme notamment : Keshtiari et al., 2015; Laukka et al., 2016 ; Busso et al., 2017). Cela aiderait le locuteur à exprimer l'émotion souhaitée naturellement. C'est également plus représentatif de la vie de tous les jours, dans laquelle nous nous exprimons très fréquemment avec émotion. En fin de compte, les banques de phrases de prosodie émotionnelle avec un contenu sémantique ou neutre peuvent servir de ressources pour l'étude et la détection des émotions. Cependant, il est important de prendre en compte les différences individuelles et culturelles dans l'expression et la perception des émotions. Par exemple, dans le cadre d'une banque d'accent régional, les différentes utilisations langagières doivent être prises en compte lors de la création des phrases car certains mots n'ont pas la même signification ou des mots différents sont utiliser pour une même signification. Cela arrive beaucoup entre le langage français et québécois par exemple.

2.1.2. Construction des phrases. Lors de la construction des phrases, il est important de contrôler leurs structures afin d'éviter les biais. Par exemple, il est préférable d'utiliser une variabilité de types de phrases comme : exclamative, déclarative, question etc. (Saratxaga et al., 2006 ; Gong et al., 2023). Il est aussi important de contrôler le nombre de mots et syllabes ainsi que le champ lexical, le type de vocabulaire (Saratxaga et al., 2006). La variabilité de l'intensité d'une émotion est également à prendre en compte lors de la création d'un corpus. En effet, une émotion s'exprime dans notre quotidien selon différente intensité allant de faible à forte (Gong et al., 2022). Par exemple, la banque audio-visuelle RAVDESS de Livingstone et Russo (2018)

est basée sur deux phrases émotionnelles avec 6 émotions primaires dans deux niveaux d'intensité (normal et élevé).

La méta-analyse de Diconne et al. (2022) présente KADOPI, un stockage de multiples banques de stimuli émotionnelles en science ouverte. Les auteurs expliquent comment les différentes banques sont créées et amènent des limites globales. Notamment, ils avancent qu'il n'y ait pas une définition claire d'émotion par la communauté scientifique. Ils amènent également des recommandations pour les études futures de création de banque de stimuli. Par exemple, la validation des stimuli devrait suivre un échantillonnage important avec une sélection bien justifiées, c'est-à-dire que les participants répondant à l'évaluation de ces stimuli doivent avoir les mêmes critères spécifiques que les groupes cibles de la banque.

2.1.3. Choix d'enregistrement. En ce qui concerne la façon d'enregistrer les corpus, certains utilisent des locuteurs naïfs (non professionnel) afin de capter une émotion la plus authentique possible. Parfois les locuteurs sont mis dans une situation spécifique, les poussant à réagir avec une émotion (Saratxaga et al., 2006 ; Busso et al., 2017). D'autres utilisent des enregistrements avec des locuteurs spontanés comme des enregistrements dans des situations naturelles (Busso et al, 2017), ou des enregistrements venant de la télévision (Grimm et al., 2008). Les limites de ces méthodes sont que l'énoncé d'une émotion spontanée peut se mélanger avec d'autres émotions (Pravena et Govind, 2017). D'autres études utilisent des acteurs professionnels car ils ont l'habitude de manipuler et simuler les émotions (ex: Livingston et Russo, 2018 ; Castro et Lima, 2010; Burkhardt et al., 2005). Avec cette technique, les émotions peuvent être davantage exagérées mais sont souvent mieux reconnues. Cette technique est aussi la seule qui permet un total contrôle expérimental (Saratxaga et al., 2006).

Différentes variétés de banques de prosodie émotionnelle existent. Par exemple, les chercheurs créent des banques avec différents types d'émotions. La majorité recensée est avec des émotions

primaires, des phrases ou de pseudo-phrases (Castro et Lima, 2010) et avec des accents étrangers. Par exemple, pour créer leurs stimuli, Paulmann et Uskul (2014) ont fait prononcer des pseudo-phrases dans les six émotions primaires à des Anglais et des Chinois. Afin d'étudier les caractéristiques acoustiques des voix françaises et québécoises dans l'expression émotionnelle, nous avons donc décidé de construire une banque de phrases de la langue française, prononcées avec des accents français et québécois. Il n'existe pas de banque de phrases de prosodie émotionnelle avec des accents régionaux différents. Dans un cas similaire au notre Mauchand et Pell (2020) ont enregistré 480 phrases émotionnelles de plainte et neutre par 8 acteurs québécois et français. En utilisant cette banque de données, les auteurs ont étudié comment la prosodie varie en fonction de l'émotion exprimée dans la plainte selon les accents régionaux Français et Québécois.

2.2. Validation des stimuli: reconnaissance des émotions prosodiques

Dans un premier temps, nous avons réalisé une étude de validation de nos stimuli. Nos phrases ont été validées par juges intra-groupe. C'est-à-dire que les Français écoutaient les Français et les Québécois écoutaient les Québécois. Nous avons obtenu un pourcentage de validation de 95.8% pour la colère, 93.5% pour la tristesse, 88.7% pour la fierté ; 87.3% pour la joie et 60.7% pour la honte. Nos observations vont dans le sens de Laukka et al. (2016), sauf pour la joie et la fierté qui se trouvent à des places interchangeables. En effet, Laukka et al. (2016) fournissent des informations intéressantes sur l'expression et la reconnaissance de la honte et de la fierté. Ils ont créé une banque de stimuli de phrases anglaises avec des acteurs australiens, indiens, kenyans, singapouriens et américains dans 11 émotions : joie, colère, tristesse, mépris, peur, bonheur, intérêt, désir, neutralité, fierté, soulagement, honte et orgueil. Dans l'étude de validation similaire à la nôtre, afin de ne retenir que les meilleures phrases reconnues, chaque participant de chaque culture a évalué la reconnaissance des émotions de sa propre culture et de celles des autres de manière égale. Parmi les 11 émotions étudiées, la colère et la tristesse

sont les mieux reconnus, puis la fierté et la honte sont les deux dernières émotions du classement. Les auteurs interprètent que les émotions sont exprimées différemment d'une culture à l'autre, cependant nous obtenons les mêmes résultats alors que les évaluations des phrases émotionnelles était réalisée par des personnes de la même culture. Il existerait donc des différences interculturelle et intra-culturelle dans la reconnaissance de la prosodie émotionnelle. Il est donc pertinent d'aller regarder au niveau des caractéristiques sonores pour comprendre ces différences.

2.2.1. Paramètres vocaux: profils interculturels. Comme nous, Laukka et al. (2016) ont analysé les paramètres acoustiques de chaque émotion et ont trouvé des différences émotionnelles sur certains paramètres vocaux. Par exemple, ils ont observé que les valeurs de F0M, IntM et l'indice Hammarberg étaient différent lors des expressions de la joie en comparant les cinq pays. En ce qui nous concerne, dans les expressions prosodiques des Québécois et Français, nous avons aussi observé des différences au niveau de F0, de IntM et de l'indice Hammarberg mais aussi au niveau de la durée et de la pente spectrale. Au niveau de la honte et de la colère nous observons des différences au niveau du F0M et de l'indice Hammarberg et IntM (juste pour la colère).

Certains paramètres tels que F0M, indice Hammarberg et IntM, ainsi que des paramètres de la qualité de la voix semblent ainsi avoir un rôle clé dans la différenciation d'expression d'émotions entre culture. D'autre part, notons dans leur étude, la présence de valeurs significativement différentes dans les profils acoustiques pour chaque émotion et chaque culture. Par exemple pour l'expression de la joie, l'IntM n'est pas au même niveau entre les locuteurs des US et du Kenya, alors qu'il n'y a pas de différences avec l'Australie, Singapour et l'Inde. L'indice Hammarberg diffère uniquement entre le Kenya, l'Australie et les États-Unis.

Ainsi, ce que semble montrer nos deux études mises en commun est qu'il existerait bien des différences au niveau des profils acoustiques lors de l'expression prosodique entre chaque pays et même entre accent régionaux.

2.2.2. Profils acoustiques des émotions. Si l'on examine le diagramme des paramètres acoustiques des émotions dans la littérature (e.g., Rijn et Larrouy-Maestri, 2023; Laukka et al., 2016), les émotions suivent des schémas acoustiques distincts. La colère et la joie se caractérisent par une augmentation de la fréquence fondamentale ($F0m$), c'est-à-dire que la voix devient plus aiguë, l'intensité est plus forte et la durée plus courte. L'indice de Hammarberg est élevé, ce qui confère à la voix puissance et projection. La pente spectrale est plus plate, ce qui conduit à une perception de rugosité dans la voix (par exemple, Scherer, 1996 ; Banse et Scherer 1996 ; Zhang 2006; Ekberg et al., 2023). À l'inverse, les émotions telles que la peur ou la tristesse se caractérisent par une baisse de $F0m$, de l'intensité et de la durée (par exemple, Yuan et al., 2002), ainsi que par une diminution de l'indice de Hammarberg (Ekberg et al., 2023) et une pente spectrale plus raide (Guzmann et al., 2013). Ce schéma conduit à une perception auditive dans laquelle les hautes fréquences sont prédominantes et donne l'impression d'une voix claire, lumineuse et facilement perceptible (Kuang et Liberman, 2015).

La littérature montre des schémas d'émotions de manière globale. Les études transculturelles racontent une autre histoire. Il existe en effet des différences dans les schémas émotionnels propres à chaque culture, même lorsque la même langue est parlée mais avec des accents régionaux différents. C'est ce que les résultats de notre étude actuelle semblent montrer avec les Québécois et les Français, deux peuples parlant la même langue mais avec des accents très distincts et une culture relativement voisine.

3. Cultural aspect of prosody: Québec-France

3.1. Les échelles de Hofstede

Nos résultats montrent que les Québécois suivent un profil acoustique avec une plus grande intensité, une plus longue durée et une pente spectrale plus courte par rapport aux Français, ce qui conduit à une perception d'une expression plus intense des émotions. Les différences dans la façon dont les émotions sont exprimées par la voix sont façonnées par la culture, qui module la façon dont les émotions sont communiquées. Selon le modèle de Hofstede, six dimensions (Distance hiérarchique, Individualisme/Collectivisme, Masculinité/Féminité, Evitement d'incertitude, Orientation à long-terme/court-terme, Indulgence/Retenuet) permettent de comparer les pays. Quatre d'entre elles pourraient expliquer les différences observées. Pour commencer, ces résultats pourraient être expliqués par l'indice de distance de pouvoir, (54 Qc -68 Fr) qui mesure le degré auquel un pays accepte la hiérarchie de pouvoir et l'autorité. Les pays ayant un score élevé expriment moins d'émotions (Fernandez, 2000). La France a obtenu un score plus élevé que le Québec, ce qui signifie que ce dernier a une culture qui favorise l'interdépendance avec une faible distance émotionnelle par rapport à la France. En outre, la France est une culture de discrétion émotionnelle dans laquelle les individus évitent d'exprimer leurs émotions ouvertement et intensément en public afin de préserver leur estime de soi et leur réputation sociale. En effet, selon Gross et John (2003), les individus issus de cultures telles que la France, où l'image publique est valorisée, ont tendance à être plus conscients et préoccupés par leur expression émotionnelle en public que les individus issus de cultures où cela n'est pas aussi important, comme le Québec. Ce serait une explication plausible du fait que les Québécois ont une prosodie émotionnelle plus expressive que les Français.

De même, sur l'échelle de l'indulgence, le Canada a un score de 68 et la France de 48 (aucun score disponible pour les Québécois), ce qui signifie que le Canada a une culture plus indulgente que les Français. Cela se traduit par une attitude plus positive et optimiste, favorisant la

jouissance de la vie. Cela pourrait expliquer pourquoi les Québécois expriment plus intensément des émotions positives telles que la joie et la fierté.

Enfin, l'échelle Individualisme ou Collectivisme (74: Qc – 71: Fr) détermine le degré de rattachement des individus aux groupes. Les deux cultures ont un score similaire, indiquant une culture individualiste, se traduisant par une priorité de s'occuper de soi et de sa famille proche. L'épanouissement personnel est important, notamment à travers la réussite professionnelle, l'embauche et la promotion étant régies par le mérite. Cette culture individualiste est légèrement plus élevée chez les Québécois, ce qui pourrait expliquer le fait qu'ils expriment plus fortement leur fierté, afin de montrer leur réussite personnelle à la société. D'autre part, lors de l'expression de la colère, il a été observé que la F0m est plus élevée chez les Français que chez les Québécois. L'augmentation de F0m est reconnue par la littérature comme l'un des principaux indices de la perception de la colère (par exemple, Scherer, 1996 ; Banse et Scherer, 1996, Zhang 2006). Il a été constaté que les hommes québécois obtiennent des scores plus faibles que les hommes français pour l'indice d'Hammarberg, qui est également caractéristique de la colère. Il convient toutefois de noter que l'utilisation de l'indice de Hammarberg est assez récente dans la littérature et qu'il reste encore beaucoup à apprendre sur ce paramètre. Ces résultats pourraient s'expliquer par l'indice d'évitement de l'incertitude (Qc: 60 - Fr: 86) qui mesure dans quelle mesure les membres d'une culture se sentent menacés par des situations ambiguës ou inconnues, les Français obtenant un score beaucoup plus élevé que les Québécois. Les pays ayant un score élevé connaissent plus de stress, de conflits entre les normes et les expériences et ont peu de contrôle sur l'intériorisation de leurs émotions négatives (Hofstede, 1991; Gudykunst & Nishida, 1994 ; Basabe et al.,1999). Cela expliquerait que les Français expriment plus intensément leur colère par le biais de la prosodie. En fait, les Québécois accepteraient volontiers de nouvelles idées et essaieraient de nouvelles choses, telles que de nouvelles technologies ou pratiques commerciales. Cela témoigne également d'une tolérance à l'égard des

idées ou des opinions d'autrui. À l'inverse, les Français n'aiment pas les surprises et ont besoin de structure et de planification. Ils aiment avoir toutes les informations en main et développer de nouvelles idées dans un environnement stable.

En résumé, les scores obtenus sur ces différentes échelles culturelles comparatives nous montrent que le Québec et la France sont des cultures similaires avec quelques différences notables. En raison de la combinaison d'un score élevé à l'indice d'évitement de l'incertitude, de l'individualisme et de la distance de pouvoir, les Français auraient besoin d'un filet de sécurité émotionnel. Les Québécois sont plus enclins à s'autoriser à exprimer leurs émotions en général. Divers facteurs façonnent ces normes sociales comme l'éducation.

3.2. Education

L'éducation est fondamentale pour l'expression et la répression des comportements. En effet, l'éducation dès le plus jeune âge joue un rôle dans ces différences interculturelles. Au Québec, le programme d'éducation préscolaire est considéré comme un moyen d'enrichir la capacité des enfants à interagir avec les autres et avec leur environnement. En revanche, l'école maternelle française adopte des objectifs de préparation à l'école primaire ainsi qu'un complément d'éducation familiale (Brougères, 2023), qui est donc beaucoup moins axé sur la gestion des émotions et la communication. Cela pourrait également contribuer à expliquer pourquoi les Québécois et les Québécoises semblent exprimer leurs émotions davantage par la voix. Ils seraient encouragés à exprimer leurs émotions dès le plus jeune âge.

3.3. Différences de prosodie selon le sexe

La grande majorité de nos résultats montrent des différences entre les prosodies des hommes français et des hommes québécois. En effet, les hommes québécois ont une FOM moins élevée que les hommes français, et l'inverse est vrai pour la tristesse. Il est important de noter ce résultat car, en général, les hommes québécois ont une FOM plus basse que les hommes

français, mais ce n'est que pour la tristesse que l'on trouve l'inverse. Il s'agirait donc d'un marqueur important dans l'expression prosodique de la tristesse chez les hommes français. Par ailleurs, l'intensité moyenne (IntM) est plus élevée chez les hommes québécois pour la joie, la fierté et la honte. Aussi, ils mettent plus de temps (durée) à exprimer la joie et la tristesse, et l'inverse est vrai pour la fierté. Ils ont également une pente spectrale plus plate pour toutes les émotions sauf la colère. Enfin, l'indice de Hammarberg est plus élevé pour toutes les émotions.

Les normes culturelles de genre jouent également un rôle dans ces différences d'expression émotionnelle. La France et le Québec ont des constructions sociales différentes en ce qui concerne les attentes à l'égard des hommes. En effet, les hommes québécois seraient moins contraints de respecter les normes traditionnelles de masculinité, ce qui leur permettrait de se sentir plus à l'aise pour exprimer leurs émotions (Roy et al. 2014) et notamment leurs émotions négatives. A l'inverse, les hommes français pourraient être plus attachés aux attentes traditionnelles de masculinité, ce qui pourrait les conduire à réprimer leurs émotions (Anxo et al., 2011).

4. Fierté et honte : Des émotions culturelles ?

4.1. Profils acoustiques de la honte et de la fierté

Nos résultats sur la honte et la fierté sont en accord avec l'une des rares études qui se sont penchées sur ces émotions prosodiques. En ce qui concerne le profil de la honte, nous avons observé des différences entre les deux cultures et les deux sexes. La honte a été peu étudiée en prosodie. Sulins et al. (2020) ont entrepris d'étudier les schémas prosodiques de la honte dans un contexte universitaire. En enregistrant les expériences des participants immédiatement après une expérience de honte et en analysant leurs paramètres vocaux, les auteurs ont observé que la prosodie de la honte était caractérisée par une fréquence fondamentale plus basse, une intensité plus faible et plus monotone. Nous avons observé que

les Français avaient des valeurs plus élevées dans le modèle de la honte que les Québécois pour à l'exception de F0m pour les femmes et de la pente spectrale.

Par rapport à la fierté, à notre connaissance, nous sommes les premiers à présenter des données sur la prosodie de la fierté en français. Nous montrons que chaque culture a des paramètres dominants dans le schéma de la fierté. Les Français montrent une fréquence fondamentale moyenne plus élevée, une plus grande durée, tandis que les Québécois montrent une plus grande intensité, une pente spectrale plus plate et un indice de Hammarberg plus petit que les Français. Nous avons décrit ici les profils acoustiques de la honte et de la fierté. Les études sur la prosodie émotionnelle ont très peu étudié les caractéristiques acoustiques de ces deux émotions, et encore moins les études interculturelles.

4.2. Expression et reconnaissance de la fierté et de la honte

Les études existantes montrent que ces émotions sont exprimées très différemment selon les cultures. Laukka et al. (2016), dans leur deuxième étude, visant à comprendre les effets de la perception des émotions sur la culture, ont présenté les phrases validées à de nouveaux participants issus des mêmes pays. Leur tâche consistait à reconnaître l'émotion qu'ils écoutaient parmi un choix de 11 émotions. La honte et la fierté sont les deux seules émotions qui ont été plus souvent mal classées que correctement classées. La fierté a souvent été confondue avec la neutralité et l'intérêt. La honte a été confondue avec la tristesse et la neutralité. Il est à noter que ces erreurs n'apparaissent que dans l'évaluation de l'out-group culturel, ce qui n'est pas le cas pour la fierté, qui est également confondue par les personnes de leur propre culture. Ces résultats vont dans le sens de notre étude, qui met en évidence le fait qu'il existe bel et bien des émotions dont le schéma prosodique émotionnel diffère fortement d'une culture à l'autre. Concrètement, nous montrons à l'aide de paramètres sonores que les différences sont plus prononcées pour la honte et la fierté, même à travers la même langue et des cultures apparentées. Laukka et al.

(2016) montrent ces différences dans la reconnaissance et la perception de la prosodie émotionnelle.

4.3. *Émotions de conscience de soi*

Selon Tangney et Tracy (2012), la fierté et la honte sont des émotions de conscience de soi, forgées par l'approbation sociale car nous avons peur de perdre notre statut social. Pour eux, la honte et la fierté font partie des émotions qui nous poussent à agir de manière morale et socialement appropriée. Ils décrivent un modèle pour expliquer leurs fondements: 1) Les émotions conscientes de soi nécessitent une représentation et une conscience de soi. 2) Elles apparaissent plus tard dans l'enfance que les émotions de base. 3) Elles facilitent la réalisation d'objectifs sociaux complexes. 4) Les émotions conscientes n'ont pas d'expressions faciales discrètes et universellement reconnues. 5) Elles sont cognitivement complexes. En bref, ces émotions conscientes de soi se construisent tout au long de notre vie, sont façonnées par la société qui nous entoure et diffèrent donc d'une culture à l'autre. Nos résultats soutiennent ce point de vue, en montrant que ces différences sont également perceptives, à travers la même langue, mais que les différences prosodiques dans les émotions agissent comme des marqueurs culturels.

4.4. *Fierté et honte : Différences entre les sexes*

Des différences entre les sexes et les pays ont également été observées. Il est important de noter que les seules différences de genre entre les femmes concernent les émotions prosodiques de la fierté et de la honte.

En ce qui concerne la fierté, les femmes françaises ont une fréquence fondamentale plus élevée, ce qui renvoie à une voix plus aiguë. De plus, nous avons noté que les femmes françaises expriment la fierté plus rapidement que les femmes québécoises et que les hommes français expriment la fierté plus longtemps que les hommes québécois.

De plus, les hommes québécois expriment la fierté avec une plus grande intensité, une pente spectrale plus raide et un indice de Hammarberg plus élevé. Un indice de Hammarberg élevé se traduit par une voix puissante et claire avec une forte projection et une pente spectrale raide apporte une perception de rugosité présente dans les émotions chaudes (Kuang et Libermann, 2015).

Dans l'ensemble, les Québécoises et les Québécois semblent avoir un profil de fierté acoustique plus intense que les Françaises et les Français. En revanche, des différences intéressantes ont été notées en termes de durée. C'est la seule émotion pour laquelle les hommes français s'expriment plus longtemps que les hommes québécois.

En ce qui concerne la honte, Français et françaises tendent plus vers le modèle de honte prosodique (expliqué ci-dessus) que les Québécois de par leurs valeurs acoustiques plus faibles (par exemple pour les hommes, $F0m : Qc = 200.27 - Fr = 214.64$; $IntM : Qc = 64.75 - Fr : 58.83$), Cependant, on peut aussi interpréter à l'inverse que les Québécois et Québécoises expriment plus intensément la honte. Les femmes et les hommes se sentiraient plus libres d'exprimer ce type d'émotion avec des connotations plus négatives en raison d'un contexte culturel et politique.

4.5. Contexte politique

La fierté et la honte sont les seules émotions qui révèlent des différences prosodiques entre les Québécoises et les Françaises. Pour expliquer ces résultats, on peut s'interroger sur la place des femmes dans ces deux cultures qui, selon Revillard (2008 et 2016), est « marginalement divergente ». Elle explique que ces différences s'expliquent en partie par le contexte politique et social de chaque pays. En France, le familialisme est profondément ancré dans l'État et soutenu par un puissant mouvement familial distinct du mouvement des femmes. Cette situation crée des tensions entre les défenseurs des femmes et les intérêts familiaux, ce qui limite les progrès en matière d'égalité professionnelle.

Au Québec, en revanche, la cause des femmes est institutionnalisée avant les intérêts familiaux. Cela a conduit à une plus grande collaboration entre le mouvement des femmes et le mouvement familial, ce qui a permis de réduire les relations conflictuelles. Les défenseurs des droits des femmes au Québec considèrent également les politiques familiales comme un moyen de promouvoir l'autonomie économique des femmes. Diverses stratégies gouvernementales ont donc été adoptées pour soutenir la cause des femmes (Leboucher, 2023).

Ainsi, les différences entre les émotions de honte et de fierté sont sous-tendues par l'essence même de ces émotions construites par les mœurs du pays. Cela se traduirait de manière importante dans la façon dont ces émotions sont exprimées par les hommes et les femmes françaises et québécoises.

5. Limites et futures recherches

De futures recherches pourraient être menées pour mieux comprendre les différences prosodiques que nous avons montrées. Quelques limites qui empêchent la généralisation de nos résultats sont d'abord à noter. Tout d'abord, la banque de stimuli n'a été validée que par des femmes du fait d'un manque important de participants hommes. Il serait donc pertinent de faire valider cette base de données aussi par les hommes. D'autre part, nous avons choisi d'enregistrer des comédiens dans un laboratoire en leur donnant des contextes à chaque phrase afin qu'ils expriment comme ils sentent les émotions. Il serait pertinent de comparer nos données avec non pas des acteurs, mais des enregistrements vocaux pris dans des moments quotidiens de la vie (Busso et al., 2017). Aussi, les futures recherches pourraient comparer avec des phrases sans contextes sémantiques, (Navas et al., 2006; Paulmann et al., 2013 ; Gong et al., 2023) afin de comprendre le rôle de la prosodie de manière isolée et totalement dissociée d'indice littéral. Par ailleurs, du fait d'un très grand nombre de stimuli à valider nous avons enlevé les émotions de la jalousie et de la surprise, préalablement prévues. De futures études pourraient reprendre et valider ces dernières données préliminaires.

D'autre part, nous avons pris des acteurs venant de Montréal et de Paris. Il serait pertinent de réaliser des études avec un plus vaste type d'accent québécois (ex: Mauricie, Saguenay, Gaspésie) et français (ex: Picardie (chti), Marseillais, Grenoble, Strasbourg) afin de comprendre d'avantage les différences d'expression et de perception de la prosodie émotionnelle des accents régionaux (Woehrling et de Mareüil, 2006 ; Boula de Mareüil, 2008 ; Puska et al., 2019).

Pour finir, des études en neuroimagerie devraient être réalisées afin de comprendre les mécanismes cérébraux sous-jacents aux phénomènes observés.

5.1. Etude EEG en cours

D'ailleurs, nous avons commencé notre étude en électroencéphalogramme, les données en cours d'acquisition feront l'objet d'un article futur. Notre objectif est d'étudier la différence de reconnaissance de la prosodie émotionnelle de la langue française selon des accents français et québécois au niveau cérébral. En suivant le modèle de Paulmann et Kotz (2011), nous pensons que les composantes évoquées N100, P200 et P600 seront sensibles aux variations acoustiques liées à l'accent régional et à leur interprétation (effet précoce et tardif). Plus exactement, nous émettons l'hypothèse que les participants écoutant les phrases émotionnelles du groupe culturel régional étranger (out-group) auront un signal précoce N100 et P200 inexistant et une augmentation à P600 comparé au groupe contrôle (*in-group*).

Nous avons créé avec E-Prime une étude expérimentale qui consiste à montrer 15 phrases X 4 émotions (Fierté, Colère, Joie, Honte) X 2 genre X 2 Origine (Qc et Fr) de nos phrases validées. Ainsi nous avons deux émotions à valence positive et deux émotions à valence négative. Il est prévu de recruter 40 Québécois et 40 Français, participantes femmes (car notre banque stimuli a été validée que par des femmes) ont pour tâche d'écouter le stimuli en regardant un écran noir, puis une croix de fixation d'une seconde apparaît, ensuit ils doivent

répondre à la question : ‘A quelle intensité l’émotion écoutée est-elle perçue ? et doit répondre à l’aide d’une échelle analogue visuelle entre « Minimum » et « Maximum ».

L’acquisition des données se fera grâce à un casque EEG à 64 canaux Neuroscan (basé sur le système international 10-20) avec une référence bi-mastoïdienne. Également afin d’enregistrer et d’enlever les artefacts liés aux mouvements oculaires, une électrode sera placée sous l’œil gauche et deux autres aux coins externes des yeux seront utilisées. Un filtre passe-bande de 0,1 Hz à 40 Hz sera appliqué et la fréquence d’échantillonnage sera de 1024 Hz. Après avoir référencé les données des électrodes d’enregistrement à la moyenne des électrodes, des fenêtres de données seront extraites pour chaque bonne réponse, allant de 200 ms avant à 800 ms après la présentation du stimulus, avec une correction pré-stimulus de 200 ms. Les moyennes de latence (ms) et d’amplitude (μV) seront calculées à partir des composantes N100, P200 et P600, et utilisées comme variables dépendantes dans l’analyse.

Les données EEG seront prétraitées selon les procédures standard (Lopez-Calderon et Luck, 2014) avec EEGLab (Lopez-Calderon et Luck, 2014) et ERPLab (Event-Related Potential) (Delorme et Makeig, 2004). Avec les mesures EEG d’amplitude et de la latence des ERPs, en variable dépendantes et les accents et les émotions en variables indépendantes nous réaliserons un les mêmes statistiques que notre étude sur les paramètres vocaux. Nous réaliserons donc, avec R-studio, un Modèle Général Mixte (Gelman, et al., 2020) avec les packages lme4 (Bates et al., 2015) et nlme (Kuznetsova et al., 2017) suivi d’une ANOVA (Type II Wald chisquare tests) pour explorer les effets inter-groupes avec les packages car et emmeans.

Conclusion

La présente étude visait à créer une banque de stimuli afin de comprendre la production et la perception de la prosodie émotionnelle selon les accents français et québécois. Nous sommes les premiers à montrer des résultats de profils acoustiques spécifiques aux Français et

aux Québécois selon plusieurs émotions basiques et culturelles. Nous montrons aussi des différences marquées entre les genres, surtout entre les hommes français et québécois. Nous discutons de la façon dont la culture façonne les émotions et surtout comment les émotions sont exprimées entre différentes cultures qui partagent pourtant la même langue.

Références

- Anxo, D., Mencarini, L., Pailhé, A., Solaz, A., Tanturri, M. L., & Flood, L. (2011). Gender Differences in Time Use over the Life Course in France, Italy, Sweden, and the US. *Feminist Economics*, 17(3), 159-195. <https://doi.org/10.1080/13545701.2011.582822>
- Audrin, C. (2020). Les émotions dans la formation enseignante : Une perspective historique. *Recherches en éducation*, 41, Article 41. <https://doi.org/10.4000/ree.541>
- Ayako, I., & Hansen, J. (2007). The Effect of Listener Accent Background on Accent Perception and Comprehension. *EURASIP Journal on Audio, Speech, and Music Processing*, 2007. <https://doi.org/10.1155/2007/76030>

- Banse, R., & Scherer, K. R. (1996a). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614-636. <https://doi.org/10.1037/0022-3514.70.3.614>
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv. arXiv preprint arXiv:1506.04967, 10.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). *Fitting Linear Mixed-Effects Models using lme4* (arXiv:1406.5823). arXiv. <https://doi.org/10.48550/arXiv.1406.5823>
- Bhatara, A., Laukka, P., Boll-Avetisyan, N., Granjon, L., Elfenbein, H. A., & Bänziger, T. (2016). Second Language Ability and Emotional Prosody Perception. *PLOS ONE*, 11(6), e0156855. <https://doi.org/10.1371/journal.pone.0156855>
- Boersma, P., & Weenink, D. (2001). PRAAT, a system for doing phonetics by computer. *Glott international*, 5, 341-345.
- Boula de Mareüil, P., Vieru-Dimulescu, B., Woehrling, C., & Adda-Decker, M. (2008). Accents étrangers et régionaux en français. Caractérisation et identification [Foreign and regional accents in French. Characterisation and identification]. *Traitement Automatique des Langues, Volume 49, Numéro 3 : Recherches actuelles en phonologie et en phonétique : interfaces avec le traitement automatique des langues [Current Research in Phonology and Phonetics: Interfaces with Natural-Language Processing]*, 135-163. <https://aclanthology.org/2008.tal-3.6>
- Brougère, G. (1997). Jeu et objectifs pédagogiques : Une approche comparative de l'éducation préscolaire. *Revue française de pédagogie*, 119, 47-56.
- Brunswik, E. 1956. Perception and the Representative Design of Psychological Experiments. 2nd ed. Berkeley, CA: University of California Press.

- Bryant, G. A., & Fox Tree, J. E. (2005). Is there an ironic tone of voice? *Language and Speech*, 48(3), 257-277. <https://doi.org/10.1177/00238309050480030101>
- Buchanan, T. W., Lutz, K., Mirzazade, S., Specht, K., Shah, N. J., Zilles, K., & Jäncke, L. (2000). Recognition of emotional prosody and verbal components of spoken language : An fMRI study. *Cognitive Brain Research*, 9(3), 227-238. [https://doi.org/10.1016/S0926-6410\(99\)00060-9](https://doi.org/10.1016/S0926-6410(99)00060-9)
- Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., & Weiss, B. (2005). *A database of German emotional speech*. 5, 1517-1520. <https://doi.org/10.21437/Interspeech.2005-446>
- Busso, C., Parthasarathy, S., Burmania, A., AbdelWahab, M., Sadoughi, N., & Provost, E. M. (2017). MSP-IMPROV : An Acted Corpus of Dyadic Interactions to Study Emotion Perception. *IEEE Transactions on Affective Computing*, 8(1), 67-80. <https://doi.org/10.1109/TAFFC.2016.2515617>
- Canpolat, S. F., Ormanoğlu, Z., & Zeyrek, D. (2020, May). Turkish Emotion Voice Database (TurEV-DB). In Proceedings of the 1st Joint Workshop on Spoken Language Technologies for Under-Resourced Languages (SLTU) and Collaboration and Computing for Under-Resourced Languages (CCURL) (pp. 368-375).
- Castro, S. L., & Lima, C. F. (2010a). Recognizing emotions in spoken language : A validated set of Portuguese sentences and pseudosentences for research on emotional prosody. *Behavior Research Methods*, 42(1), 74-81. <https://doi.org/10.3758/BRM.42.1.74>
- Ceravolo, L., Frühholz, S., Pierce, J., Grandjean, D., & Péron, J. (2021). Basal ganglia and cerebellum contributions to vocal emotion processing as revealed by high-resolution fMRI. *Scientific Reports*, 11(1), Article 1. <https://doi.org/10.1038/s41598-021-90222-6>

- Chu, J., Yang, C., & Liu, G. (2019). Analysis of Second Language Acquisition (SLA) Speech Perception Model & the Perception of Second Language Prosody. *Revista de Cercetare Si Interventie Sociala*, 64, 334-351. <https://doi.org/10.33788/reis.64.25>
- Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse processing. *Brain and cognition*, 46(1-2), 73-82.
- Cohen, P. R., Morgan, J., & Pollack, M. E. (Éds.). (2003). The Meaning of Intonational Contours in the Interpretation of Discourse. In *Intentions in Communication*. The MIT Press. <https://doi.org/10.7551/mitpress/3839.003.0016>
- Cowen, A. S., Laukka, P., Elfenbein, H. A., Liu, R., & Keltner, D. (2019). The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nature Human Behaviour*, 3(4), Article 4. <https://doi.org/10.1038/s41562-019-0533-6>
- Craig, D., A. (2009). How do you feel — now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, 10(1), Article 1. <https://doi.org/10.1038/nrn2555>
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language : A literature review. *Language and Speech*, 40 (Pt 2), 141-201. <https://doi.org/10.1177/002383099704000203>
- Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Van De Heyning, P., Remacle, M., Woisard, V., & Committee on Phoniatics of the European Laryngological Society (ELS). (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatics of the European Laryngological Society (ELS). *European Archives of Oto-Rhino-Laryngology: Official Journal of the European Federation of Oto-Rhino-Laryngological*

Societies (EUFOS): Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery, 258(2), 77-82. <https://doi.org/10.1007/s004050000299>

Delorme A & Makeig S (2004) EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics, *Journal of Neuroscience Methods* 134:9-21.

Diconne, K., Kountouriotis, G. K., Paltoglou, A. E., Parker, A., & Hostler, T. J. (2022). Presenting KAPODI – The Searchable Database of Emotional Stimuli Sets. *Emotion Review*, 14(1), 84-95. <https://doi.org/10.1177/17540739211072803>

Ekberg, M., Stavrinou, G., Andin, J., Stenfelt, S., & Dahlström, Ö. (2023). Acoustic Features Distinguishing Emotions in Swedish Speech. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2023.03.010>

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4), 169-200. <https://doi.org/10.1080/02699939208411068>

Ekman, P. (1999). Basic emotions. In *Handbook of cognition and emotion* (p. 45-60). John Wiley & Sons Ltd. <https://doi.org/10.1002/0470013494.ch3>

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 124-129. <https://doi.org/10.1037/h0030377>

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the human face : Guidelines for research and an integration of findings* (p. xii, 191). Pergamon Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science (New York, N.Y.)*, 164(3875), 86-88. <https://doi.org/10.1126/science.164.3875.86>

Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological Bulletin*, 128(2), 243-249. <https://doi.org/10.1037/0033-2909.128.2.243>

- Floccia, C., Butler, J., Goslin, J., & Ellis, L. (2009). Regional and Foreign Accent Processing in English : Can Listeners Adapt? *Journal of Psycholinguistic Research*, 38(4), 379-412.
<https://doi.org/10.1007/s10936-008-9097-8>
- Floccia, C., Goslin, J., Girard, F., & Konopczynski, G. (2006). Does a regional accent perturb speech processing? *Journal of Experimental Psychology: Human Perception and Performance*, 32(5), 1276-1293. <https://doi.org/10.1037/0096-1523.32.5.1276>
- Fredouille, C. (2016). Traitement Automatique et Troubles de la Voix et de la Parole: champs d'application, contraintes et limites (Doctoral dissertation, Université d'Avignon et du Pays de Vaucluse).
- Frühholz, S., & Grandjean, D. (2012). Towards a fronto-temporal neural network for the decoding of angry vocal expressions. *NeuroImage*, 62(3), 1658-1666.
<https://doi.org/10.1016/j.neuroimage.2012.06.015>
- Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories* (1^{re} éd.). Cambridge University Press. <https://doi.org/10.1017/9781139161879>
- Gholipour, A., Sedaaghi, M. H., & Shamsi, M. (2012). The contribution of prosody to the identification of Persian regional accents. *2012 IEEE Symposium on Industrial Electronics and Applications*.
https://www.academia.edu/27611610/The_contribution_of_prosody_to_the_identification_of_Persian_regional_accents
- Giles, H. (1970). Accent Mobility : A Model and Some Data. *Anthropological Linguistics*, 15.
- Golestani, N., Price, C. J., & Scott, S. K. (2011). Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *The Journal of Neuroscience: The Official Journal of the*

Society for Neuroscience, 31(11), 4213-4220. <https://doi.org/10.1523/JNEUROSCI.3891-10.2011>

- Gong, B., Li, N., Li, Q., Yan, X., Chen, J., Li, L., Wu, X., & Wu, C. (2023). The Mandarin Chinese auditory emotions stimulus database : A validated set of Chinese pseudo-sentences. *Behavior Research Methods*, 55(3), 1441-1459. <https://doi.org/10.3758/s13428-022-01868-7>
- Goslin, J., Duffy, H., & Floccia, C. (2012). An ERP investigation of regional and foreign accent processing. *Brain and Language*, 122(2), 92-102. <https://doi.org/10.1016/j.bandl.2012.04.017>
- Grandjean, D. (2021). Brain Networks of Emotional Prosody Processing. *Emotion Review*, 13(1), 34-43. <https://doi.org/10.1177/1754073919898522>
- Grimm, M., Kroschel, K., & Narayanan, S. (2008). The Vera am Mittag German audio-visual emotional speech database. *2008 IEEE International Conference on Multimedia and Expo*, 865-868. <https://doi.org/10.1109/ICME.2008.4607572>
- Gross, J. J., & Barrett, L. F. (2011). Emotion Generation and Emotion Regulation : One or Two Depends on Your Point of View. *Emotion review*, 3(1), 8-16. <https://doi.org/10.1177/1754073910380974>
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes : Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, 85(2), 348-362. <https://doi.org/10.1037/0022-3514.85.2.348>
- Gudykunst, W. (1994). *Bridging Differences : Effective Intergroup Communication* (4^e éd.). <https://doi.org/10.4135/9781452229706>
- Guzman, M., Correa, S., Muñoz, D., & Mayerhoff, R. (2013). Influence on Spectral Energy Distribution of Emotional Expression. *Journal of Voice*, 27(1), 129.e1-129.e10. <https://doi.org/10.1016/j.jvoice.2012.08.008>

- Hammerschmidt, K., & Jürgens, U. (2007). Acoustical Correlates of Affective Prosody. *Journal of Voice*, 21(5), 531-540. <https://doi.org/10.1016/j.jvoice.2006.03.002>
- Hirano, M. (1981). Clinical examination of voice. *Disorders of human communication*, 5, 1-99.
- Hirschfeld, U., & Trouvain, J. (2008). Teaching prosody in German as foreign language. In *Teaching prosody in German as foreign* (p. 171-188). De Gruyter Mouton. <https://doi.org/10.1515/9783110198751.2.171>
- Hofstede, G. (1980). *Culture's consequences : International differences in work-related values* (Vol. 5). sage.
- Hofstede, G., & Minkov, M. (2010). Long-versus short-term orientation: new perspectives. *Asia Pacific business review*, 16(4), 493-504.
- Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7241-7244. <https://doi.org/10.1073/pnas.1200155109>
- Jiang, X. (2017). Experimental Approaches to Socio-Linguistics : Usage and Interpretation of Non-Verbal and Verbal Expressions in Cross- Cultural Communication. In *Sociolinguistics— Interdisciplinary Perspectives*, 125. IntechOpen. <https://doi.org/10.5772/intechopen.69879>
- Jiang, X., Gossack-Keenan, K., & Pell, M. D. (2020a). To believe or not to believe? How voice and accent information in speech alter listener impressions of trust. *Quarterly Journal of Experimental Psychology*, 73(1), 55-79. <https://doi.org/10.1177/1747021819865833>
- Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate : Counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 27(33), 8877-8884. <https://doi.org/10.1523/JNEUROSCI.2063-07.2007>

- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance : Different channels, same code? *Psychological Bulletin*, *129*(5), 770-814.
<https://doi.org/10.1037/0033-2909.129.5.770>
- Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music : The need to consider underlying mechanisms. *The Behavioral and Brain Sciences*, *31*(5), 559-575; discussion 575-621. <https://doi.org/10.1017/S0140525X08005293>
- Keltner, D., Ekman, P., Gonzaga, G. C., & Beer, J. (2003). Facial expression of emotion. In *Handbook of affective sciences* (p. 415-432). Oxford University Press.
- Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (2019). Emotional Expression : Advances in Basic Emotion Theory. *Journal of Nonverbal Behavior*, *43*(2), 133-160.
<https://doi.org/10.1007/s10919-019-00293-3>
- Keshtiari, N., Kuhlmann, M., Eslami, M., & Klann-Delius, G. (2015). Recognizing emotional speech in Persian : A validated database of Persian emotional speech (Persian ESD). *Behavior Research Methods*, *47*(1), 275-294. <https://doi.org/10.3758/s13428-014-0467-x>
- Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, emotion, and well-being : Good feelings in Japan and the United States. *Cognition and Emotion*, *14*(1), 93-124.
<https://doi.org/10.1080/026999300379003>
- Kitayama, S., Mesquita, B., & Karasawa, M. (2006). Cultural affordances and emotional experience : Socially engaging and disengaging emotions in Japan and the United States. *Journal of personality and social psychology*, *91*, 890-903. <https://doi.org/10.1037/0022-3514.91.5.890>
- Kjellin, O. (1999). *Accent Addition : Prosody and Perception Facilitate Second Language Learning* (p. 373-398).

- Kotz, S. A., & Paulmann, S. (2011). Emotion, Language, and the Brain. *Language and Linguistics Compass*, 5(3), 108.
- Krüger, M. (2018). *Prosodic Decoding and Encoding of Referential Givenness in Adults with Autism Spectrum Disorders* [Text.thesis.doctoral, Universität zu Köln]. <http://www.uni-koeln.de/>
- Kuang, J., & Liberman, M. (2015). The effect of spectral slope on pitch perception. In Sixteenth Annual Conference of the International Speech Communication Association.
- Kumaran, D., & Maguire, E. A. (2005). The human hippocampus : Cognitive maps or relational memory? *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 25(31), 7254-7259. <https://doi.org/10.1523/JNEUROSCI.1103-05.2005>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package : Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1-26. <https://doi.org/10.18637/jss.v082.i13>
- Labov, W., Ash, S., & Boberg, C. (2008). The Atlas of North American English : Phonetics, Phonology and Sound Change. In *The Atlas of North American English*. De Gruyter Mouton. <https://doi.org/10.1515/9783110167467>
- Ladd, D. R. (1996). Intonational Phonology. (Cambridge Studies in Linguistics 79.) Cambridge : Cambridge University Press. pp. XV+334. *Phonology*, 15(1), 115-118. <https://doi.org/10.1017/s0952675798003546>
- Laukka, P., & Elfenbein, H. A. (2021). Cross-Cultural Emotion Recognition and In-Group Advantage in Vocal Expression : A Meta-Analysis. *Emotion Review*, 13(1), 3-11. <https://doi.org/10.1177/1754073919897295>

- Laukka, P., Elfenbein, H. A., Thingujam, N. S., Rockstuhl, T., Iraki, F. K., Chui, W., & Althoff, J. (2016). The expression and recognition of emotions in the voice across five nations : A lens model analysis based on acoustic features. *Journal of Personality and Social Psychology*, *111*(5), 686-705. <https://doi.org/10.1037/pspi0000066>
- Laver, J. (1980). The phonetic description of voice quality. (Pp. Ix + 186. Cambridge University Press, 1980.). *Journal of the International Phonetic Association*, *11*(2), 78-84. <https://doi.org/10.1017/S0025100300002292>
- Leboucher, M. (2023). *Les coûts de la protection contre les violences de genre : Autonomie et dépendance dans les groupes des tiers secteurs au Québec et en France*, 1-359. <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/28332>
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155-184. <https://doi.org/10.1146/annurev.neuro.23.1.155>
- Lengeris, A. (2012). *Prosody and Second Language Teaching : Lessons from L2 Speech Perception and Production Research* (Vol. 15, p. 25-40). https://doi.org/10.1007/978-94-007-3883-6_3
- Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition and Emotion*, *13*(5), 481-504. <https://doi.org/10.1080/026999399379159>
- Leventhal, H., & Scherer, K. (1987). The Relationship of Emotion to Cognition : A Functional Approach to a Semantic Controversy. *Cognition and Emotion*, *1*, 3-28. <https://doi.org/10.1080/02699938708408361>
- Lichev, V., Sacher, J., Ihme, K., Rosenberg, N., Quirin, M., Lepsien, J., Pampel, A., Rufer, M., Grabe, H.-J., Kugel, H., Kersting, A., Villringer, A., Lane, R. D., & Suslow, T. (2015). Automatic emotion processing as a function of trait emotional awareness : An fMRI study.

Social Cognitive and Affective Neuroscience, 10(5), 680-689.

<https://doi.org/10.1093/scan/nsu104>

Lippi-Green, R. (2011). *English with an Accent : Language, Ideology and Discrimination in the United States* (2^e éd.), 384. Routledge. <https://doi.org/10.4324/9780203348802>

Liu, P., & Pell, M. (2014). *Processing emotional prosody in Mandarin Chinese : A cross-language comparison* (p. 99). <https://doi.org/10.21437/SpeechProsody.2014-7>

Livingstone, S. R., & Russo, F. A. (2018a). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) : A dynamic, multimodal set of facial and vocal expressions in North American English. *PloS One*, 13(5), e0196391.

<https://doi.org/10.1371/journal.pone.0196391>

Mareüil, P. B. de, & Vieru-Dimulescu, B. (2006). The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica*, 63(4), 247-267. <https://doi.org/10.1159/000097308>

Markus, H. R., & Kitayama, S. (2010). Cultures and selves : A cycle of mutual constitution. *Perspectives on Psychological Science*, 5(4), 420-430.

<https://doi.org/10.1177/1745691610375557>

Matsumoto, D., Yoo, S. H., Hirayama, S., & Petrova, G. (2005). Development and validation of a measure of display rule knowledge : The display rule assessment inventory. *Emotion (Washington, D.C.)*, 5(1), 23-40. <https://doi.org/10.1037/1528-3542.5.1.23>

Mauchand, M., & Pell, M. D. (2020). Emotivity in the Voice : Prosodic, Lexical, and Cultural Appraisal of Complaining Speech. *Frontiers in Psychology*, 11, 619222.

<https://doi.org/10.3389/fpsyg.2020.619222>

Mauchand, M., Vergis, N., & Pell, M. (2019). Irony, Prosody, and Social Impressions of Affective Stance. *Discourse Processes*, 57, 1-17. <https://doi.org/10.1080/0163853X.2019.1581588>

- Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions : A review. *Psychological Bulletin*, 112(2), 179-204. <https://doi.org/10.1037/0033-2909.112.2.179>
- Mesquita, B. (2003). Emotions as dynamic cultural phenomena. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 871–890). Oxford University Press.
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in human neuroscience*, 8, 213.
- Morling, B., Kitayama, S., & Miyamoto, Y. (2002). Cultural practices emphasize influence in the United States and adjustment in Japan. *Personality and Social Psychology Bulletin*, 28(3), 311-323. <https://doi.org/10.1177/0146167202286003>
- Navas, E., Hernaez, I., & Luengo, I. (2006). An objective and subjective study of the role of semantics and prosodic features in building corpora for emotional TTS. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(4), 1117-1127. <https://doi.org/10.1109/TASL.2006.876121>
- Ohala, J. (1994). *Sound Symbolism*.
- Panksepp, J. (1998). *Affective neuroscience : The foundations of human and animal emotions* (p. xii, 466). Oxford University Press.
- Paulmann, S., Bleichner, M., & Kotz, S. A. (s. d.). *Frontiers | Valence, arousal, and task effects in emotional prosody processing*. Consulté 17 août 2023, à l'adresse <https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00345/full>
- Paulmann, S., & Kotz, S. A. (2008). An ERP investigation on the temporal dynamics of emotional prosody and emotional semantics in pseudo- and lexical-sentence context. *Brain and Language*, 105(1), 59-69. <https://doi.org/10.1016/j.bandl.2007.11.005>

- Paulmann, S., Ott, D. V. M., & Kotz, S. A. (2011). Emotional Speech Perception Unfolding in Time : The Role of the Basal Ganglia. *PLOS ONE*, 6(3), e17694.
<https://doi.org/10.1371/journal.pone.0017694>
- Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). Functional contributions of the basal ganglia to emotional prosody : Evidence from ERPs. *Brain Research*, 1217, 171-178.
<https://doi.org/10.1016/j.brainres.2008.04.032>
- Paulmann, S., & Uskul, A. K. (2014). Cross-cultural emotional prosody recognition : Evidence from Chinese and British listeners. *Cognition and Emotion*, 28(2), 230-244.
<https://doi.org/10.1080/02699931.2013.812033>
- Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing Emotions in a Foreign Language. *Journal of Nonverbal Behavior*, 33(2), 107-120. <https://doi.org/10.1007/s10919-008-0065-7>
- Pell, M. D., & Skorup, V. (2008). Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication*, 50(6), 519-530.
<https://doi.org/10.1016/j.specom.2008.03.006>
- Péron, J., Dondaine, T., Le Jeune, F., Grandjean, D., & V́erin, M. (2012). Emotional processing in Parkinson's disease : A systematic review. *Movement Disorders*, 27(2), 186-199.
<https://doi.org/10.1002/mds.24025>
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the Amygdala to Emotion Processing : From Animal Models to Human Behavior. *Neuron*, 48(2), 175-187.
<https://doi.org/10.1016/j.neuron.2005.09.025>
- Pierrehumbert, J., & Hirschberg, J. B. (1990). *The Meaning of Intonational Contours in the Interpretation of Discourse*. 271-311. <https://doi.org/10.7916/D8KD24FP>

- Pravena, D., & Govind, D. (2017). Development of simulated emotion speech database for excitation source analysis. *International Journal of Speech Technology*, 20(2), 327-338. <https://doi.org/10.1007/s10772-017-9407-3>
- Rao, K., & Koolagudi, S. (2011). Identification of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech. *Journal of Systemics, Cybernetics and Informatics*, 9.
- Rao, K. S., Koolagudi, S. G., & Vempada, R. R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, 16(2), 143-160. <https://doi.org/10.1007/s10772-012-9172-2>
- Regel, S., Gunter, T. C., & Friederici, A. D. (2011). *Isn't It Ironic ? An Electrophysiological Exploration of Figurative Language Processing*. 23(2).
- Revillard, A. (2008). Quelle politique pour les femmes ? Une comparaison France-Québec. *Revue internationale de politique comparée*, 15(4), 687-704. <https://doi.org/10.3917/ripc.154.0687>
- Riches, P., & Foddy, M. (1989). Ethnic Accent as a Status Cue. *Social Psychology Quarterly*, 52(3), 197-206. <https://doi.org/10.2307/2786714>
- Rigoulot, S., Wassiliwizky, E., & Pell, M. (2013). Feeling backwards? How temporal order in speech affects the time course of vocal emotion recognition. *Frontiers in Psychology*, 4. <https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00367>
- Roy, J., Tremblay, G., & Guilmette, D. (2014). *Un point de vue sociologique pour les pratiques*.
- Sander, D., Grafman, J., & Zalla, T. (2003). The Human Amygdala : An Evolved System for Relevance Detection. *Reviews in the neurosciences*, 14, 303-316. <https://doi.org/10.1515/REVNEURO.2003.14.4.303>

- Saratxaga, I., Navas, E., Hernáez, I., & Aholab, I. (2006). Designing and Recording an Emotional Speech Database for Corpus Based Synthesis in Basque. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*.
- Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences*, 107(6), 2408-2412. <https://doi.org/10.1073/pnas.0908239106>
- Scharenborg, O., Kakouros, S., & Koemans, J. (2018). *The Effect of Noise on Emotion Perception in an Unknown Language* (p. 368). <https://doi.org/10.21437/SpeechProsody.2018-74>
- Scherer, K. R. (1986). Vocal affect expression : A review and a model for future research. *Psychological Bulletin*, 99(2), 143-165.
- Scherer, K. R. (1996). *Adding the Affective Dimension : A New Look in Speech Analysis and Synthesis. ICSLP*.
- Scherer, K. R. (2003). Vocal communication of emotion : A review of research paradigms. *Speech Communication*, 40(1-2), 227-256. [https://doi.org/10.1016/S0167-6393\(02\)00084-5](https://doi.org/10.1016/S0167-6393(02)00084-5)
- Scherer, K. R., & Bänziger, T. (2004). Emotional expression in prosody: a review and an agenda for future research. In *Speech prosody 2004, international conference*.
- Scherer, K. R., Sundberg, J., Tamarit, L., & Salomão, G. L. (2015). Comparing the acoustic expression of emotion in the speaking and the singing voice. *Computer Speech & Language*, 29(1), 218-235. <https://doi.org/10.1016/j.csl.2013.10.002>
- Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere : Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, 10(1), 24-30. <https://doi.org/10.1016/j.tics.2005.11.009>

- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Research. Cognitive Brain Research*, 14(2), 228-233. [https://doi.org/10.1016/s0926-6410\(02\)00108-8](https://doi.org/10.1016/s0926-6410(02)00108-8)
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech : Sex differences revisited. *Brain Research. Cognitive Brain Research*, 24(3), 442-452. <https://doi.org/10.1016/j.cogbrainres.2005.02.022>
- Schirmer, A., Striano, T., & Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *NeuroReport: For Rapid Communication of Neuroscience Research*, 16(6), 635-639. <https://doi.org/10.1097/00001756-200504250-00024>
- Schneider, K. P., & Barron, A. (Éds.). (2008). *Variational pragmatics : A focus on regional varieties in pluricentric languages*. John Benjamins.
- Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences*, 13(8), 334-340. <https://doi.org/10.1016/j.tics.2009.05.001>
- Sobin, C., & Alpert, M. (1999). Emotion in speech : The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research*, 28(4), 347-365. <https://doi.org/10.1023/a:1023237014909>
- Sullins, J., Camp, L. G., Little, A., Waldron, L., & Wilson, A. (2020a). *A Prosodic Exploration of Academic Shame*.
- Thompson, W. F., & Balkwill, L.-L. (2006). *Decoding speech prosody in five languages*. 2006(158), 407-424. <https://doi.org/10.1515/SEM.2006.017>
- Titze, I. R., & Martin, D. W. (1998). *Principles of voice production*. Acoustical Society of America.

- Van Rijn, P., & Larrouy-Maestri, P. (2023). Modelling individual and cross-cultural variation in the mapping of emotions to speech prosody. *Nature Human Behaviour*, 7(3), 386-396.
<https://doi.org/10.1038/s41562-022-01505-5>
- Veaux, C., Yamagishi, J., & King, S. (2013). *The voice bank corpus : Design, collection and data analysis of a large regional accent speech database* (p. 4).
<https://doi.org/10.1109/ICSDA.2013.6709856>
- Vicsi, K., & Szaszák, G. (2010). Using prosody to improve automatic speech recognition. *Speech Communication*, 52(5), 413-426. <https://doi.org/10.1016/j.specom.2010.01.003>
- Voeten, C. C., & Levelt, C. C. (2019). ERP Responses to Regional Accent Reflect Two Distinct Processes of Perceptual Compensation. *Frontiers in Neuroscience*, 13.
<https://www.frontiersin.org/articles/10.3389/fnins.2019.00546>
- Woehrling, C., & de Mareüil, P. B. (2006). *Identification d'accents régionaux en français : Perception et catégorisation*.
- Wu, C., Li, H., Tian, Q., Wu, X., Wang, C.-Y., & Li, L. (2013). Disappearance of the unmasking effect of temporally pre-presented lipreading cues on speech recognition in people with chronic schizophrenia. *Schizophrenia research*, 150.
<https://doi.org/10.1016/j.schres.2013.08.017>
- Yuan, J., Shen, L., & Chen, F. (2002). The acoustic realization of anger, fear, joy and sadness in Chinese. *7th International Conference on Spoken Language Processing (ICSLP 2002)*, 2025-2028. <https://doi.org/10.21437/ICSLP.2002-556>
- Zhang, S. (2008). Emotion Recognition in Chinese Natural Speech by Combining Prosody and Voice Quality Features. In F. Sun, J. Zhang, Y. Tan, J. Cao, & W. Yu (Éds.), *Advances in*

Neural Networks—ISNN 2008 (p. 457-464). Springer. https://doi.org/10.1007/978-3-540-87734-9_52

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica*, 67(6), 361-370.

ANNEXE A

Tableau des phrases émotionnelles (banque de stimuli)

Joie	Surprise
1 Quelle merveilleuse journée	1 Tu m'offres un cadeau
2 Je suis si content(e) de te rencontrer	2 Mon père arrive demain
3 On va avoir un enfant	3 Je ne pensais pas avoir maigri autant cet hiver
4 Je vais voir mes amis ce soir	4 Elle est partie sans me dire au revoir
5 Je vais finalement avoir un chat	5 Pourquoi ma soeur appelle ce soir
6 J'ai réussi le concours d'entrée de l'école que je voulais	6 Elle n'en ai jamais rien su de cette histoire
7 Je pars en voyage avec mon amoureux	7 Tu arrives déjà de ton week-end
8 Demain c'est enfin le week-end	8 Tu as un frère jumeau
9 Mon frère m'a fait mon gâteau préféré	9 Je ne savais pas qu'il était encore vivant
10 J'ai gagné la course contre mon ami	10 Tu m'offres un bouquet de fleurs
11 C'était génial on n'a eu que du beau temps	11 Tu pars demain pour six mois
12 On m'a offert un nouveau vélo tout neuf	12 Je ne peux pas le croire
13 Je suis tellement contente de te revoir	13 Tu m'as cuisiné mon plat préféré
14 Je retrouve enfin ma famille	14 Il danse vraiment bien
15 Je me suis achetée une nouvelle guitare	15 Il a déménagé loin de son travail
16 On va à la plage cet après-midi	16 Ils n'habitent plus ensemble
17 On va passer un bon moment	17 Ils se sont séparés depuis longtemps
18 C'est mon anniversaire aujourd'hui	18 Que fais-tu ici
19 Ce soir on fait la fête	19 C'était toi à la télévision hier soir
20 Je suis de bonne humeur aujourd'hui	20 Tu t'es teint les cheveux en rouge
21 Je me suis tellement bien reposée ce week-end	21 Tu plaisantes elle n'a pas fait ça
22 Félicitation vous avez gagné le match de foot	22 Je ne m'attendais pas à ça
23 C'était une belle soirée tous enfin réunis	23 Tu as mis une surprise dans mon sac
24 On va pouvoir finalement aller au concert ce soir	24 Ton cousin est a joué dans ce film
25 J'aime cette ambiance chaleureuse	25 Tu es vraiment championne régional de gymnastique
26 On s'est vraiment bien amusé aujourd'hui	26 Tu vas partir vivre à l'étranger
27 Toute la famille se réunit pour passer Noël	27 Tu as une impressionnante collection d'oeuvres d'art
28 J'ai gagné un voyage en Thaïlande	28 Je n'aurais vraiment pas cru ça d'elle
29 Mon vœu s'est enfin réalisé	29 J'ai une grande soeur
30 Ça me rappelle des bons souvenirs	30 Tu n'aimes pas le chocolat
31 Nous avons mangé à notre restaurant préféré	31 Tu es devenu si grand
32 Je suis enfin en vacances	32 Tu as écrit cette pièce de théâtre toute seule
33 J'ai rencontré une personne formidable	33 Ton père est une personnalité connue
34 Nous avons goûté de délicieux cocktails	34 Tu as dix ans de plus que moi
35 J'ai pu récupérer tout mon argent	35 Je ne savais pas que c'était aussi grave
36 J'ai réalisé mon rêve cet été	36 Tu as voté pour ce candidat lors des élections
37 Cette visite avec toi était mémorable	37 Tu as repeint ta chambre toute seule
38 On a préparé une grande fête pour notre fille	38 J'ai bu toute la bouteille de vin
39 C'est le plus beau jour de ma vie	39 Tu n'as pas vu ton père depuis dix ans
40 Merci pour ce magnifique cadeau	40 Tu as une bonne note sans réviser

Tristesse	Honte
1 Le livre était à mon grand-père	1 Elle est tombée par terre devant tout le monde
2 Paul est parti de la maison	2 J'ai oublié l'anniversaire de mon frère
3 J'ai une mauvaise nouvelle à t'annoncer	3 J'ai brûlé mon plat je suis désolée
4 J'ai perdu mon téléphone dans le parc	4 Je lui ai menti
5 Mon ami ne veut plus me voir	5 Mon chien a reveillé tous les voisins cette nuit
6 Il n'y a plus de places pour le concert	6 Je me suis endormie pendant qu'elle parlait
7 Je peux t'appeler ça ne va pas très bien	7 J'ai joué faux durant tout le concert
8 Ma famille me manque tellement	8 J'ai rougi devant tout le monde
9 J'ai déchiré ma nouvelle robe	9 ils se sont moqués de moi
10 Mon chat est mort hier	10 Je me suis fait licencier de mon travail
11 J'ai envie de pleurer	11 C'est sale chez moi
12 Je ne peux pas avoir d'enfants	12 Je n'ai pas de cadeau pour toi
13 Mes amis sont partis sans moi	13 Je ne sais pas nagé
14 Je n'ai pas pu assister à son spectacle	14 Je l'ai appelé par un autre prénom
15 Le restaurant que nous adorons à fermé	15 J'ai eu un fou rire pendant l'enterrement
16 Mon vêtement préféré n'est plus récupérable	16 J'ai mis de la nourriture partout sur moi
17 Ils ont annulé le mariage	17 Je suis entrée chez la mauvaise personne
18 Plus rien ne va en ce moment	18 J'ai ronflé toute la nuit
19 J'ai perdu mon travail	19 Je l'ai confondu avec une autre personne
20 Les vacances sont déjà finies	20 Tu as de la salade entre les dents
21 Mon frère est parti loin de moi	21 Il l'a pris pour quelqu'un d'autre
22 Il m'a quitté hier soir	22 Personne ne m'a entendu parler
23 Ma grand-mère est malade	23 Je suis plus petite que tout le monde
24 J'ai perdu toute la motivation que j'avais	24 Mon mari ne sait pas faire la cuisine
25 Je me sens si seul en ce moment	25 Je me suis perdue plusieurs fois au même endroit
26 J'ai raté mon examen très important	26 J'arrive toujours en retard
27 Il n'a pas pu venir au rendez-vous	27 Ma maison n'est pas entretenue
28 J'ai raté mon train pour venir te voir	28 Je ne me rappelle plus d'elle
29 J'ai été disqualifié de la course	29 J'ai échoué devant tout le monde
30 On ne va plus jamais se revoir	30 Je n'ai pas beaucoup d'argent
31 Je ne peux pas venir aujourd'hui	31 Je n'ai pas les moyens de me payer mon loyer
32 Vous m'avez tous oublié	32 J'ai arrêté mes études
33 Je me sens si seul(e)	33 Il a tâché la veste du professeur
34 Cette musique me fait penser à mon ami décédé	34 J'ai renversé toute la soupe sur mon voisin
35 Le médecin lui a appris la mauvaise nouvelle	35 Je ne sais pas parler anglais
36 Je n'ai plus envie de sortir de chez moi	36 J'ai eu la plus mauvaise note de la classe
37 Je n'ai plus le goût à rien	37 Je me suis sentie humilié en classe
38 Nous n'avons plus aucun ami ici	38 Je me suis fait arrêter par la police
39 J'ai perdu tout ce que j'avais	39 Je me suis énervé contre la mauvaise personne
40 Je n'ai pas atteint mes objectifs cette année	40 J'ai encore oublié toutes mes affaires

Fierté	Colère	Jalousie
1 J'ai eu une bonne note à mon examen	1 Tu es encore en retard	1 Il a eu l'université que je voulais
2 J'ai enfin réussi à résoudre ce problème	2 Tu as perdu mon courrier	2 Ces chaussures lui vont mieux qu'à moi
3 J'ai gagné contre toi	3 Mes voisins font trop de bruit	3 C'est à elle que tu as envoyé un message
4 Mon fils a remporté la coupe	4 Toutes ces heures de travail pour rien	4 Elle est en train de lui parler depuis longtemps
5 J'ai acheté mon premier appartement	5 J'ai perdu toutes mes données	5 Leur groupe est meilleur que nous
6 Regardez ma nouvelle voiture	6 Tu as tout sali alors que j'avais tout nettoyé	6 Il réussit toujours mieux que moi
7 Mon fils est le premier de sa classe	7 Tu ne réponds jamais à mes questions	7 Tu es encore sortie avec cette fille
8 J'ai une pièce rare dans ma collection d'œuvres d'art	8 Il ne m'a jamais envoyé de message	8 Il a de la chance de manger sans grossir
9 Je porte la bague de ma grand-mère	9 Je n'arrive pas à faire cet exercice	9 Elle a reçu le prix d'excellence et pas moi
10 J'ai fait les travaux de ma maison toute seule	10 Tu as cassé ma nouvelle voiture	10 Elle a pu atteindre ses objectifs avant moi
11 Mon fils a dit son premier mot	11 Ça fait des heures que j'attends votre appel	11 Il me plaît mais c'est toi qu'il regarde
12 C'est moi qui ai écrit cette chanson	12 Le coiffeur a raté ma coupe de cheveux	12 De quoi vous discutiez depuis tout ce temps
13 Ma fille a vraiment fait preuve de courage	13 Je ne te supporte plus	13 Tu as triché pour gagner
14 Pour une fois je suis arrivée en avance	14 Tu m'empêches de dormir	14 Il est meilleur que moi dans tous les domaines
15 Mon fils est très mature pour son âge	15 Tu fais trop de bruits	15 Ils sont partis tout le week-end ensemble
16 J'ai eu une meilleure note que toi	16 Arrête de me crier dans les oreilles	16 Sa voiture est deux fois plus rapide
17 Pour une fois je me suis levée tôt ce matin	17 On m'a volé mon portefeuille	17 Tu t'es acheté une maison au bord de la mer
18 Je gagne très bien ma vie	18 L'assurance ne veut pas me rembourser	18 Elle a plus de temps libre que moi
19 J'ai récité ce poème sans faute	19 Taisez-vous tout le monde	19 Comment a-t-il fait pour gagner tout cet argent
20 J'ai appris à parler quatre langues	20 J'en ai marre de toi	20 Vous l'écoutez plus que moi
21 Je me suis trouvé un nouveau travail	21 Cet appareil ne marche jamais	21 Il gagne plus d'argent que moi
22 J'ai atteint mes objectifs cette année	22 Tu n'as encore pas fait tes devoirs	22 Il dessine mieux que moi
23 Mon patron m'a donné une promotion	23 Elle est tellement égoïste	23 Tu es plus proche de tes nouveaux amis
24 J'ai eu mon premier emploi	24 Tu n'as toujours pas fait ce que je t'ai demandé	24 Qui est cet homme qui te parle tout le temps
25 Je suis devenue quelqu'un de connu	25 Arrête de me couper la parole	25 Je ne suis pas invitée à cette soirée
26 J'ai acheté un tableau d'une grande valeur	26 Il m'a menti et m'a tout pris	26 Pourquoi tu la regardes encore
27 Elle a vraiment écrit un bon bouquin	27 Vous êtes en train de me ridiculiser	27 Tu passes plus de temps avec elle
28 J'ai été élu pour représenter mon école	28 Tu ne m'écoutes pas quand je te parle	28 Elle a de beaux cheveux longs
29 Je n'ai pas eu besoin d'aide pour réussir ce travail	29 Arrête d'avoir peur pour rien	29 J'aimerais tellement avoir ses yeux
30 J'ai fait une bonne affaire en achetant ce meuble	30 Tu ne fais jamais sa vaisselle	30 Il gagne tout le temps chaque compétition
31 Regardez ce que j'ai réussi à faire	31 Qui a encore laissé la porte ouverte	31 Ils ont une plus belle complicité
32 On a gagné le tournoi de football	32 Tu as perdu la bague de ma grand-mère	32 Je n'ai pas son talent
33 Nous avons fait le tour d'Europe en vélo	33 Je perds tout le temps contre toi	33 Elle a eu plus de chocolat que moi
34 Regarde mon nouvel exploit	34 Il n'aurait jamais dû faire ça	34 Elle s'exprime vraiment mieux que moi
35 Je maîtrise parfaitement cette danse maintenant	35 Je me suis pris une amende	35 Elle joue mieux que moi du piano
36 On m'a offert le poste de directeur	36 On a encore amené ma voiture à la fourrière	36 Tu pars avec tes amis ce week-end sans moi
37 Je suis la meilleure dans ma discipline	37 Je suis de mauvaise humeur j'ai mal dormi	37 J'aimerais nager aussi bien que ma sœur
38 Je connais tout mon texte par cœur	38 Tu ne comprends rien à ce que je te dis	38 Il a eu la promotion que je voulais avoir
39 On m'a longuement félicité pour mon prix	39 Il y a encore des bouchons	39 Il plaît à toutes les filles
40 J'ai une excellente mémoire	40 Je suis bloquée depuis deux jours sur ce dossier	40 Elle le dévore des yeux

ANNEXE B

Questionnaire HADS (Zigmond et Snaith, 1983)

Le questionnaire HADS (de l'anglais *Hospital Anxiety and Depression Scale*)

Dans la série de questions ci-dessous, cochez la réponse qui exprime le mieux ce que vous avez éprouvé au cours de la semaine qui vient de s'écouler. Ne vous attardez pas sur la réponse à faire : votre réaction immédiate à chaque question fournira probablement une meilleure indication de ce que vous éprouvez, qu'une réponse longuement méditée.

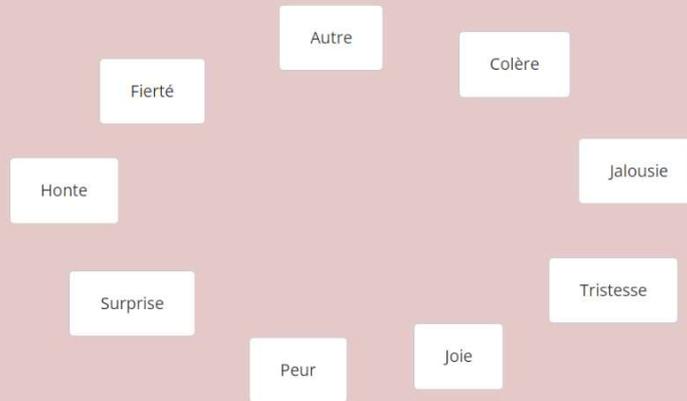
Score	Anxiété	Score	Dépression
3 2 1 0	Je me sens tendu ou énervé : <input type="checkbox"/> la plupart du temps <input type="checkbox"/> souvent <input type="checkbox"/> de temps en temps <input type="checkbox"/> jamais	0 1 2 3	Je prends plaisir aux mêmes choses qu'autrefois <input type="checkbox"/> oui, tout autant <input type="checkbox"/> pas autant <input type="checkbox"/> un peu seulement <input type="checkbox"/> presque plus
3 2 1 0	J'ai une sensation de peur comme si quelque chose d'horrible allait m'arriver <input type="checkbox"/> oui, très nettement <input type="checkbox"/> oui, mais ce n'est pas grave <input type="checkbox"/> un peu, mais cela ne m'inquiète pas <input type="checkbox"/> pas du tout	0 1 2 3	Je ris facilement et vois le bon côté des choses <input type="checkbox"/> autant que par le passé <input type="checkbox"/> plus autant qu'avant <input type="checkbox"/> vraiment moins qu'avant <input type="checkbox"/> plus du tout
3 2 1 0	Je me fais du souci : <input type="checkbox"/> très souvent <input type="checkbox"/> assez souvent <input type="checkbox"/> occasionnellement <input type="checkbox"/> très occasionnellement	3 2 1 0	Je suis de bonne humeur : <input type="checkbox"/> jamais <input type="checkbox"/> rarement <input type="checkbox"/> assez souvent <input type="checkbox"/> la plupart du temps
0 1 2 3	Je peux rester tranquillement assis à ne rien faire et me sentir décontracté : <input type="checkbox"/> oui, quoi qu'il arrive <input type="checkbox"/> oui, en général <input type="checkbox"/> rarement <input type="checkbox"/> jamais	3 2 1 0	J'ai l'impression de fonctionner au ralenti : <input type="checkbox"/> presque toujours <input type="checkbox"/> très souvent <input type="checkbox"/> parfois <input type="checkbox"/> jamais
0 1 2 3	J'éprouve des sensations de peur et j'ai l'estomac noué : <input type="checkbox"/> jamais <input type="checkbox"/> parfois <input type="checkbox"/> assez souvent <input type="checkbox"/> très souvent	3 2 1 0	Je ne m'intéresse plus à mon apparence : <input type="checkbox"/> plus du tout <input type="checkbox"/> je n'y accorde pas autant d'attention que je le devrais <input type="checkbox"/> il se peut que je n'y fasse plus autant attention <input type="checkbox"/> j'y prête autant d'attention que par le passé
3 2 1 0	J'ai la bougeotte et n'arrive pas à tenir en place : <input type="checkbox"/> oui, c'est tout à fait le cas <input type="checkbox"/> un peu <input type="checkbox"/> pas tellement <input type="checkbox"/> pas du tout	0 1 2 3	Je me réjouis d'avance à l'idée de faire certaines choses : <input type="checkbox"/> autant qu'auparavant <input type="checkbox"/> un peu moins qu'avant <input type="checkbox"/> bien moins qu'avant <input type="checkbox"/> presque jamais
3 2 1 0	J'éprouve des sensations soudaines de panique : <input type="checkbox"/> vraiment très souvent <input type="checkbox"/> assez souvent <input type="checkbox"/> pas très souvent <input type="checkbox"/> jamais	0 1 2 3	Je peux prendre plaisir à un bon livre ou à une bonne émission radio ou de télévision : <input type="checkbox"/> souvent <input type="checkbox"/> parfois <input type="checkbox"/> rarement <input type="checkbox"/> très rarement
	Total du score pour l'anxiété		Total du score pour la dépression

Chaque réponse correspond à un chiffre. En additionnant ces chiffres, on obtient un score total par colonne (anxiété et dépression). Si le score d'une colonne est supérieur ou égal à 11, cela signifie que vous souffrez d'anxiété ou de dépression (selon la colonne concernée).

ANNEXE C

Tâche de l'expérience de validation en ligne

Quelle émotion le locuteur exprime t-il?



A quelle intensité est exprimée l'émotion?



Continue