

Voix et émotions: impact de l'intensité émotionnelle sur la hauteur tonale en fonction de l'étendue fréquentielle des participants

Auteur : Sasserath, Juliette

Promoteur(s) : MORSOMME, Dominique; 4709

Faculté : Faculté de Psychologie, Logopédie et Sciences de l'Éducation

Diplôme : Master en logopédie, à finalité spécialisée en voix

Année académique : 2017-2018

URI/URL : <http://hdl.handle.net/2268.2/5888>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.



Université de Liège

Faculté des Sciences Psychologiques et de l'Éducation

Unité de logopédie de la voix

Voix et émotions :

*Impact de l'intensité émotionnelle sur la hauteur tonale
en fonction de l'étendue fréquentielle des participants*

Promotrice : Dominique MORSOMME

Co-Promotrice : Ingrid VERDUYCKT

Mémoire présenté en vue de l'obtention du grade de Master en logopédie à
finalité spécialisée en voix par Juliette SASSERATH

Année académique 2017 – 2018

Remerciements :

Je tenais à remercier les personnes qui m'ont soutenue tout au long de la réalisation de ce mémoire.

Tout d'abord, je remercie Dominique Morsomme, ma promotrice, pour son soutien, sa disponibilité et ses qualités humaines qui ont confirmé mon envie de travailler dans ce domaine.

Je remercie tout particulièrement Ingrid Verduyckt qui, malgré la distance, s'est montrée extrêmement disponible, m'a suivie et soutenue tout au long de ce travail et m'a donné de nombreux conseils pertinents.

Je remercie également Madame Comblain et le Docteur Lagier, mes lectrices qui prendront le temps de lire et de s'intéresser à mon travail.

Je remercie Ellen Blanckaert pour ses conseils avisés quant à l'utilisation du matériel et à l'emploi des logiciels d'analyse.

Ensuite, je remercie l'ensemble des participants, pour le temps qu'ils m'ont accordé, leur intérêt pour le projet et leur soutien.

Un merci particulier à ma famille, mon compagnon et mes amis pour leur aide et leurs encouragements perpétuels tout au long de la réalisation de ce travail.

Table des matières

I. Introduction.....	2
II. Les caractéristiques acoustiques de la voix.....	4
A. L'analyse spectrale.....	4
B. Le niveau de pression sonore.....	5
C. La hauteur tonale.....	5
1. La fréquence fondamentale.....	5
2. Les harmoniques.....	6
3. Les formants.....	6
4. Le triangle vocalique.....	7
D. Le timbre.....	8
E. L'instabilité en hauteur – le Jitter.....	8
F. L'instabilité en intensité – le Shimmer.....	8
G. L'étendue vocale.....	9
H. Les différents mécanismes.....	12
I. Le rapport harmonique bruit (HNR).....	12
III. Les émotions.....	13
A. Définition d'une émotion.....	13
B. L'émotion est-elle universelle ou propice aux variations culturelles ?.....	13
1. Ekman : conception universelle.....	14
2. Averill : variations culturelles.....	14
3. Scherer : conciliation de l'universel et du culturel.....	15
C. Le système limbique.....	16
1. L'hypothalamus.....	16
2. Le corps amygdaloïde.....	16
3. L'hippocampe.....	16
D. Perception et expression émotionnelles.....	17
IV. Voix et émotions.....	19
A. Liens entre les émotions et la voix.....	19
B. Paramètres vocaux influencés par l'état émotionnel.....	21
C. La perception émotionnelle.....	22
D. Perception émotionnelle interculturelle.....	23
E. Perception et production émotionnelle selon l'âge.....	23
F. Perception et production émotionnelle selon le genre.....	26
G. Perception émotionnelle et intelligence artificielle.....	27
H. Physiologie et émotions.....	28
V. Objectifs et hypothèses.....	32
A. Objectifs de ce mémoire.....	32
B. Hypothèses de ce mémoire.....	33

VI. Méthodologie.....	34
A. Population	34
1. Population de l'étude	34
2. Population de ce mémoire.....	34
B. Matériel	34
1. Questionnaires psychologiques	34
a) Questionnaire de Sensibilité à la Punition et de Sensibilité à la Récompense (QSPSR)	34
b) L'échelle de Régulation Émotionnelle (DERS-F).....	35
c) L'Échelle d'Anxiété Sociale de Liebowitz (ESCL)	35
d) L'Inventaire de Dépression de Beck Seconde Edition (IDB-II).....	35
e) La Positive and Negative Affect Scale (PANAS).....	35
2. Matériel pour l'enregistrement vocal.....	36
a) Le microphone et le casque	36
b) La carte son	36
c) Les logiciels d'enregistrement et d'analyse vocale	37
3. Matériel pour les tâches de productions vocales émotionnelles	37
C. Lieu des enregistrements	38
VII. Procédure générale.....	39
A. Évaluation par questionnaires à domicile	39
B. Administratif et questionnaires psychologiques.....	39
1. Accueil.....	39
2. Administratif	39
3. Inventaire de Dépression de Beck (BECK-II).....	40
4. La Positive and Negative Affect Scale (PANAS).....	40
C. Tâches vocales	40
1. Tâche d'étendue fréquentielle	40
2. Production des sons pour la tâche de justesse vocale.....	41
3. Préparation pour la suite de l'expérience	41
4. Tâche de justesse vocale	41
D. Expérience.....	42
1. Set 1 : spontané	43
2. Set 2 : imposé.....	43
3. Set 3 : simulé/neutre	43
4. Set 4 : masqué	43
E. Durée de l'expérience.....	44
F. Variables recueillies et analysées	44
VIII. Résultats	45
A. Description de la population	46
B. Étendues fréquentielles des participants en demi-tons	46
1. Étendues fréquentielles et physiologiques.....	46
C. Vocalisations émotionnelles sur les sets 1 et set 2.....	48
1. Nombre de vocalisations pour les sets 1 et 2.....	48
2. Comparaison des étendues fréquentielles émotionnelles entre les sets 1 et 2.....	48
3. Comparaison des étendues fréquentielles conversationnelles et émotionnelles.....	49
D. Vocalisations émotionnelles des sets 1 et 2 selon le niveau d'intensité émotionnelle estimé 49	
1. Valeurs d'intensité et de valence émotionnelles relevées dans le set 2.....	49
2. Étendues fréquentielles des vocalisations émotionnelles du set 2 selon l'intensité émotionnelle faible, moyenne et élevée	50
E. Résultats statistiques au regard de l'hypothèse 1 :	51

1.	Analyse de l'étendue fréquentielle des vocalisations émotionnelles selon les niveaux d'intensité cotée par les participants.....	52
a)	Analyse de l'étendue fréquentielle maximale en fonction du niveau d'intensité émotionnelle.....	52
b)	Analyse de l'étendue fréquentielle sectionnée en fonction du niveau d'intensité émotionnel.....	53
2.	Analyse de distance entre l'étendue fréquentielle des échantillons vocaux émotionnels et l'étendue fréquentielle conversationnelle des participants.....	55
a)	Analyse de la distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels des participants (DistMin).....	55
b)	Analyse de la distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels des participants (DistMax).....	57
c)	Analyse de distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels des participants (DistSect).....	58
3.	Réponse à l'hypothèse 1.....	59
F.	Résultats statistiques au regard de l'hypothèse 2 :	61
1.	Liens entre les intensités et valences estimées du participant.....	61
2.	Analyses des différences d'étendues fréquentielles émotionnelles selon le set.....	61
a)	Analyses des différences d'étendues fréquentielles émotionnelles maximales en fonction du niveau d'intensité émotionnelle estimé et du set.....	62
b)	Analyses des différences d'étendues fréquentielles émotionnelles sectionnées en fonction du niveau d'intensité émotionnelle estimé et du set.....	64
c)	Analyses de la distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé et du set (DistMin).....	66
d)	Analyses de la distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé et du set (DistSect).....	67
e)	Analyses de la distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé et du set (DistMax).....	69
3.	Réponse à l'hypothèse 2.....	71
G.	Résultats statistiques au regard de l'hypothèse 3 :	72
1.	Comparaison des étendues fréquentielles émotionnelles et des étendues fréquentielles conversationnelles.....	72
a)	Comparaison des étendues fréquentielles émotionnelles et des étendues fréquentielles conversationnelles maximales.....	72
b)	Comparaison des étendues fréquentielles émotionnelles et des étendues fréquentielles conversationnelles sectionnées.....	74
2.	Réponse à l'hypothèse 3.....	75
IX.	Discussion.....	76
A.	Analyses des statistiques descriptives et choix méthodologiques.....	76
1.	Description de la population.....	76
a)	Origine de la population.....	76
b)	Effectif de la population.....	76
2.	Étendues fréquentielles des participants.....	76
a)	Le quantile 25.....	76
b)	Le quantile 75.....	77
c)	Étendue fréquentielle conversationnelle.....	77
3.	Conversion des fréquences hertziennes en demi-tons.....	78
4.	Nombre de vocalisations émotionnelles.....	79
5.	Intensité émotionnelle cotée par le participant.....	79
B.	Discussion des résultats et des hypothèses.....	80
1.	Discussion au regard des résultats de l'hypothèse 1.....	80
a)	Analyses sur le set imposé.....	80
b)	Résultats : validation de l'hypothèse.....	80
c)	Limites.....	81
d)	Différences d'étendues fréquentielles entre les niveaux d'intensité cotée par le participant.....	81
e)	Différences inter-genres.....	82

2.	Discussion au regard des résultats de l'hypothèse 2	82
a)	Estimation des intensités émotionnelles à l'aide de la valence cotée par le participant.....	82
b)	Résultats : hypothèse non-vérifiée	83
c)	Distances inter-quantiles et étendues fréquentielles des vocalisations	84
3.	Discussion au regard des résultats de l'hypothèse 3	85
a)	Résultats : validation de l'hypothèse.....	85
b)	Différences inter-genres	86
c)	Différences interculturelles.....	87
C.	Forces et faiblesses de cette recherche.....	87
1.	Limites.....	87
a)	Population.....	87
b)	Méthodologie.....	87
c)	Mesures physiologiques	88
2.	Point positifs	88
a)	Population non professionnelle de la voix.....	88
b)	Étude internationale	89
X.	Conclusions et perspectives	90
XI.	Bibliographie.....	92
XII.	Annexes	100

Liste des figures

Figure 1.	Triangle vocalique des voyelles (The international phonetic alphabet, revised to 2015)	7
Figure 2.	Valeurs formantiques des voyelles du français (Méchain, C., 2014).....	7
Figure 3.	Étendues vocales des enseignants (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002)	10
Figure 4.	Étendues vocales pour les voyelles /a/ et /i/ (Titze, I, 1992, adapté de Gramming, 1988)	11
Figure 5.	Clavier de Pierce, J.	11
Figure 6.	Comparaison étendues conversationnelles en demi-tons entre le quantile 25 et la F0 maximale et entre les quantiles 25 et 75	47
Figure 7.	Comparaison des étendues fréquentielles conversationnelles et physiologiques.....	47
Figure 8.	Comparaison des étendues fréquentielles conversationnelles et émotionnelles en demi-tons : A= moyennes entre les fréquences des quantiles 25 et des fréquences maximales ; B= moyennes entre les fréquences des quantiles 25 et quantiles 75.....	49
Figure 9.	Catégories d'intensité émotionnelle basées sur les cotations individuelles d'intensité émotionnelle	50
Figure 10.	Histogrammes de distribution des données fréquentielles émotionnelles selon l'intensité émotionnelle, et courbe gaussienne : 1→3 : étendue entre la fréquence du quantile 25 et le quantile 75 ; 4→6 : étendue entre la fréquence du quantile 25 et la fréquence maximale.....	51
Figure 11.	Étendues fréquentielles maximales des vocalisations émotionnelles en fonction de l'intensité...	53
Figure 12.	Étendues fréquentielles sectionnées des vocalisations émotionnelles en fonction de l'intensité .	54
Figure 13.	Distances séparant les quantiles 25 des échantillons émotionnels et de l'étendue conversationnelle	56
Figure 14.	Distances séparant les fréquences maximales des échantillons émotionnels et de l'étendue conversationnelle.....	57
Figure 15.	Distances séparant les fréquences des quantiles 75 des échantillons émotionnels et de l'étendue conversationnelle.....	59
Figure 16.	Comparaison des excursions des étendues fréquentielles maximales et sectionnées en fonction de l'intensité	60
Figure 17.	Comparaison des excursions fréquentielles maximales et sectionnées en fonction de l'intensité et du set.	72
Figure 18.	Comparaison des étendues fréquentielles conversationnelles et émotionnelles sans distinction d'intensité.....	75
Figure 19.	Taux de réponses des vocalisations émotionnelles en fonction de l'intensité calculée sur base des scores de valence du set 1.....	121

Liste des tableaux

Tableau 1 : Description de l'effectif et de l'âge de la population de l'étude	46
Tableau 2 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale, et entre les moyennes des quantiles 25 et 75 pour le conversationnel et entre la fréquence minimale et maximale pour le physiologique.....	47
Tableau 3 : Nombre de vocalisations émotionnelles pour les sets 1 et 2 par pays et par genre	48
Tableau 4 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale, et entre les moyennes des quantiles 25 et 75 pour les vocalisations émotionnelles en demi-tons pour les sets 1 et 2.....	48
Tableau 5 : Étendues fréquentielles des vocalisations émotionnelles en demi-tons mesurées en fonction de l'intensité émotionnelle cotée par les participants.....	50
Tableau 6 : Étendues fréquentielles maximales des vocalisations émotionnelles en demi-tons mesurées en fonction de l'intensité émotionnelle cotée par le participant.....	52
Tableau 7 : Test des effets intra-participants pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en fonction de l'intensité émotionnelle.....	53
Tableau 8 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en fonction de l'intensité émotionnelle.....	53
Tableau 9 : Étendues fréquentielles sectionnées des vocalisations émotionnelles en demi-tons mesurées en fonction de l'intensité émotionnelle cotée par les participants.....	54
Tableau 10 : Test multivarié pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en fonction de l'intensité émotionnelle cotée par les participants.....	54
Tableau 11 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en fonction de l'intensité émotionnelle cotée par les participants.....	55
Tableau 12 : Distance entre les fréquences du quantile 25 des échantillons émotionnels et les fréquences du quantile 25 des étendues fréquentielles conversationnelles.....	56
Tableau 13 : Test multivarié pour la distance entre les fréquences du quantile 25 des échantillons émotionnels et les fréquences du quantile 25 des étendues fréquentielles conversationnelles.....	56
Tableau 14 : Test des effets inter-participants pour la distance entre les fréquences du quantile 25 des échantillons émotionnels et du quantile 25 des étendues fréquentielles conversationnelles.....	56
Tableau 15 : Distance entre les fréquences maximales des échantillons émotionnels et les fréquences maximales des étendues fréquentielles conversationnelles.....	57
Tableau 16 : Test multivarié pour la distance entre les fréquences maximales des échantillons émotionnels et les fréquences maximales des étendues fréquentielles conversationnelles.....	58
Tableau 17 : Test des effets inter-participants pour la distance entre les fréquences maximales des échantillons émotionnels et les fréquences maximales des étendues fréquentielles conversationnelles.....	58
Tableau 18 : Distance entre les fréquences des quantiles 75 des échantillons émotionnels et les fréquences des quantiles 75 des étendues fréquentielles conversationnelles.....	58
Tableau 19 : Test multivarié pour la distance entre les fréquences du quantile 75 des échantillons émotionnels et les fréquences du quantile 75 des étendues fréquentielles conversationnelles.....	59
Tableau 20 : Test des effets inter-participants pour la distance entre les fréquences du quantile 75 des échantillons émotionnels et du quantile 75 des étendues fréquentielles conversationnelles.....	59
Tableau 21 : Étendues fréquentielles maximales des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 1.....	63
Tableau 22 : Étendues fréquentielles maximales des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 2.....	63
Tableau 23 : Test multivarié pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimée.....	64
Tableau 24 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimée.....	64
Tableau 25 : Étendues fréquentielles sectionnées des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 1.....	64

Tableau 26 : Étendues fréquentielles sectionnées des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 2.....	65
Tableau 27 : Test multivarié pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimés.....	65
Tableau 28 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimés.....	66
Tableau 29 : Distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 1.....	66
Tableau 30 : Distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 2.....	66
Tableau 31 : Test multivarié pour l'analyse des distances entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé.....	67
Tableau 32 : Test des effets inter-participants pour l'analyse des distances entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé.....	67
Tableau 33 : Distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons en fonction du niveau d'intensité émotionnelle estimé pour le set 1.....	68
Tableau 34 : Distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 2.....	68
Tableau 35 : Test des effets intra-participants pour l'analyse des distances entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé.....	69
Tableau 36 : Test des effets inter-participants pour l'analyse des distances entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé.....	69
Tableau 37 : Distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 1.....	70
Tableau 38 : Distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 2.....	70
Tableau 39 : Test des effets intra-participants pour l'analyse des distances entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé.....	71
Tableau 40 : Test des effets inter-participants pour l'analyse des distances entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé.....	71
Tableau 41 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale pour les vocalisations émotionnelles et conversationnelles.....	73
Tableau 42 : Test des effets intra-participants pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale pour les vocalisations émotionnelles et conversationnelles.....	73
Tableau 43 : Test des effets inter-participants pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale pour les vocalisations émotionnelles et conversationnelles.....	73
Tableau 44 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de quantile 75 pour les vocalisations émotionnelles et conversationnelles.....	74
Tableau 45 : Test multivarié pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et quantile 75 pour les vocalisations émotionnelles et conversationnelles.....	74
Tableau 46 : Test des effets inter-participants pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et 75 pour les vocalisations émotionnelles et conversationnelles....	74

PARTIE
THEORIQUE

I. Introduction

Les émotions sont exprimées de multiples manières, y compris par la voix, mais lors de la présentation d'un stimulus visuel, quelles sont les différentes variables acoustiques impactées par la valence et l'intensité émotionnelles ? L'émotion peut être définie comme une expérience affective relativement intense, dont la cause est clairement identifiée, et qui ne dure pas très longtemps. On peut différencier les émotions utilitaires, menant à l'action (la colère, la joie, la peur ou le dégoût) des émotions esthétiques, ne menant pas à l'action (admiration, extase, harmonie, solennité) (Scherer, K.R., 2005). La présentation d'un stimulus émotionnel engendre des modifications du système limbique, mais également des modifications des expressions faciales et vocales. Ces modifications faciales sont souvent visibles, tandis que les modifications vocales sont perceptibles auditivement. Celles-ci permettent à la personne de transmettre un état émotionnel vécu, de manière instantanée aux personnes qui l'entourent. La voix est donc un bon indicateur de l'état émotionnel d'une personne. Aussi, cette question de recherche mérite attention car il est également connu qu'un impact émotionnel, à répétition sur l'appareil vocal, peut endommager celui-ci et entraîner des pathologies vocales fréquemment rencontrées en clinique (Dietrich, M. Verdolini Abbott, K. Gartner-Schmidt, J et Rosen, C.A., 2008 ; Van Mersbergen, M., 2011). Dans le milieu phoniatrique, nous distinguons principalement les pathologies vocales fonctionnelles et les pathologies vocales organiques. La répartition moyenne de ces pathologies dans une population adulte est de 30% de cas de dysphonies fonctionnelles contre 43 à 65 % de cas de dysphonies organiques d'origine phonotraumatique (Cohen, S.M., Kim, J., Roy, N., Asche, C. et Courey, M., 2012 ; Van Houtte, E., Van Lierde, K., D'Haeseleer, E. et Claeys, S., 2010). Par contre, les dysphonies organiques phonotraumatiques représentent 77,7% des cas d'une population infantile (Coyle, S.M., Weinrich, B.D. et Stemple J.C., 2001 ; Dobres, R., Lee, L., Stemple, J.C., Kummer, A.W. et Kretschmer, L.W., 1990 ; Van Houtte, E., Van Lierde, K., D'Haeseleer, E. et Claeys, S., 2010).

Si l'intérêt d'étudier l'impact émotionnel sur les différents paramètres vocaux dans le contexte des pathologies vocales est justifié, il apparaît que les études existantes à ce sujet ont ciblé majoritairement des acteurs et comédiens. Il existe donc peu d'études relatant les liens directs entre les émotions et les paramètres vocaux en situation naturelle, et cela dans une population diversifiée. De plus, il semble important de vérifier la stabilité de ces paramètres vocaux en fonction du contexte communicationnel et de l'influence environnementale sur le participant.

Lorsqu'une personne est confrontée à un stimulus émotionnel, on peut mesurer sa réaction par des changements physiologiques, comportementaux et cognitifs (Genter, A., 2014). Pour induire des émotions et ensuite en mesurer les conséquences, il est donc important de prendre en compte trois composantes. La première concerne la composante cognitive, qui correspond aux changements d'état mental. La seconde est la composante comportementale, qui correspond aux manifestations comportementales et expressives d'une émotion. La troisième est la composante physiologique, qui se réfère aux manifestations physiologiques concomitantes à un événement émotionnel (Humphreys, M. S., et Revelle, W., 1984 ; Kim, J., Bouchard, C., Ryu, H., Omhover, J.-F., et Aoussat, A., 2012). D'un point de vue statistique, les émotions peuvent être étudiées en tant que variables dépendantes afin de les étudier en tant que telles, mais elles peuvent également être étudiées comme des variables indépendantes afin d'en recueillir, quantifier et analyser les conséquences cognitives, comportementales et physiologiques. La variation des paramètres vocaux acoustiques tels que la hauteur tonale, est une des conséquences physiologiques des émotions. Il est donc important de prendre en compte le point de vue physiologique, mais aussi le point de vue psychologique et de faire le lien entre ces domaines.

Pour pallier ce manque de données, je vais tenter, à travers ce mémoire, d'investiguer les variations de fréquences fondamentales en fonction d'un stimulus émotionnel visuel présenté à des participants neurotypiques. Des stimuli visuels seront employés car cette méthodologie s'inscrit dans l'étude « Voix et émotions » réalisée à l'Université de Montréal et qui emploie les mêmes stimuli. Ce travail a toute son importance dans le cadre où comprendre les processus mis en place dans l'expressivité vocale peut participer à l'amélioration de la pratique clinique des logopèdes. Nous pouvons nous retrouver confronter à des patients présentant des déficits lors de l'expression émotionnelle vocale. Dans ce cas, connaître les modifications des paramètres vocaux peut les aider à améliorer ces déficits. Cette recherche s'inscrit également dans le développement grandissant des intelligences artificielles, plus précisément en ce qui concerne la synthèse vocale. En effet, toutes ces technologies présentent un manque de connaissances quant à la reconnaissance émotionnelle.

Dans la partie théorique de ce travail, je m'intéresserai d'abord aux paramètres vocaux principalement employés dans l'analyse vocale. Ensuite, une description des émotions sera faite, s'en suivront les liens entre les émotions et la voix. Grâce à cette section, j'aboutirai à la formulation de mes objectifs et hypothèses de travail.

II. Les caractéristiques acoustiques de la voix

La production vocale met en œuvre l'appareil respiratoire, le système phonatoire et l'appareil articulatoire. Ces systèmes sont dès lors employés afin de produire la voix, mais également dans la respiration et la déglutition. C'est la combinaison des muscles abdominaux, des muscles dorsaux, des poumons, du larynx, de la cage thoracique et des cavités nasale et buccale, qui permet aux humains de communiquer entre eux, de transmettre leurs émotions, sentiments ou pensées (Sataloff, R.T., 1995). Il est donc nécessaire de tenir compte de l'ensemble de ces facteurs lorsque nous réalisons des analyses acoustiques vocales. La phonation provient d'une suite de mouvements d'accolements et d'ouvertures des plis vocaux produisant des ondes sonores. Cela dit, depuis plusieurs années, la parole « naturelle » est opposée à la parole « de laboratoire ». En effet, l'analyse de parole « naturelle » est la source d'une importante variabilité de sons, par le fait qu'elle soit non-contrôlée mais surtout parce qu'elle est affectée par un grand nombre de variations majoritairement influencées par une multitude de sources externes, telles que le bruit de fond ou l'environnement influençant positivement ou négativement les interlocuteurs (Meunier, C., 2007).

La voix peut être analysée selon certains paramètres dont les principaux sont la hauteur, le niveau de pression sonore, le timbre, le débit et la prosodie.

A. L'analyse spectrale

Afin d'analyser un son sur ordinateur, on réalise un spectrogramme. Celui-ci permet d'illustrer l'ensemble de la décomposition spectrale de la voix. On peut y observer les harmoniques via les différentes intensités de couleurs (plus claires ou plus foncées). Par conséquent, il nous est loisible d'observer l'évolution des harmoniques en fonction du temps. Débuter par l'analyse de la voyelle /a/ est enrichissant grâce à la distribution riche des harmoniques, fournissant un nombre important de données telles que la fréquence fondamentale, l'attaque et l'extinction, les formants, les ruptures de sonorisation, les instabilités, les bruits, la durée du son et la dégradation des composantes spectrales (Menin-Sicard, A. Sicard, E., 2016).

B. Le niveau de pression sonore

Le niveau de pression sonore (NPS) correspond à l'amplitude de variation de la pression sonore, elle varie donc en fonction de la pression sous-glottique (Le Huche, F., Allali, A., 2010). En voix conversationnelle, le NPS se situe aux environs de 60 dB (Juslin, P.N., Laukka, P., 2001). Cela dit, le NPS a tendance à être plus élevé lors d'un état d'excitation plus important. S'opposant à une baisse du NPS lors d'un état émotionnel à faible niveau d'excitation. Le NPS est corrélé à l'amplitude de vibration des plis vocaux, ces deux aspects évoluent dans le même sens mais ne sont pas proportionnels (Belyk, M., Brown, S., 2014).

C. La hauteur tonale

La hauteur du son correspond au nombre de périodes par seconde et est exprimée en Hertz. Plus le son est aigu, plus le nombre de cycles par seconde est élevé.

1. La fréquence fondamentale

La fréquence fondamentale (F_0) est une mesure du nombre de cycles vibratoires des plis vocaux par seconde. Celle-ci correspond à la première harmonique d'un son. La F_0 se définit comme étant le fondamental correspondant à la hauteur tonale. Celle-ci peut varier par les modifications de tension et de rigidité des cordes vocales et par la pression d'air (Titze, I.R., Liang, H., 1994, cité par Morsomme, D. 2001). Généralement, l'analyse de la fréquence fondamentale se fait sur base d'une voyelle tenue, de la lecture d'un texte phonétiquement équilibré ou de la parole spontanée (Ghio, A., 2007). Détecter correctement la F_0 de la voix est une condition sine qua non dans l'évaluation vocale car cela permet de repérer des altérations, d'identifier l'étendue vocale et de déterminer le fondamental usuel moyen, c'est-à-dire la moyenne des fréquences fondamentales instantanées.

Jusqu'à la puberté, une descente du larynx s'effectue, accentuant les différences histologiques des cordes vocales entre les hommes et les femmes. Pendant la puberté, la taille et la forme du larynx et des plis vocaux sont modifiées. Les plis vocaux augmentent en longueur d'environ 60% pour les garçons, et 34% pour les filles (Kahane, J.C., 1987. Cité par Morsomme, D, 2001). En général, à l'âge adulte, la fréquence fondamentale des femmes se situe entre 200 et 250 Hz tandis que celle des hommes se trouve entre 110 et 140 Hz. Ces valeurs s'expliquent par la différence de longueur et d'épaisseur des plis vocaux (Le Huche, F., Allali, A., 2010).

La F0 augmente lors d'un état émotionnel de forte excitation, tandis qu'elle diminue lors d'un état émotionnel à faible excitation. La F0 permettrait donc à elle-seule d'identifier des émotions de peur ou de soulagement, de colère ou de joie (Patel, S., Scherer, K.R., Sundberg, J. et Björkner, E., 2011).

2. Les harmoniques

La fréquence fondamentale contient des harmoniques, correspondant aux multiples de la fréquence fondamentale, également visibles sur un spectrogramme. Par exemple, la deuxième harmonique est égal au doublement de la fréquence fondamentale. Une voix très riche en harmoniques semble plus ronde et chaude qu'une voix pauvre en harmoniques.

3. Les formants

Les harmoniques peuvent se regrouper dans une zone fréquentielle prédéfinie afin de constituer les formants. Différents points de résonance du conduit vocal (des plis vocaux jusqu'à la bouche) sont à l'origine des formants, ce conduit mesure entre 8 et 18 cm en fonction du genre et de l'âge du participant. Le conduit vocal résonne de manière naturelle dans des zones de résonance, appelées formants, axées sur les fréquences f_1 , f_2 , f_3 , etc. (Menin-Sicard, A., Sicard, E., 2016). Sur le spectre d'un signal vocal, les formants sont des zones où apparaît un maximum d'énergie, surtout lors de la production de voyelles (Le Bagousse, S. 2008). Le son produit est donc décomposé en harmoniques mais l'amplitude de celles-ci est modulée selon leurs fréquences, produisant les formants. La valeur fréquentielle des différents formants est influencée par la configuration de la cavité de résonance ainsi que la voyelle employée (Ghio, A., 2007).

Un spectrogramme à bandes larges est plus précis pour identifier l'amplitude donc les formants, tandis qu'un spectrogramme à bandes étroites est plus précis pour identifier les fréquences, donc les harmoniques (Meunier, C., 2007).

Lors de l'analyse phonétique, il est courant de se limiter aux valeurs des formants f_1 et f_2 car elles suffisent à caractériser et reconnaître une voyelle (Le Bagousse, S. 2008). Cependant, il existe un formant particulier, appelé formant du chanteur. Ce formant désigne généralement une bande fréquentielle renforcée, avoisinant les 3000 Hz (Sundberg, J., 1972), zone la plus sensible de l'oreille humaine (Le Bagousse, S. 2008). Ce formant est recherché

dans les voix chantées car il permet aux chanteurs d'être entendus malgré l'orchestre, qui présente une fréquence avoisinant les 450 Hz au maximum (Sundberg, J., 1972).

4. Le triangle vocalique

Les formants sont donc dépendants de la forme du conduit vocal. Sa forme change en fonction de la position de la langue et de l'aperture buccale. Ces changements entraînent une modification des valeurs des fréquences de résonances des formants selon le triangle vocalique. Ce dernier réalise un parallèle entre la position de langue (antérieure/postérieure) et celle de la bouche (ouverte/fermée). Dans ce triangle, l'axe des Y représente le premier formant ($f1$), correspondant au degré d'aperture buccale, avec une fréquence décroissante. Sur l'axe des X, on peut placer le deuxième formant ($f2$) qui représente la position linguale, avec fréquence croissante (Le Huche, F., Allali, A., 2010).

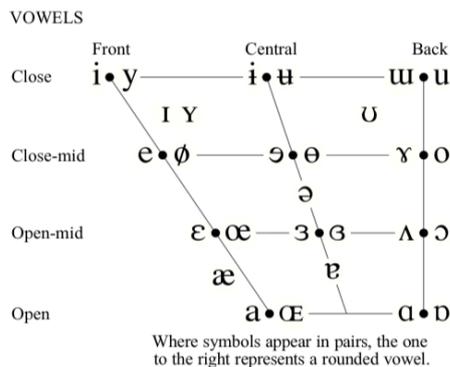
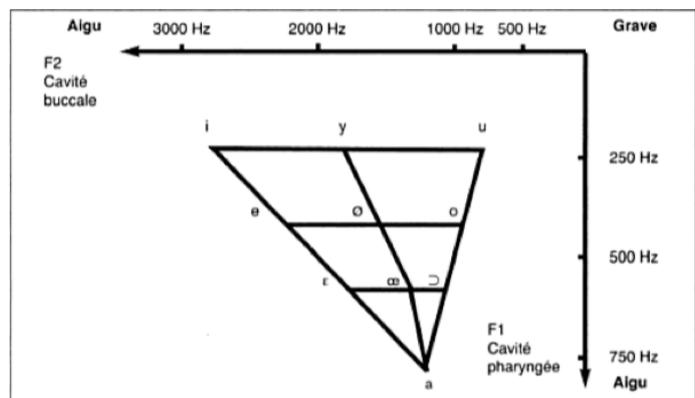


Figure 1. Triangle vocalique des voyelles (The international phonetic alphabet, revised to 2015)

Figure 2. Valeurs formantiques des voyelles du français (Méchain, C., 2014)



D. Le timbre

La définition du timbre soutient que lorsque deux sons sont équivalents en hauteur tonale et en NPS mais semblent différents auditivement, ils diffèrent en timbre (Sundberg, J., 1994). Il correspond à la répartition et à l'intensité des harmoniques et du bruit dans la production d'un instrument ou de la voix (Le Huche, F., Allali, A., 2010). D'un point de vue vocal, le timbre est composé de plusieurs éléments : la longueur et la structure des cordes vocales, la qualité de l'accolement des cordes vocales, le volume et la configuration des résonateurs (Heuillet-Martin, G., Garson-Bavard, H. et Legré, A., 1995). Le timbre fournit une grande quantité d'informations concernant la voix, comme la position des articulateurs ainsi que l'état et le volume des résonateurs car ceux-ci peuvent amplifier ou inhiber certaines harmoniques. Le timbre est lié à la source vocale et aux formants, grâce à ces combinaisons, nous obtenons des caractéristiques de la voix du locuteur, telles que l'effort vocal ou le mécanisme employé (Sundberg, J., 1994).

E. L'instabilité en hauteur – le Jitter

L'indice *jitter* correspond à une variation cycle à cycle, à court terme de la fréquence fondamentale (Titze, I.R., 1992). Le *Jitter* se caractérise par une altération de la périodicité de la fréquence fondamentale (Le Bagousse, S. 2008). L'évaluation de ce paramètre est généralement réalisée sur la production de la voyelle /a/ et certaines normes ont été établies afin d'identifier le caractère pathologique de la voix. Les valeurs typiques du *jitter* se situent entre 0,5 et 1,0 % en phonation soutenue chez le jeune adulte. Ce pourcentage correspond au pourcentage moyen de variation de fréquence cycle à cycle. Cependant, la valeur de *jitter* diffère entre les hommes et les femmes. Pour les hommes, le *jitter* est de 0,7 % tandis que pour les femmes, il est de 0,5 % (Giovanni, A., Robert, D., Estublier, N., Teston, B., Zanaret, M., Cannoni, M., 1996). Les valeurs du *jitter* augmentent lorsque la voix est pathologique (Teixeira, J.P., Oliveira, C., Lopes, C., 2013).

F. L'instabilité en intensité – le Shimmer

L'indice *shimmer* est modifié avec une diminution de la résistance glottique et est corrélé avec la présence d'un souffle (Teixeira, J.P., Oliveira, C., Lopes, C., 2013). Le *shimmer* se caractérise plus précisément par une « variation d'amplitude entre périodes successives

quand un individu tente de tenir la phonation à une fréquence et une intensité constantes » (Lhote, E., 1982). Le seuil pathologique du *shimmer* se situe à des valeurs inférieures à 3% pour les adultes, mais entre 0,4 et 1% pour les enfants (Guimarães, I., 2007, citée par Teixeira, J.P., Oliveira, C., Lopes, C., 2013).

G. L'étendue vocale

L'étendue vocale (Voice Range Profil – VRP) correspond à l'étendue entre les fréquences les plus basses et les plus élevées qu'une personne peut produire. Elle est apparentée à la tessiture chez le chanteur, à la différence que celle-ci exclut les notes extrêmes inconfortables. La tessiture vocale est donc la partie de l'étendue vocale qu'une personne peut produire à pleine voix, de manière régulière et confortable.

De manière plus précise, l'étendue vocale est une représentation graphique de la capacité d'un locuteur à produire des niveaux de pressions sonores minimales et maximales à des pourcentages fixes de son étendue fréquentielle maximale de phonation (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002). Le VRP est donc un affichage des NPS par rapport à la fréquence fondamentale. Généralement, le niveau de pression sonore en dB, est tracé verticalement et la fréquence est représentée horizontalement de manière linéaire ou logarithmique. Dans cet affichage, l'étendue des NPS varie avec l'étendue de fréquences (Titze, I. 1992). Afin de calculer l'étendue vocale d'un participant, celui-ci est invité à produire une sirène ascendante, de la note la plus grave à la plus aiguë, ainsi qu'une sirène descendante, de la note la plus aiguë à la note la plus grave afin d'obtenir une représentation fidèle des performances en hauteur du participant.

Différents types de VRP ont vu le jour ces dernières années afin de pallier aux manquements du VRP traditionnel. Le VRP normatif basé sur des transformations mathématiques, et qui fait partie de ces nouveautés, contient 95 % d'intervalles de prédiction pour les valeurs de NPS et de fréquences. Le VRP normatif permet d'obtenir une étendue vocale de forme ovale typique (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002). En effet, malgré les différences interindividuelles au niveau des voyelles et du sexe, le VRP conserve une forme bien caractéristique (Gramming, P., Sundberg, J., Terström, S., Leanderson, R., et Perkins, W.H., 1988). Ce VRP normatif a été établi sur des professionnels de la voix, locuteurs dont tout problème vocal entrave l'exécution anormale de la profession (Koufman, J.A. et Isaacson, G., 1991, cités par Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002). Afin de mesurer l'étendue vocale de leurs participants, ceux-ci ont

invité les locuteurs à explorer leurs capacités vocales dans les fréquences les plus élevées et les plus basses. Le VRP normatif est établi selon quatre étapes, la première consiste à convertir les fréquences en demi-tons afin de pallier un biais logarithmique fréquentielle. La deuxième étape consiste à faire débiter toutes les étendues vocales au même demi-ton. La troisième étape consiste, quant à elle, à faire finir toutes les étendues vocales sur le même demi-ton. Finalement, la quatrième étape correspond à l'interpolation des étendues vocales individuelles pour établir les intensités correspondant à chaque demi-ton, suivie du calcul des intervalles de prédiction à 95% pour chaque demi-ton. La transformation des demi-tons en fréquences correspond à l'étape ultime de ce VRP normatif.

Grâce à cette étude, Heylen, Wuyts, De Bodt et Van de Heyning ont pu établir des normes d'étendue vocale pour des enseignants. En effet, pour les enseignantes, l'étendue vocale se situe entre 156 et 880 Hz tandis que pour les enseignants, celles-ci se situent entre 77,8 Hz et 622 Hz (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002). En demi-tons, les femmes ont une étendue vocale de 2 octaves et 7 demi-tons, et les hommes de 3 octaves.

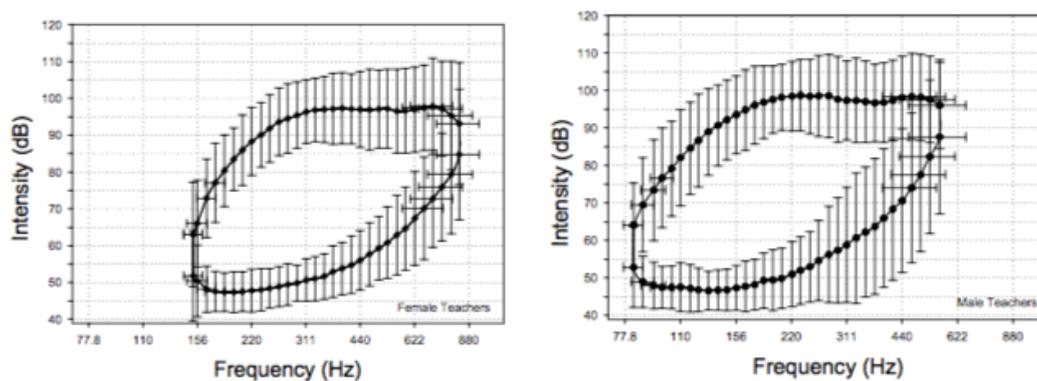


Figure 3. Étendues vocales des enseignants (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002)

En regard des figures 3 et 4, à gauche sont représentées les normes d'étendue vocale pour les enseignantes. A droite sont représentées les normes pour les enseignants. Les lignes horizontales et verticales représentent les 95% d'intervalles de prédiction. Ces intervalles comprennent 95% des valeurs d'intensité et de fréquence de la population (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002).

Ces étendues vocales typiques sont en accord avec celles établies par Gramming en 1988. En effet, celle-ci a établi des courbes pour les participants masculins et féminins sur les voyelles /a/ et /i/. D'après la forme ovale de l'étendue vocale, il est mentionné explicitement

H. Les différents mécanismes

L'étendue vocale représente donc l'ensemble des hauteurs, en fréquences, qu'un participant peut produire sans forçage. Les différents registres se rapportent aux différents modes de vibrations des cordes vocales. On recense quatre mécanismes : qui sont, tout d'abord, le mécanisme 0 appelé le *fry*, qui est le registre le plus grave, dans une tonalité plus grave que la voix parlée. Ensuite, le mécanisme I appelé *voix de poitrine*, qui est le plus utilisé en voix parlée ou chantée. Le mécanisme II est appelé *voix de tête* pour les femmes et *voix de fausset* pour les hommes, la fréquence employée est supérieure à celle de la voix parlée. Finalement, le mécanisme III, appelé *voix de sifflet*, correspond au plus haut registre avec des aigus extrêmes (Le Bagousse, S. 2008).

I. Le rapport harmonique bruit (HNR)

Le rapport harmonique bruit, désigné par les initiales HNR (harmonic to noise ratio), représente le rapport entre l'amplitude des harmoniques et l'amplitude du bruit exprimée en décibels. En d'autres termes, on évalue la puissance du bruit par rapport aux harmoniques. (Barsties, B., Maryn, Y., 2015). Pour calculer celui-ci, on emploie une méthode temporelle (Ghio, A., 2007) où deux composants interviennent. Tout d'abord, un composant provenant de la vibration des plis vocaux et un second composant provenant du bruit glottal. Un HNR élevé est le signe d'une phonation efficace tandis qu'un faible HNR représente une voix asthénique ou une dysphonie. Une production vocale présentant un HNR inférieur à 7 dB est considéré comme pathologique (Teixeira, J.P., Oliveira, C., Lopes, C., 2013).

III. Les émotions

A. Définition d'une émotion

Initialement, un épisode émotionnel demande une évaluation pertinente de l'évènement en fonction du but. Cet épisode crée ensuite un état de préparation à agir en urgence qui interrompt les autres actions en cours. Finalement, survient l'expérience subjective qui active un état mental différent de celui dans lequel l'individu est habituellement (Frijda, N.H., 1988).

Actuellement, l'émotion est considérée comme un système à trois composantes dont la première correspond aux réponses physiologiques. Ces réponses peuvent se traduire par une accélération cardiaque, une augmentation de la température corporelle ou par une modification d'un rythme et de l'intensité respiratoire, par exemple. Ensuite, la deuxième composante se penche sur les réponses comportementales et expressives, c'est-à-dire les changements, volontaires ou involontaires du visage, de la voix et des postures qui surviennent après un épisode émotionnel. A noter, que cet aspect se focalise essentiellement sur les modifications du visage car c'est par ce biais que l'individu communique le plus son émotion à autrui. En effet, dans le cas de la surprise par exemple, les sourcils s'élèvent subitement. Cela dit, l'ensemble du corps et de la posture participe à cette communication émotionnelle. Par exemple, lorsque l'individu est en colère, le corps et les mains sont très agités, tandis que dans le cas de la tristesse, le corps est en retrait et les mouvements sont lents. Un autre élément important pour partager une émotion est la voix qui subit des modifications émotionnelles qui seront explicitées plus tard. La troisième composante concerne les réponses cognitives et expérientielles où le terme cognitif renvoie aux processus grâce auxquels l'organisme reçoit des informations environnementales et y répond adéquatement. En revanche, le terme expérientiel correspond au vécu subjectif et unique d'une expérience émotionnelle (Lang, P.J., 1979).

B. L'émotion est-elle universelle ou propice aux variations culturelles ?

Une des grandes questions est de savoir si les manifestations émotionnelles sont universelles ou si elles subissent des variations culturelles. A ce sujet, trois points de vue sont retenus concernant la psychologie des émotions. Celui d'Ekman qui est plutôt en faveur d'une

conception universelle des émotions. Suivi d'Averill qui est plus en faveur de variations culturelles et de Scherer qui prétend à un mélange de ces deux perspectives (Luminet, O., 2007)

1. Ekman : conception universelle

Selon Paul Ekman, les émotions ont subi une évolution au cours de l'histoire en fonction des exigences de chaque période (Luminet, O., 2007). Il suggère que les émotions sont donc un produit de notre évolution, accompagnées de données biologiques. Des caractéristiques universelles permettraient de représenter un ensemble d'émotions de base (Ekman, P., 1992). Pour lui, il est difficile que le contexte social intervienne dans une réaction émotionnelle parce que tout sentiment subjectif est précédé par un mécanisme d'évaluation cognitive automatique et inconscient. De plus, les éléments contextuels qui modifient la réaction émotionnelle sont également universels (Luminet, O., 2007). Cela dit, il ne nie pas totalement les processus culturels et d'apprentissage social qui, selon lui, interviennent dans le contrôle des expressions émotionnelles, la manière d'évaluer une situation émotionnelle, ou comment faire face à une émotion. Sa théorie est appuyée par l'universalité des expressions faciales. Par exemple, les expressions faciales représentant la peur ou la colère sont identiques chez les primates. Il met également en avant le caractère physiologique de l'émotion qui profère à l'individu la possibilité de réagir très rapidement (Ekman, P., 1992). Malgré tous ces arguments, les théories d'Ekman sont très décriées par la communauté scientifique. De plus, Ekman ne tient compte que de deux composantes émotionnelles et laisse de côté la composante cognitive et expérientielle de l'état émotionnel (Luminet, O., 2007).

2. Averill : variations culturelles

Averill s'oppose à la théorie universaliste d'Ekman car il lui reproche de ne prendre en compte que le caractère biologique dans la définition des émotions de base, renvoyant au caractère vital des émotions dans la survie des espèces (Luminet, O., 2007). En effet, selon lui, une émotion est effectivement vitale d'un point de vue biologique. Mais elle l'est tout autant d'un point de vue social ou psychologique dans l'optique que selon la manière dont une émotion affecte une personne, cela constitue également un élément fondamental dans le fonctionnement émotionnel et humain. Des facteurs de personnalité et de culture ont été relevés lors d'une étude sur l'expression émotionnelle entre différentes cultures qui sont la culture japonaise et la culture américaine. Les participants étaient confrontés à des stimuli visuels provoquant un choc

émotionnel. Même si certains éléments tels que la nature ou l'intensité de la réaction étaient identiques, d'autres facteurs comme le niveau d'activation émotionnel diffèrent entre les cultures, les Japonais étant en constante activation même sans scène choquante (Averill, J.R., Opton Jr, E.M., Lazarus, R.S., 1969). Il est vrai que, de manière générale, les émotions ayant le plus marqués les personnes sont celles qui les ont bouleversées. Cependant, il est reproché à Averill de ne pas appuyer ses dires avec suffisamment de données empiriques (Luminet, O., 2007).

3. Scherer : conciliation de l'universel et du culturel

Scherer vient marquer un compromis entre ces différents points de vue en soutenant à la fois une théorie d'universalité des schémas émotionnels différentiels ainsi qu'une théorie prônant des différences culturelles importantes dans l'expressivité et la régulation émotionnelle, la représentation symbolique de l'émotion, ainsi que le partage social autour des émotions (Scherer, K.R., 1994). Il appuie d'abord cette théorie par le fait que des émotions se produisent de manière similaire chez des personnes de cultures différentes (Luminet, O., 2007). D'après les travaux réalisés afin d'élucider cette opposition entre l'universalité et l'aspect culturel, il a été mis en évidence que la spécificité et l'aspect culturel des émotions pouvaient expliquer de nombreuses différences au niveau des interactions émotionnelles (Scherer, K.R., 1994). En conclusion, en réponse à un stimulus interne ou externe, une émotion va engendrer une séquence de changements reliés dans l'ensemble de l'organisme. Le processus émotionnel serait donc un processus cumulatif où les différents composants émotionnels se synchronisent et puis se séparent, en dépendant toujours d'aspects individuels (Luminet, O., 2007). Scherer marque l'importance des représentations sociales, qui dépendent des schémas biopsychologiques universels. En effet, dans le cas contraire, si les représentations sociales étaient indépendantes de ces schémas, elles seraient très différentes d'une culture à l'autre (Scherer, K.R., 1994).

Comme tout processus, l'émotion engendre des modifications physiologiques, ayant pour point de départ les structures composant le système limbique.

C. Le système limbique

Le système limbique correspond à l'ensemble des structures cérébrales qui participent aux réactions émotionnelles et aux comportements motivés, c'est-à-dire réalisés pour atteindre un but. Celui-ci est composé de l'hypothalamus, du corps amygdaloïde et de l'hippocampe. Ce système joue un rôle majeur au niveau émotionnel, surtout en ce qui concerne les émotions de rage et de peur (Tavris, C., Wade, C., 2007). Il nous permet de réagir sans délai face à un danger ou une menace (Damasio, A.R., 2010).

1. L'hypothalamus

L'hypothalamus est situé sous le thalamus et, malgré sa petite taille, permet de réguler les émotions, la faim, la soif et la reproduction qui sont des pulsions indispensables à la survie. Cette structure contrôle la température par la production de sueur ou de frissons et gère les différentes opérations complexes établies dans le système nerveux autonome (Tavris, C., Wade, C., 2007).

2. Le corps amygdaloïde

Le corps amygdaloïde est une structure qui s'occupe de stimuler et réguler les émotions, en s'appuyant sur l'association de souvenirs apportant du sens à l'émotion. Il sert donc à évaluer les stimulations sensorielles et à les relier aux composantes sensorielles en mémoire. Cette structure est responsable de réponses émotionnelles rapides lors de changements brusques de situations (Tavris, C., Wade, C., 2007). Il joue un rôle essentiel dans l'expression de l'anxiété et de la dépression. Lorsque ces deux éléments sont proéminents chez une personne, l'activité électrique est élevée dans le corps amygdaloïde (Drevets, W.C., 2000). En revanche, si cette structure est endommagée, les personnes présentent souvent des difficultés à interpréter et percevoir leur propre peur ou celle des autres (Damasio, A.R., 2010).

3. L'hippocampe

L'hippocampe permet de stocker en mémoire les informations nouvelles, mais également de comparer les stimulations sensorielles avec celles que le cerveau reçoit habituellement. Si elles sont similaires, l'hippocampe transmet l'ordre de rester calme (Tavris, C., Wade, C.,

2007). En plus de cela, l'hippocampe a un rôle très important dans la formation, en mémoire, des souvenirs de lieux et des représentations de l'espace (Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J. et Frith, C.D., 2000)

D. Perception et expression émotionnelles

Les émotions nous permettent de communiquer rapidement nos différents besoins. En effet, l'expression émotionnelle traduit notre état interne en produisant des indices visibles de l'extérieur. Cela suscite une réaction des interlocuteurs, telle que poursuivre leurs actions si l'expression est positive ou les dissuader de continuer si l'expression est négative. Cela permet donc d'améliorer la coordination des groupes sociaux (Blair, R.J.R., 2003). La perception émotionnelle se fait de manière automatique à l'aide de normes sociales établies. Cette faculté perceptive s'ajoute à la communication verbale tout en étant plus rapide et très efficace. Les humains communiquent leurs émotions à travers des modifications faciales, vocales, corporelles mais également grâce à la communication écrite, la musique ou l'art (Jang, D., Efenbein, H.A., 2015).

Les émotions résultent d'un stimulus extérieur qui attire notre attention. Afin de répondre à ces stimuli, nous sommes engagés dans un processus d'évaluation cognitive qui permet d'évaluer les stimuli suivant 5 critères (Frijda, N.H., 1988). Premièrement, la valence qui indique si le stimulus est positif, négatif ou neutre. Deuxièmement, la nouveauté précisant si le stimulus est familier ou non familier. Le troisième critère est l'obstruction de l'objectif qui identifie les obstacles pour atteindre ses objectifs. Quatrièmement, le potentiel contrôle du stimulus par soi-même ou une autre personne. Le cinquième et dernier critère est l'équité, qui est la mesure de conformité aux normes sociales ou personnelles (Smith, C.A., Ellsworth, P.C., 1985).

Les expressions émotionnelles sont de nature multimodale. En effet, les émotions étant composées d'états sentimentaux subjectifs et de réponses physiologiques (James, W., 1884), il est loisible d'identifier les marqueurs physiologiques de l'émotion. Ils sont composés des expressions faciales, des expressions vocales, des modifications de la peau, de la respiration, du rythme cardiaque, de la posture et de l'activité musculaire (Banse, R. et Scherer, K.R., 1996). C'est donc l'ensemble de ces indices qui permet de fournir des indices émotionnels aux autres mais nous sommes également capables de manipuler certains indices tels que les expressions faciales, vocales ou la posture. Parmi tous ces indices, les expressions faciales sont les plus

saillantes chez l'être humain grâce à leurs structures musculaires très complexes (Jang, D., Elfenbein, H.A., 2015).

Les émotions entraînent donc des changements de l'organisme en mobilisant un ensemble de ressources permettant de faire face à une situation, qu'elle soit positive ou négative. Cependant, le mécanisme émotionnel est pourvu d'un temps de latence permettant de jauger la suite de la situation, permettant par exemple d'éviter d'envenimer la situation et d'encourir des risques pour sa propre personne. Le temps de latence permet de prédire la réaction probable de son interlocuteur (Scherer, K.R., 1995).

IV. Voix et émotions

A. Liens entre les émotions et la voix

Il est connu que les émotions peuvent être décelées grâce à la voix par d'autres locuteurs. Cela dit, nous manquons de preuves empiriques démontrant que les humains sont capables de reconnaître et d'identifier une émotion dans un signal vocal isolé, en dehors du contexte situationnel. En général, les échantillons vocaux de types émotionnels sont reconnus correctement pour 60 % d'entre eux. En revanche, toutes les émotions ne sont pas reconnues de manière équivalente. En effet, la tristesse et la colère sont mieux identifiées que la peur et la joie. Le dégoût, quant à lui, est très mal reconnu, cela relève plus souvent du hasard (Scherer, K.R., 1995).

Des liens entre la personnalité et la voix ont été établis et permettent d'associer certaines pathologies vocales avec certains traits de personnalité. L'anxiété, le névrosisme et l'introversion sont corrélés avec les dysphonies dysfonctionnelles tandis que l'extraversion et l'impulsivité sont corrélées aux dysphonies organiques (Roy, N., Bless, D.M., Heisey, D., 2000).

Depuis plusieurs années, un débat portant sur les liens de causalité entre la personnalité et les troubles vocaux anime le domaine scientifique. En effet, pour des cas de dysphonies fonctionnelles, organiques ou spasmodiques, on ne sait pas si la personnalité, y compris l'inadaptation émotionnelle, est une des causes des troubles vocaux ou bien si les troubles vocaux entraînent des problèmes psychologiques et de personnalité (Roy, N., Bless, D.M., 2000).

Malgré ces interrogations, il est tout à fait reconnu que les émotions et la personnalité tiennent un rôle important dans l'origine des troubles vocaux. Par exemple, chez les patients dysphoniques, on a remarqué que certains traits de personnalité pouvaient altérer les capacités à vivre une émotion, qu'elle soit positive ou négative. De plus, la personnalité peut altérer, par un déséquilibre des tensions musculaires laryngées, la production vocale en la rendant impulsive, excessive ou inhibée. (Dietrich, M., Verdolini Abbott, K., Cartner-Schmidt, J., Rosen, C.A., 2008).

Dans le domaine des relations entre la voix et les émotions, une des grandes difficultés est de trouver un paramètre de mesure suffisamment sensible pour détecter les modifications de la voix dans des états émotionnels transitoires. Pour pallier ces difficultés, les études dans ce secteur emploient généralement des acteurs pour vocaliser les émotions. A force de travailler dans cette ligne de recherche, plusieurs facteurs ont été identifiés comme bons prédicteurs d'une émotion. Tout d'abord, une augmentation de la fréquence fondamentale et du NPS traduit une émotion avec un fort arousal, c'est-à-dire une forte intensité émotionnelle, telle que la colère. Au contraire, une diminution de la fréquence fondamentale et du NPS peut représenter une émotion telle que la tristesse. Une autre difficulté de ce domaine de recherche est que la majeure partie des émotions vocales produites par les acteurs n'est pas le reflet réel des vocalisations émotionnelles en environnement naturel, ou bien produites par des personnes qui ne sont pas des comédiens.

Malgré ces difficultés, l'étude des expressions émotionnelles vocales non verbales pourrait être un outil très intéressant pour comprendre le développement de certains troubles vocaux. En effet, la compréhension des différences de tempérament pourrait expliquer les différences vocales présentes chez les personnes souffrant d'un trouble vocal lorsqu'elles produisent une émotion vocalement. La théorie des traits des troubles de la voix de Roy (TTVD) postule que les personnes présentant des troubles vocaux acquis par le comportement, tels que des nodules ou une dysphonie fonctionnelle, possèderaient des différences de tempérament en comparaison avec des personnes atteintes de troubles vocaux comme une paralysie unilatérale du pli vocal (Roy, N., Bless, DM., et Heisey, D. 2000, cités par Van Mersbergen, M. et Delany, M. 2014). Pour formuler cette hypothèse, les auteurs de la théorie des traits des troubles de la voix ont effectué des mesures psychométriques du tempérament et ont constaté que les participants possédant des nodules des plis vocaux présentaient un tempérament accompagné d'une extraversion accrue, d'un neuroticisme ainsi que d'une diminution de l'inhibition comportementale. En ce qui concerne les personnes atteintes d'une dysphonie fonctionnelle, elles présentaient un tempérament accompagné d'une introversion, d'un neuroticisme et d'une augmentation de l'inhibition comportementale. Cela signifie que ces personnes inhiberaient leurs expressions émotionnelles, surtout au niveau des muscles du visage, ce qui diminue leurs capacités à communiquer leurs émotions. Ces différences de tempérament sont la preuve que les troubles de la voix comportementaux pourraient être le résultat de l'expression de certains traits de la personnalité (Van Mersbergen, M. et Delany, M. 2014).

Même si la communication émotionnelle vocale est moins étudiée que l'expression faciale, les recherches ont démontré que pour de nombreuses espèces animales, les différents états émotionnels affectifs sont exprimés par des vocalisations et permettent de communiquer de manière spécifique les besoins immédiats de l'individu (Scherer, K.R., 1995). D'ailleurs, nous retrouvons dans la parole humaine, un ensemble d'interjections spontanées communes à beaucoup de cultures, telles que « Ouch », « Aïe », « Berk ». (Kleinpaul, R., cité par Scherer, K.R., 1995).

B. Paramètres vocaux influencés par l'état émotionnel

L'ensemble des indices acoustiques véhiculant l'émotion vocale sont la fréquence fondamentale (hauteur tonale), le niveau de pression sonore, le rythme (durée des segments et des pauses) et le timbre (distribution de l'énergie spectrale) (Banse et Scherer, 1996 ; Grandjean et al., 2006, cité par Liebenthal, E., Silbersweig, D.A. et Stern, E., 2016). Ces paramètres dépendent largement de facteurs physiologiques tels que la fréquence cardiaque, le flux sanguin et la tension musculaire, qui varient eux-mêmes avec l'état émotionnel d'une personne. La prosodie (fréquence, contour sonore et rythme) et la qualité de la vocalisation non verbale (timbre et changements spectraux) doivent transmettre l'état émotionnel d'une personne (Banse et Scherer, 1996 ; Grandjean et al., 2006, cité par Liebenthal, E., Silbersweig, D.A. et Stern, E., 2016). Grâce à certains de ces paramètres, la tristesse serait, par exemple, traduite par une hauteur tonale plus grave, une variation faible du niveau de pression sonore et un discours lent, alors que le bonheur et la colère seraient exprimés par une variation forte du niveau de pression sonore (Enquête de Pell, Paulmann et al. 2009, cité par Altrov, 2013).

Actuellement, de nombreuses études démontrent que les émotions individuelles peuvent être caractérisées par des signaux acoustiques (Juslin et Laukka 2003; Patel et Scherer 2013; Scherer 2003) mais très peu d'études se sont penchées sur la perception vocale de ces signaux acoustiques. Ceci peut être dû à la difficulté d'évaluer de manière fiable la perception des caractéristiques vocales par un auditeur, ainsi qu'au manque de modèle consensuel des mécanismes des effets de l'émotion sur la voix et la parole dans le langage normal. Bänziger, T., Patel, S. et Scherer, K.R (2014) ont tenté d'établir une liste standardisée des fonctions vocales pouvant être évaluées de manière fiable par des auditeurs non-experts. Pour atteindre cet objectif, ils ont réalisé deux études en parallèle. Dans la première étude, ils emploient un panel de 144 expressions vocales, produites par 9 acteurs germanophones, réparties en 8 catégories émotionnelles. Dans la seconde étude, ils ont employé 160 expressions vocales,

provenant d'un autre corpus d'expression, produites par 10 acteurs francophones. Pour les deux études, deux groupes de participants ont été sélectionnés pour écouter et classer les 144 et 160 expressions vocales présentées à travers des haut-parleurs. Dans la première étude, les participants devaient indiquer leurs réponses, correspondant à la valence de l'émotion, sur un ordinateur, sur une échelle analogique allant de 0 à 10. En revanche, pour la seconde étude, cette échelle allait de 0 à 100. Finalement, ils ont sélectionné les paramètres les plus fiables permettant l'identification des émotions. Pour les deux études, il s'agit du « loudness », du taux de parole, de l'intonation et de la hauteur. Ces paramètres vocaux permettent de discriminer les huit émotions incluses dans les deux études avec un haut niveau de précision. Cette recherche confirme la faisabilité de l'évaluation des indices vocaux perçus dans la communication émotionnelle (Bänziger, T., Patel, S. et Scherer, K.R (2014).

C. La perception émotionnelle

Dans la communication orale, la majorité des actes de communication implique des émotions. Leur reconnaissance s'avère indispensable à la compréhension du message, du comportement ou de l'état émotionnel de la personne (Cowie et al. 2011, cité par Altrov 2013). Le degré de précision dans la reconnaissance des émotions est similaire pour l'expression faciale que pour l'expression vocale (Scherer et al., 2011, cité par Bänziger, T., Patel, S. et Scherer, K.R. 2014). Il en est de même pour le temps de perception nécessaire au traitement neuronal (120 ms), cela laisse penser que les traitements émotionnels vocaux et faciaux sont régis par les voies neuronales supramodales (Sauter et Eimer, 2010, cités par Liebenthal, E., Silbersweig, D.A. et Stern, E., 2016). Cependant, le bonheur et le dégoût sont mieux reconnus dans le visage, tandis que la tristesse et la colère sont mieux perçues à travers la voix (Scherer et al., 2011, cité par Bänziger, T., Patel, S. et Scherer, K.R. 2014). Cela dit, la compréhension émotionnelle vocale est d'autant plus importante qu'elle permet de décoder certains messages lorsque l'interlocuteur n'est pas observable (Cheang et Pell 2008, cité par Altrov 2013). Néanmoins, elle ne s'avère pas toujours aisée, puisqu'il a été démontré que les différences culturelles peuvent compliquer l'analyse de l'état émotionnel de son interlocuteur. Autrement dit, partager la même culture facilite la communication émotionnelle (Soto et Levenson, 2009, cité par Altrov, 2013). Malgré cela, les individus sont capables d'identifier un panel d'émotions de base (la colère, le bonheur, la tristesse, la peur et le dégoût) au-delà de leurs cultures et de leurs langues (Bryant et Barret 2008, cité par Altrov, 2013). La faculté des auditeurs à décoder l'émotion vocale dans une langue inconnue prouve d'ailleurs l'existence d'indices prosodiques

universels, notons que ceux-ci sont améliorés dans une culture et langue commune (Thompson et Balkwill, 2006, cité par Altrov, 2013). Pell et al. (2009, cité par Altrov, 2013) ajoutent que les émotions les plus reconnues universellement sont la colère et la tristesse, tandis que le dégoût est moins bien perçu.

D. Perception émotionnelle interculturelle

Comme le montrent les résultats de l'étude d'Altrov (2013), qui porte sur les indices permettant d'identifier une émotion lorsque des personnes habitent un même pays mais parlent différentes langues, il y a une influence culturelle dans la reconnaissance émotionnelle. Les deux questions de recherche étaient de savoir si les Russes qui vivent en Estonie et dont la langue seconde (L2) est l'estonien reconnaissent les émotions exprimées en estonien et si les Russes vivant en Estonie reconnaissent les émotions de la même manière que les Estoniens ou les Russes qui vivent en Russie. Le matériel de passation a été extrait du corpus de parole émotionnel estonien, où les énoncés ont été isolés du contexte. Le test a été réalisé via un site web, où les participants devaient décider si le ton du locuteur était joyeux, fâché, triste ou neutre. Les résultats ont montré que la compréhension des émotions dépend de normes culturelles, apprises lors d'échanges sociaux. En effet, les émotions n'ont pas toujours été reconnues de la même manière entre les participants dont la langue maternelle est l'estonien et les participants dont l'estonien est leur seconde langue. Ceci s'explique par le fait que les caractéristiques vocales émotionnelles varient d'une langue à l'autre.

E. Perception et production émotionnelle selon l'âge

L'étude de Söderstrom et al. s'interrogeait sur les capacités de très jeunes enfants à discriminer le soulagement et la victoire, exprimés de manière non-verbale. Les bébés sont âgés de 4, 6 ou 8 mois, et inclus sur base d'une liste de critères. Trois expériences sont conduites. Dans la première, le groupe de participants a toujours été exposé à la même personne qui n'a jamais exprimé ses émotions non-verbale. Un locuteur adulte produit le soulagement, par le soupir, et la victoire, par des acclamations. Il a été démontré que les bébés de 6 mois parvenaient à distinguer le soulagement de la victoire. Dans la deuxième expérience, les bébés étaient, contrairement à la première expérience, exposés à quatre locuteurs, deux féminins et deux masculins, durant la phase d'habituation. Ceci a été fait dans le but que les bébés n'aient pas un seul locuteur de référence, de manière à observer leurs capacités d'adaptation. Les

résultats de cette seconde expérience ont montré que les bébés, dès 6 mois, étaient tout à fait capables de catégoriser une émotion, sur bases de signes comportementaux, et ce chez différents locuteurs. De plus, ils sont en mesure de s'adapter à chaque locuteur pour en traduire leur état émotionnel. Dans la troisième expérience, les signes émotionnels présentés pendant le test différaient de ceux présentés pendant la phase d'adaptation, les locuteurs n'employaient pas les mêmes attitudes émotionnelles vocales de par leur identité culturelle et linguistique différente. Suite à ces changements, les bébés de 6 mois ne parvenaient pas toujours à déceler l'émotion présentée (soulagement ou victoire) dans une langue étrangère. Seuls les enfants de 10 à 12 mois parvenaient à discriminer ces deux catégories d'émotions lorsqu'elles étaient présentées par des locuteurs de cultures étrangères (Söderstrom et al. 2017).

La capacité à percevoir les émotions se développe très tôt, en même temps que la période d'enrichissement lexicale, cette idée a été soutenue par les recherches menées en imagerie cérébrale (Blasi et al., 2011; Grossman, Oberecker, Koch et Friederici, 2010, cité par Söderstrom et al. 2017). De plus, dès cinq ans, les enfants sont capables de produire une gamme d'états émotionnels sur base de vocalisations non linguistiques (cris, rires, soupirs) (Sauter, Eisner, Calder et Scott, 2010; Sauter, Panattoni, & Happe, 2013; Sauter & Scott, 2007, cité par Söderstrom et al, 2017). L'ensemble de ces vocalisations non verbales peuvent être illustrées par le Modèle Organon de Bühler, celui-ci stipule que les différentes expressions vocales peuvent soit représenter un symptôme de l'état du locuteur, c'est-à-dire les émotions, les intentions ou les attitudes, soit servir de signal poussant l'auditeur à produire une réaction, soit servir de symbole qui représente un objet ou évènement joyeux ou triste par exemple (Scherer, K.R., 1988, cité par Laukka et Elfenbein, 2012). L'étude de Laukka et Elfenbein (2012), portant sur la fonction symbolique de l'expression vocale, stipule que les émotions sont provoquées et évaluées par l'observateur comme étant la représentation cognitive d'un évènement antérieur (Johnstone, van Reekum et Scherer, 2001 ; Scherer, 1986, cité par Laukka et Elfenbein 2012). Pour constituer le matériel de cette étude, les chercheurs ont d'abord demandé à des acteurs de jouer des émotions vocales en s'adaptant à différentes situations émotionnelles comme la joie ou la tristesse. Les acteurs ont reçu comme consignes de produire des énoncés courts au contenu neutre tout en exprimant l'émotion adaptée à la situation. Ensuite, les acteurs devaient évaluer leurs expressions produites précédemment afin de les valider. Suite à cela, les chercheurs ont pu sélectionner des paramètres vocaux permettant d'analyser les productions émotionnelles. Ces paramètres sont la fréquence fondamentale, le niveau de pression sonore, les formants, les marqueurs temporels et la qualité de la voix (Juslin & Laukka, 2003, cité par Laukka et Elfenbein 2012). Lors de la phase de test, les auditeurs devaient écouter une série de phrases

prononcées par les différents acteurs et répondre à des questions en donnant une note entre 1 et 5. Ces questions concernaient : l'effet d'apparition du stimulus suscitant l'émotion (1= pas du tout soudain, 5= très brusque) ; le plaisir que procurait le stimulus (1= très désagréable, 5= très agréable) ; l'éventuel but atteint ou besoin satisfait grâce au stimulus (1= but ou besoin prévu, 5= but ou besoin résolu) ; l'état d'urgence de la réponse de l'auditeur (1= pas du tout urgent, 5= très urgent) ; le pouvoir d'une éventuelle action humaine adaptée pouvant modifier le résultat du stimulus (1= impuissant, 5= puissant) ; l'auto-responsabilité de l'intervenant (1= pas de responsabilité, 5= grande responsabilité) ; la responsabilité d'une autre personne par rapport à l'évènement (1= Aucune, 5= grande responsabilité) ; l'accord entre les normes des stimuli et celles des hauts parleurs (1= non respecté, 5= très en accord). Les résultats montrent que les auditeurs peuvent interpréter plusieurs facettes de l'émotion au niveau situationnel, telles que la nouveauté, le plaisir tiré du stimulus, l'avancement des objectifs, l'urgence de la réponse, le pouvoir de l'évaluateur, et la norme de comptabilité entre le stimulus et le haut-parleur. Les expressions vocales porteraient ainsi de nombreuses informations, y compris sur la représentation cognitive de la nature de la situation qui a suscité la réponse émotionnelle chez le locuteur (Scherer, K.R., 1988, cité par Laukka et Elfenbein, 2012).

Comme précédemment cité, les jeunes enfants parviennent à déceler les émotions de leurs interlocuteurs. Mais qu'en est-il des personnes âgées ? C'est sur ce point qu'ont travaillé Schmidt, J., Janse, E. et Scharenbord, O. (2016). Ces chercheuses ont étudié l'effet de l'âge et des différences de sensibilité auditive sur la perception des émotions dans le discours émotionnel. Pour ce faire, elles ont employé des énoncés émotionnels courts en y incluant des paramètres vocaux acoustiques tels que la hauteur tonale ou le NPS. Deux groupes de participants ont été sélectionnés. Le premier était composé de 20 adultes âgés de 22 ans en moyenne, et le second était composé de 20 personnes plus âgées, de 72 ans en moyenne. Les participants devaient écouter ces énoncés et noter l'intensité émotionnelle et la valence des énoncés perçus sur une échelle analogique allant de 1 à 5. Cette étude a pu démontrer que tous les participants se basent sur le NPS moyen pour l'arousal et sur la fréquence fondamentale moyenne pour la valence. Cependant, les différences de NPS sont moins utilisées par le groupe de participants âgés pour les évaluations d'arousal, alors que les différences de fréquences fondamentales ont une plus importante influence sur les notes de valence chez ces mêmes participants. La perception de l'intensité émotionnelle semble être résistante à une perte auditive légère à modérée.

F. Perception et production émotionnelle selon le genre

L'expression émotionnelle est légèrement différente entre les hommes et les femmes. En effet, les femmes ont tendance à être plus expressives (Brody L.R., Hall, J.A., 1993, cité par Chaplin, T.M., 2015). Les femmes sont donc plus expressives pour les émotions positives (LaFrance, M., Hecht, M.A., Levy Paluck, E.L., 2003) tandis qu'elles masquent plus les émotions négatives telles que la tristesse (Allen, J.G., Haccoun, D.M., 1976). En revanche, les hommes ont tendance à exprimer plus d'agressivité et de colère (Archer, J., 2004). On peut expliquer cela par le fait que les hommes ont tendance à garder leurs émotions pour eux en étant donc excités intérieurement tandis que les femmes laissent aller leurs émotions (Buck, R. ; 1977) tout en présentant le même taux d'excitabilité physiologique (Chaplin, T.M., Hong, K.A., Bergquist, K., Sinha, R., 2008).

Ces différences dans l'expression émotionnelle s'explique généralement par un modèle bio-psycho-social, alliant donc les trois modèles théoriques de départ. Le premier de ces modèles est un modèle biologique qui soutient que des différences innées dues à des facteurs biologiques créent ces différences comportementales entre les filles et les garçons. Ces différences biologiques entraîneraient une baisse de l'inhibition des comportements négatifs chez les garçons (Chaplin, T.M., 2015). Le second modèle est psychosocial et stipule que les enfants sont conditionnés à réagir d'une certaine manière en fonction d'un apprentissage cognitif, de l'expérience et de la socialisation qui se développent (Liben, L.S., Bigler, R.S., 2002). Dans ce modèle, la théorie du schéma de genre propose que les enfants développent des schémas cognitifs en fonction du genre, sur des observations de l'environnement (Martin, C.L., Halverson, C.F.Jr., 1981). Ensuite, les théories de l'apprentissage social soutiennent que les enfants sont conditionnés par un enseignement explicite, une modélisation par un encouragement implicite des comportements conformes à leur genre (Bandura, A., 1969, cité par Chaplin, T.M., 2015). Finalement, le troisième modèle est constructionniste. Les théoriciens de ce modèle pensent que les émotions ne sont pas contenues de manière brute par les individus mais qu'elles sont créées par les interactions avec l'environnement. Les enfants ont tendance à réagir émotionnellement aux attentes des personnes qui les entourent lors d'interactions sociales (Zeman, J., Garber, J., 1996).

Lorsque l'on combine ces trois théories, un modèle bio-psycho-social émerge en ce qui concerne l'expression émotionnelle. Ce modèle émet l'idée que les différences inter-genres dans l'expression émotionnelle provient d'une combinaison de facteurs tels que les différences biologiques innées, l'influence de l'environnement social, la socialisation et les attentes de la

société (Chaplin, T.M., 2015). Cette théorie aborde des règles expressives liées au genre présentes dans la majorité des cultures. Selon ce modèle, les filles présenteraient une expressivité émotionnelle plus importante pour les émotions positives tandis qu'elles intériorisent les émotions négatives. A côté de cela, les garçons présenteraient de manière plus importante les émotions extériorisées comme la colère, le dégoût ou le mépris, par un manque d'inhibition de ces comportements, tandis les émotions considérées comme négatives (tristesse, anxiété) sont intériorisées (Brody L.R., Hall, J.A., 1993, cité par Chaplin, T.M., 2015).

G. Perception émotionnelle et intelligence artificielle

La capacité à reconnaître automatiquement les émotions pourrait être exploitée dans les interactions vocales avec des outils informatiques tels que des robots. C'est ce qu'ont essayé de démontrer Devillers et Al.(2014). En effet, leur étude portant sur les descripteurs non verbaux, s'interrogeait sur les possibilités de communication émotionnelle entre des humains et des robots. La voix est porteuse de nombreux indices tels que l'état physique, mental ou affectif d'un participant. Cette étude portait sur des échantillons vocaux non-verbaux spontanés, à l'instar de nombreuses études précédentes, souvent réalisées à partir d'émotions jouées par des comédiens. La recherche des descripteurs non-verbaux les plus robustes est fondamentale dans le domaine des émotions et de la voix, cette étude a permis d'identifier un ensemble de paramètres suffisamment robuste pour offrir un gain de temps lors de l'analyse émotionnelle. Les tâches employées ont donc permis d'identifier, parmi un large panel d'ensemble de paramètres, une famille de paramètres vocaux qui permettrait plus facilement d'identifier l'émotion du locuteur. Ces paramètres sont une combinaison entre l'emploi de l'énergie par bandes spectrales calculée à partir des coefficients de « loudness », la moyenne et l'écart type de l'énergie par bandes de Bark calculés sur les parties voisées et non voisées également. Les bandes de Bark sont utilisées en psychoacoustique. Ce sont des bandes de fréquences critiques, comprises entre 0 et 15500 Hz, qui permettent de créer un axe des abscisses sur une courbe de sonie d'un son complexe. Les paramètres identifiés, c'est-à-dire l'énergie par bandes de Bark, le loudness se rapportant aux bandes de Bark, permettent de mieux reconnaître une classe d'émotions par rapport à une autre.

H. Physiologie et émotions

Rappelons que les émotions sont la traduction physiologique d'un état psychologique (Cacioppo, J.T., Berntson, G.G., Larsen, J.T., Poehlmann, K.M. et Ito, T.A. 2000, cités par Van Mersbergen, M. et Delany, M. 2014). Il semble, de ce fait, nécessaire d'employer des techniques qui quantifient directement la physiologie des plis vocaux. L'électroglottographie (EEG) offre cette possibilité, l'EEG reflète la durée de la phase de contact entre les plis vocaux sur un cycle glottal complet. L'électroglottographe permet d'obtenir les caractéristiques d'adduction laryngée en mesurant le changement fluctuant dans la résistance électrique du tissu laryngé présenté par l'oscillation du pli vocal (Baken, R.J., Orkioff, R.E. 2000, cité par Van Mersbergen, M. et Delany, M. 2014). On pourrait, par cette technique, essayer de déterminer si la vibration du pli vocal est influencée par les stimuli émotionnels.

C'est sur base de ces constatations que Van Mersbergen, M. et Delany, M. (2014) ont mené leur étude. Elles ont voulu vérifier l'efficacité de l'électroglottographie comme une mesure viable pour différencier les changements vocaux en réponse à des stimuli émotionnels. Cette étude a donc testé l'hypothèse selon laquelle le quotient de contact EEG serait plus important lors d'état émotionnel négatif, au contraire des émotions positives dont le quotient de contact serait plus faible. Pour parvenir à leur objectif, elles ont sélectionné un groupe de participants ne présentant pas de troubles vocaux ou psychologiques (dépression, troubles anxieux,...). Pour induire l'émotion chez ces participants, elles leur ont présenté 60 diapositives représentant des images du Système International d'Images Affectives (IAPS) (Lang, P.J., Bradley, M.M., Cuthbert, B.N. 2008, cité par Van Mersbergen, M. et Delany, M. 2014). Certaines images présentaient une valence positive, d'autres une valence négative, ou encore neutre. Les participants étaient installés dans une chaise longue avec une jauge respiratoire autour de la cage thoracique, des électrodes ECG sur la poitrine (Skintact, Leonard Lang USA, Inc. Inverness, FL). Les électrodes EGG (EG2-PCX2 ; Glottal Enterprises, Inc., Syracuse, NY) ont été placées sur les bords lamellés des cartilages thyroïdiens. Après visualisation d'une image IAPS, les participants devaient vocaliser leurs émotions sur la voyelle /u/ jusqu'à disparition de l'image (3 secondes). Après l'expérience, les participants devaient remplir trois questionnaires d'auto-évaluation (BORG CR-10, SAM-A et SAM-V). Au total, 192 images ont été présentées aux participants, sous forme de 6 blocs de 32 images. Les stimuli ont été présentés via le logiciel E-Prime 2 (version 2.0 Professionnel, Psychology Software Tools, Inc., Sharpsburg, PA.), les questionnaires d'auto-évaluation ont également été enregistrés et générés à partir de ce logiciel. Les résultats de cette étude montrent que le quotient de contact EEG peut

être une mesure viable pour détecter les changements de comportements vocaux en fonction des émotions. En effet, on relève un quotient de contact EEG plus élevé pour les conditions émotionnelles négatives par rapport aux conditions neutres ou positives. Il serait intéressant que les futures recherches dans ce domaine incluent des mesures physiologiques de la phonation supplémentaires telles que des mesures de l'engagement musculaire via une électromyographie, ou des mesures de la fonction aérodynamique avec une mesure de la pression du seuil de la phonation ainsi que des estimations de la résistance laryngée. Cet ensemble de mesures permettrait de déterminer de quelle manière l'émotion et le tempérament contribuent au développement des troubles de la voix.

Dans le même thème de recherches, Waaramaa, T. et Kankare, E. (2013) ont étudié les interactions entre le quotient de contact EEG, les fréquences des formants F1, F2, F3 et F4, et la durée, dans le but d'identifier les variations de couleurs émotionnels dans la voix. Elles ont voulu déterminer si le quotient de contact EEG ainsi que les valeurs de formants variaient en fonction de l'émotion et de la valence positive, négative ou neutre. Pour mener à terme cette investigation, elles ont recruté un groupe d'acteurs professionnels qui devaient produire les phonèmes /a/, /i/ et /y/, de manière soutenue, dans différents états émotionnels (joie, intérêt, surprise, neutralité, colère, peur, dégoût et tristesse). L'électroglottographie permet de mesurer l'impédance des tissus glottiques pendant la phonation. Cette impédance est plus faible lorsque la glotte est fermée car les tissus sont de meilleurs conducteurs de courant que l'air quand la glotte est ouverte. En conséquence, le courant diminue lorsque la glotte est ouverte. Les auteurs ont noté qu'il est plus fiable de réaliser une électroglottographie chez un participant masculin que chez un participant féminin, en raison de la structure du cartilage thyroïdien plus marquée. De plus, cette technique de mesure peut se révéler inefficace chez les personnes en excès pondéral en raison de l'épaisseur du tissu adipeux au niveau du cou. Pour donner suite à ces observations, les auteurs ont trouvé intéressant d'étudier la variation de fréquences des formants, ceci a été réalisé à l'aide du logiciel Praat Software® (Boersma, 2009). Les résultats de cette étude révèlent une corrélation négative entre les fréquences de formants F3 et F4 et le quotient de contact (CQEGG) ; plus F3 et F4 sont enrichis, plus le quotient de contact diminue. En effet, chez les participants féminins, le CQEGG peut varier de manière inversement proportionnelle avec les fréquences F3 et F4 lors des expressions émotionnelles, notons tout de même que cela peut être dû aux mouvements verticaux du larynx qui modifieraient les paramètres. Ensuite, les discriminateurs de valence émotionnelle les plus fiables semblent être effectivement le CQEGG et les formants 3 et 4. Le quotient de contact a tendance à varier simultanément et inversement avec les fréquences F3 et F4 pour des expressions émotionnelles

positives (Waaramaa, T., Kankare, E., 2013). Les auteurs rappellent également que d'après des études précédentes, les fréquences des formants F3 et F4 se sont révélées plus élevées et plus riches en harmoniques dans les expressions émotionnelles positives, que dans les expressions négatives (Waaramaa, T., Laukkanen, A-M., Airas, M. et Alki, P. 2010 ; Laukkanen, A-M., Vilkmán, E., Alku, P. et Oksanen, H. ; Waaramaa, T., Laukkanen, A-M., Alku, P. 2006, cités par Waaramaa , T. et Kankare, E. 2013).

PARTIE
PRATIQUE

V. Objectifs et hypothèses

A. Objectifs de ce mémoire

L'équipe canadienne, dirigée par Ingrid Verduyck et Boutheina Jemel, cherche à développer un ensemble de stimuli visuels permettant d'étudier les réponses émotionnelles vocales et subvocales (tensions musculaires laryngées) de participants humains. Cet ensemble de stimuli sera utilisé par la suite pour, tout d'abord, étudier les facteurs propres aux stimuli (intensité et valence émotionnelle,) qui influencent la propension d'un individu à réagir vocalement ou subvocalement à des stimuli émotionnels. Ensuite, ces stimuli seront utilisés pour étudier la manière dont ces facteurs influencent la qualité du geste vocal en termes de tension musculaire et de caractéristiques acoustiques. Et finalement, ils seront utilisés pour étudier les facteurs individuels qui influencent la propension d'un individu à réagir vocalement ou subvocalement à des stimuli émotionnels. Comprendre les processus cognitivo-émotionnels sous-jacents à l'expressivité vocale améliorera sans aucun doute les pratiques cliniques (évaluation et intervention) en psychologie et en logopédie. La population neurotypique désireuse d'améliorer ses compétences émotionnelles ainsi que les patients souffrant de déficit de l'expression vocale seront particulièrement concernés. De plus, une meilleure compréhension de la variabilité interindividuelle de l'expression émotionnelle vocale devrait améliorer les technologies développées dans le cadre de la reconnaissance et de la synthèse vocale. Ces techniques sont actuellement en plein essor, mais les performances actuelles des systèmes d'identification des émotions ne sont pas encore adéquats pour des applications de la vie réelle. Leur évolution retentira positivement sur les domaines de la psychologie et de la logopédie notamment dans l'aide au diagnostic, du monitoring à distance, ou encore des programmes d'entraînement à domicile.

Afin d'atteindre cet objectif, différents travaux, dont celui-ci, ont vu le jour. L'objectif de ce mémoire en particulier est de décrire les vocalisations émotionnelles en termes d'étendues fréquentielles et de les situer par rapport à l'étendue fréquentielle physiologique et conversationnelle du locuteur. Pour se faire, une analyse fréquentielle des différentes vocalisations sera réalisée. Les analyses statistiques visent à comparer les différentes étendues fréquentielles, relevées à travers l'ensemble des enregistrements, en fonction de l'intensité émotionnelle.

B. Hypothèses de ce mémoire

L'ensemble des données de la littérature sur le rapport entre la voix et les émotions ainsi que nos nombreux échanges d'idées nous ont permis d'émettre trois hypothèses.

Notre première hypothèse stipule que l'étendue émotionnelle effectuera une ascension plus importante et dépassera plus l'étendue fréquentielle conversationnelle pour les émotions à forte intensité émotionnelle que pour les émotions à faible intensité émotionnelle de manière régulière chez tous les participants et en ne tenant pas compte de la valence. En effet, d'après des études antérieures, des moyennes de fréquences plus élevées par rapport à la fréquence fondamentale ont été relevées en ce qui concerne la colère chaude ou froide, ainsi que pour la joie intense. De plus, les pentes d'excursion ascendantes ou descendantes de la f_0 ont été étudiées et il apparaît que les pentes pour les émotions à forte activation, comme la joie ou la colère, sont plus importantes, tandis que les pentes pour les émotions à plus faible activation comme la tristesse, la joie calme ou l'anxiété, sont également plus faibles (Morel, M., Bänziger, T., 2002).

Notre seconde hypothèse concerne les différences entre les conditions spontanées et imposées. En effet, nous pensons que les étendues fréquentielles des vocalisations spontanées seront plus petites que celles de vocalisations imposées. Les locuteurs auraient tendance à exagérer les vocalisations et à être plus expressif étant donné qu'ils sont amenés à vocaliser.

Ensuite, notre troisième hypothèse est que les étendues fréquentielles des vocalisations émotionnelles se situeront au-delà de l'étendue fréquentielle conversationnelle des locuteurs. En effet, les émotions permettent aux êtres humains de communiquer rapidement leurs différents besoins à leur interlocuteur, les changements physiologiques sont saillants et donc visibles de l'extérieur (Blair, R.J.R., 2003). Comme le citait Altrov (2013), les différentes émotions sont traduites par des variations de hauteur tonale en fonction de l'intensité émotionnelle, plus grave pour la tristesse mais plus aigüe pour le bonheur et la colère.

VI. Méthodologie

A. Population

1. Population de l'étude

La population totale de l'étude « Voix et émotions » est composée de 330 participants. En Belgique et au Canada, 100 participants neurotypiques âgés de minimum 18 ans seront recrutés. S'ajouteront à cela, 50 participants belges atteints d'une dysphonie musculo-squelettique et phono-traumatique, ainsi que 50 participants souffrant de dépression. Finalement, 30 participants présentant un trouble neurologique seront inclus au Québec.

2. Population de ce mémoire

En ce qui concerne ma recherche, 20 participants belges ont été recrutés dans la population générale afin de leur administrer les tâches. Parmi ces 20 participants, 10 hommes et 10 femmes ont participé à l'expérience. S'ajoutent à cela les données des participants recrutés cette année au Canada. Les participants recrutés sont des individus neurotypiques. Les conditions d'inclusion sont d'être majeur (18 ans) afin que l'organe vocal soit mature, et de ne pas souffrir ou avoir souffert d'un trouble psychologique.

B. Matériel

1. Questionnaires psychologiques

Les questionnaires employés lors de cette expérience permettent d'évaluer la personnalité des différents participants. Dans le cadre de mon mémoire, tous les questionnaires ont été administrés mais seul un questionnaire m'a permis d'écarter un risque de dépression chez les participants. Les autres données seront employées dans le cadre des différents travaux réalisés à l'Université de Montréal.

*a) **Questionnaire de Sensibilité à la Punition et de Sensibilité à la Récompense (QSPSR)***

Cet auto-questionnaire (Lardi, C., Billieux, J., d'Acremont, M., et Van der Linden, M., 2008) se rapporte au facteur de personnalité. Il permet d'évaluer les variations individuelles des

sensibilités émotionnelles et les tendances comportementales d'approche et d'évitement des participants. Ce questionnaire évalue la manière de penser ou de se comporter des participants.

b) *L'échelle de Régulation Émotionnelle (DERS-F)*

Cette échelle (Dan-Glauser, E.S. et Scherer, K.R., 2013) se rapporte au facteur de régulation émotionnelle. Elle se compose de 6 facteurs tels que les difficultés à s'engager dans un comportement dirigé vers un but ou les difficultés de contrôle des impulsions. Elle permet donc d'évaluer à quel point les participants sont attentifs à leurs émotions dans la vie courante. La DERS-F est la traduction française de la *Difficulties in Emotion Regulation Scale (DERS)* (Gratz, K.L. et Roemer, L., 2004). D'ailleurs, la version française a démontré de très bonnes propriétés psychométriques (Côté, G., Gosselin, P., Dagenais, I., 2013)

c) *L'Échelle d'Anxiété Sociale de Liebowitz (ESCL)*

Cette échelle se rapporte aux facteurs dépressifs et anxieux. Elle est composée de 24 énoncés évaluant l'anxiété et l'évitement que ressent une personne quand elle est exposée à une situation anxiogène ou de performances. L'échelle employée est la traduction française de la *Liebowitz Social Anxiety Scale (LSAS)* (Yao, S.N., Note ; I., Fanget, F., Albuissou, E., Bouvard, M., Jalenques, I., Cottraux, J., 1999).

d) *L'Inventaire de Dépression de Beck Seconde Edition (IDB-II)*

Cet auto-questionnaire (Beck, M.D., Ward, Ch., Mendelson, M., Mock, J., Erbaugh, J., 1996) de 21 items se rapporte au facteur affectifs dépressifs et permet de mesurer la sévérité et présence de traits dépressifs chez les participants. Seul ce questionnaire me permettra de confirmer l'inclusion du participant.

e) *La Positive and Negative Affect Scale (PANAS)*

L'échelle PANAS (Watson, D., Clark, L.A., Tellegen, A., 1988) est une auto-évaluation des affects positifs et négatifs qui permet d'évaluer l'état émotionnel des personnes. Cette échelle contient 20 mots représentant des émotions ou sentiments, répartis sur une échelle d'affectivité négative et une échelle d'affectivité positive.

2. Matériel pour l'enregistrement vocal

a) Le microphone et le casque

Afin de réaliser des enregistrements vocaux de qualité, il est nécessaire de respecter certaines conditions. Tout d'abord, les participants sont équipés de microphones, réalisant la conversion de la pression acoustique en tension électrique afin d'être traitée par un logiciel. En effet, lorsque que nous parlons, la pression acoustique de l'air subit des variations, celles-ci sont captées par le microphone et transmises aux bornes du câble électrique (Menin-Sicard, A. et Sicard, E., 2016). Il est conseillé d'employer un microphone sur pied posé devant le participant ou le patient afin d'obtenir la meilleure qualité vocale possible mais dans ce cas, la distance entre la bouche du participant et le microphone risque d'être instable.

Pour pallier cela, il est possible d'utiliser une règle, une latte, maintenant une bonne distance et permettant de réaliser des analyses objectives des données. Afin de faciliter la passation des expériences de cette étude, un microphone serre-tête Shure® (Shure® WH20 XLR) a été employé. Il permet de maintenir une distance de 5 centimètres entre la bouche du participant et le microphone.

De plus, l'avantage des microphones externes, en comparaison aux microphones intégrés des ordinateurs, est qu'ils sont unidirectionnels et permettent de capter uniquement la voix des participants.

Les participants étaient également équipés d'un casque audio Bose® (Bose® Soundlink® around-ear wireless headphones II), celui-ci n'était pas anti-bruit afin que les participants puissent s'entendre lors de la répétition des vocalisations.

b) La carte son

Une carte son USB externe FocusRite iTrack Solo Lightning a été employée afin de transmettre au logiciel d'analyse vocale le son le plus fidèle possible à la réalité. Dans un même temps, celle-ci m'a permis de surveiller et de gérer les risques de saturation lors de l'enregistrement. Cela dit, cette carte son était nécessaire car le microphone employé comporte une sortie XLR. Il est judicieux d'employer une carte son car les cartes son internes des ordinateurs portables, lorsqu'elles sont présentes, sont parfois bruitées ou de mauvaise qualité. De plus, une connexion USB est plus solide et durable qu'une connexion par câble Jack qui peut s'abîmer avec le temps.

c) *Les logiciels d'enregistrement et d'analyse vocale*

Afin de relever les données acoustiques des différents participants de cette étude, le logiciel Audacity ® (version 2.2.1) a été employé. Lors des enregistrements, étant donné qu'un microphone serre-tête externe a été utilisé, la fonction mono-sound a été sélectionnée ainsi qu'une fréquence d'échantillonnage à 48000 Hz. En effet, le choix de la fréquence d'échantillonnage, c'est-à-dire la fréquence sélectionnée lors de l'enregistrement, entraîne des modifications de qualité du son et de la taille du fichier sur l'ordinateur. Dans le cas de cette étude, la fréquence employée a été de 48000 Hz car elle permet d'obtenir une excellente qualité sonore et est généralement employée par les systèmes d'enregistrement professionnels.

Ensuite, le logiciel Praat® (Boersma, 2009) a été utilisé pour analyser les différentes productions. Praat® (Boersma, 2009) est un logiciel librement téléchargeable qui offre des analyses multiparamétriques et objectives de la voix et de la parole. L'avantage principal de ce logiciel est d'être gratuit et facilement manipulable par un utilisateur averti. En revanche, il nécessite la surveillance externe de l'expérimentateur afin de ne pas commettre d'impair, tel qu'une mesure fréquentielle absurde.

3. Matériel pour les tâches de productions vocales émotionnelles

L'objectif de la recherche entamée à l'Université de Montréal est de créer un corpus de stimuli visuels permettant d'étudier les réponses émotionnelles vocales et subvocales de participants humains. Pour atteindre celui-ci, cette étude a employé des images servant de stimuli émotionnels, sélectionnées dans l'« *International Affective Picture System* » (IAPS; Lang, Bradely et Cuthbert, 2008), sur base des scores de valence et d'intensité émotionnelles proposés par les études de validation de l'IAPS. Cette base de données permet aux chercheurs d'utiliser des stimuli émotionnels normés permettant de comparer et de répliquer les recherches sur les émotions et l'attention.

Deux corpus de 60 images ont été utilisés, un pour les tâches de vocalisations spontanées et l'autre pour les tâches de vocalisations imposées. Chaque corpus contenait 20 images à valence positive, 20 images à valence négative, et 20 images à valence neutre. Pour contrôler l'intensité effective des images perçue par les participants, il leur a été demandé de porter un jugement sur l'intensité émotionnelle sur base d'une échelle analogique allant de « image provoquant une émotion de très forte intensité » à « image provoquant une émotion de très faible intensité ». Ces jugements ont été effectués après la complétion des tâches vocales.

C. Lieu des enregistrements

Les tâches de vocalisations ont été effectuées dans une cabine insonorisée dans un cabinet d'oto-rhino-laryngologie situé à Mons. La présentation des stimuli visuels et la récolte des données acoustiques ont été réalisées par le logiciel Open Sesame, (Mathôt, S., Schreij, D., & Theeuwes, J., 2012) appareillé d'un microphone monté sur un casque fixé sur la tête du participant. Ceci nous a permis de garder une distance égale (=5cm) à la bouche du participant pendant toute l'expérience. Le participant était assis dans la cabine insonorisée avec une lumière tamisée pour garder l'attention fixée sur les images. L'enregistrement vocal était synchronisé avec le temps d'apparition de l'image à l'écran. Il n'y a pas eu d'observation directe de l'expérimentateur afin d'éliminer un facteur de gêne.

VII. Procédure générale

A. Évaluation par questionnaires à domicile

Lors du recrutement des participants, un **formulaire d'informations** (annexe A.1) détaillant le but de la recherche et les différentes étapes leur a été fourni. Dès que les participants ont marqué leur accord pour participer à cette étude, différents documents leur ont été envoyés par courriel. Ces documents nous ont permis en premier lieu de vérifier qu'ils pouvaient intégrer le groupe expérimental :

- **Questionnaire anamnestique** (annexe A.2)
- **Formulaire de consentement éclairé** (annexe A.3)
- **Formulaire d'engagement des chercheurs à l'égard des participants à leur recherche** (annexe A.4)
- **Le Questionnaire de Sensibilité à la Punition et de Sensibilité à la Récompense (QSPSR)** (annexe A.5)
- **L'échelle de Régulation Émotionnelle (DERS-F)** (annexe A.6)
- **L'Échelle d'Anxiété Sociale de Liebowitz (ESCL)** (annexe A.7)

Ces données ont été encodées et rassemblées dans différents tableaux Excel et transmises aux étudiants de l'Université de Montréal via le drive internet commun.

B. Administratif et questionnaires psychologiques

1. Accueil

Le jour de l'expérience, les participants ont été accueillis dans le cabinet d'oto-rhino-laryngologie. Celui-ci étant inoccupé lors des expériences, j'étais seule avec les participants. Ils ont été accueillis et j'ai commencé par leur rappeler l'objectif de cette expérience qui est de mettre au clair la relation existante entre les émotions et la voix.

2. Administratif

Ensuite, j'ai lu l'entièreté du **formulaire de consentement éclairé**, reçu préalablement par courriel, avec le participant. Si le participant était toujours en accord avec celui-ci, nous procédions à la signature et poursuivions l'expérience. J'ai fourni au participant une copie

signée du **formulaire d'engagement des chercheurs à l'égard des participants à leur recherche.**

3. Inventaire de Dépression de Beck (BECK-II)

Le participant est resté seul quelques minutes afin de remplir cet auto-questionnaire. J'ai contrôlé le score obtenu afin d'écarter un potentiel risque de dépression. Si le participant était sain, nous pouvions continuer.

4. La Positive and Negative Affect Scale (PANAS)

En même temps que la BECK II, le participant devait remplir le questionnaire PANAS, dont les données seront employées par les étudiants de l'université de Montréal.

C. Tâches vocales

1. Tâche d'étendue fréquentielle

Le but de cette tâche était d'évaluer **l'étendue fréquentielle physiologique** et **conversationnelle** en termes de fréquences du participant. Pour **l'étendue fréquentielle physiologique**, il lui a été demandé de réaliser **trois sirènes ascendantes** et **trois sirènes descendantes**. Le participant était équipé du microphone et l'enregistrement s'est fait via Audacity ® réglé sur une fréquence de 48000 Hz afin d'obtenir un son d'excellente qualité. Tous les sons ont été enregistrés séparément. Avant chaque enregistrement, j'ai donné un exemple de sirène que le participant a reproduit afin de m'assurer qu'il avait compris l'exercice et qu'il n'y avait pas de confusion entre hauteur tonale et NPS. Lors de l'enregistrement, j'ai guidé le patient avec un mouvement de la main.

En ce qui concerne **l'étendue fréquentielle conversationnelle**, il a été demandé au participant de me raconter une **recette de cuisine** au choix pendant environ 1 minute. L'enregistrement de la voix conversationnelle est important lors d'un bilan vocal ou d'une passation d'expérience car cela permet d'enregistrer la personne en situation naturelle et spontanée.

2. Production des sons pour la tâche de justesse vocale

J'ai demandé au participant de réaliser **trois /a/** de hauteurs tonales différentes, pour une durée de deux secondes. Le premier /a/ était situé dans les hautes fréquences, le second dans la fréquence moyenne confortable pour le participant et le troisième était un /a/ de basses fréquences.

Ensuite un premier essai a été réalisé. Lors de celui-ci, je suis restée encore avec le participant dans la cabine. Un **PowerPoint de trois images** (un scorpion, un bébé et une fleur) a été diffusé sur l'ordinateur et il a été demandé au participant de tenter de me faire deviner l'émotion de l'image qu'il regardait, sans employer de mots, juste des sons. Pour cette partie, le PowerPoint a servi de support à mes explications et je pouvais donner un exemple si je voyais que le participant n'avait pas compris.

Finalement, un second **PowerPoint de 10 images** a été diffusé. Dans ce cas-ci, les vocalisations du participant ont été enregistrées et le participant était seul pendant la diffusion sur l'écran. Le participant a dû réaliser le même exercice que lors de l'essai précédent.

3. Préparation pour la suite de l'expérience

Dès que le participant a eu terminé le PowerPoint de 10 images, je l'ai invité à patienter quelques minutes le temps que je récupère les sons pertinents. Les **trois /a/** ont servi pour la **tâche de justesse vocale**, tandis que j'ai choisi **trois sons émotionnels** parmi les dix enregistrés via le dernier PowerPoint. Ces **trois sons** devaient véhiculer de manière explicite et différenciée des émotions distinctes telles que la joie, le dégoût, la peur ou l'envie. De plus, ils devaient être facilement reproduits par le participant.

Parmi ces **trois sons émotionnels**, **deux sons** ont été sélectionnés, l'un typiquement positif et l'autre négatif, afin d'être inséré dans le set 4 de l'expérience (set masqué) sur Open Sesame.

Pour préparer la **tâche de justesse vocale**, j'ai inséré dans le logiciel Open Sesame les **trois /a/** récoltés, ainsi que les **trois sons émotionnels choisis**.

4. Tâche de justesse vocale

Lors de cette tâche, le participant était seul, portait un casque audio et le microphone. Sur le logiciel Open Sesame, les **six sons** sélectionnés précédemment ont été diffusés pour le

participant plusieurs fois. Il lui a été demandé de les répéter de manière identique en termes de hauteur et de durée. Le logiciel Open Sesame a enregistré les sons de manière automatique.

D. Expérience

Lorsque les différents enregistrements précédents ont été terminés, nous sommes passés à l'expérience à proprement parler. Celle-ci était composée de quatre sets plaçant le participant dans des conditions différentes. Les images de l'IAPS ont été réparties aléatoirement via Excel pour chaque participant. Le participant devait réagir vocalement à chacune d'entre-elles, en employant la première expression vocale non-verbale qui lui venait à l'esprit. Les participants ont eu pour but de faire deviner l'image qu'ils étaient en train de visionner, sur la seule base de leur vocalisation.

Chaque image était précédée d'un écran noir avec une croix de fixation, afin de permettre aux participants de décider du moment d'apparition de l'image qu'ils activait directement de l'ordinateur. Ceci avait pour but de ne pas inonder les participants de nombreuses émotions qui pouvaient créer un phénomène de contamination, c'est-à-dire que les participants auraient pu être influencés par un stimulus à intensité émotionnelle forte ou valence négative, par exemple, ce qui aurait faussé la réponse vocale. Les participants auraient pu également ne plus réagir de manière adaptée tant à la valence qu'à l'intensité du stimuli, nous aurions fait face à un phénomène d'habituation.

Quand l'image est apparue, le participant a dû commencer sa vocalisation. Le temps d'apparition du stimulus était de 3 secondes, car c'est le temps maximum attendu lors d'une vocalisation affective non verbale. Le participant a dû cliquer sur la barre espace pour passer à l'image suivante.

Après chaque image visionnée, le participant a dû coter, via une échelle analogique, la valence et l'intensité de l'émotion ressentie afin de contrôler la concordance avec les cotations de l'IAPS. Avant le début de l'expérience, j'ai montré une capture d'écran des échelles de cotation pour expliquer correctement au participant ce qu'il devait faire.

Au début de chaque set, 5 images tests ont été projetées, afin de vérifier la compréhension de l'exercice, je suis restée avec les participants pendant ces phases tests qui n'ont pas été enregistrées.

1. Set 1 : spontané

Dans la condition spontanée, nous avons cherché à mesurer la propension d'un individu à réagir vocalement à une émotion, à identifier les stimuli émotionnels produisant le plus de réponses vocales, et à étudier l'impact de la valence et de l'intensité émotionnelles perçues des stimuli sur les caractéristiques acoustiques et temporelles des vocalisations. Dans ce cas-ci, les participants n'ont pas été obligés de réagir vocalement, le but étant de réagir spontanément, donc si rien ne leur venait à l'esprit, ils pouvaient se taire et passer à l'image suivante. Les participants ont visualisé un set de 45 images différentes.

2. Set 2 : imposé

Dans la condition imposée, nous avons cherché à prendre des mesures acoustiques et temporelles des vocalisations en fonction de la valence et de l'intensité des émotions, et à étudier les modulations de la qualité vocale induites par les variations de valence et d'intensité des stimuli émotionnels. Dans cette condition, le participant a été obligé de réagir vocalement aux 45 images qui ont défilé, en essayant de transmettre au mieux l'émotion ressentie.

3. Set 3 : simulé/neutre

Lors du troisième set, les participants ont entendu un son de fréquence neutre, ils ont dû reproduire la fréquence entendue sur un /a/ le plus stable et le plus neutre possible face aux 30 images visionnées. Nous avons cherché à identifier les paramètres instables qui traduisent l'état émotionnel de la personne.

4. Set 4 : masqué

Dans ce dernier set, les participants n'ont pas pu réagir spontanément aux images. Les sons positifs et négatifs sélectionnés lors de la production de sons pour la tâche de justesse vocale ont été insérés dans le logiciel. La vocalisation positive a été associée à un rond vert et la vocalisation négative à un rond rouge. En premier lieu, le participant s'est entraîné à produire la réponse adéquate en fonction de la couleur du rond qui apparaissait. Il a dû essayer de reproduire la même vocalisation que celle enregistrée. Ensuite, après les 5 images d'essai, le participant a vocalisé sur les 60 images du set. En effet, il y avait, dans ce set, 30 images différentes qui étaient présentées deux fois, une fois avec le rond vert et l'autre avec le rond

rouge. Nous avons cherché à étudier la stabilité de la vocalisation en fonction de la valence de l'image à laquelle le participant était confronté.

E. Durée de l'expérience

Évaluation par questionnaires à domicile	50 minutes
Anamnèse	15 minutes
QSPSR	15 minutes
DERS-F	10 minutes
ESCL	10 minutes
Administratif et questionnaires psychologiques	25 minutes
Accueil, explications et signature des documents	5 minutes
BECK - II	10 minutes
Panas	10 minutes
Tâches vocales	30 minutes
Étendue fréquentielle	10 minutes
Production des sons pour la tâche de justesse vocale	5 minutes
Préparation pour la suite de l'expérience	5 minutes
Tâche de justesse vocale	10 minutes
Expérience	60 minutes
Set 1	15 minutes
Set 2	15 minutes
Set 3	15 minutes
Set 4	15 minutes
TOTAL	2h45

F. Variables recueillies et analysées

Ce mémoire porte sur les variations de la fréquence fondamentale en fonction de l'étendue fréquentielle des participants, au travers des différentes émotions ressenties. Toutes les étapes de l'étude ont été appliquées mais je n'ai analysé que certaines tâches pertinentes par rapport à ma recherche.

Pour commencer, j'ai analysé toutes les étendues vocales des participants afin d'en extraire les valeurs de fréquences maximales et minimales dans le contexte physiologique et conversationnelle.

Ensuite, nous avons commencé par passer un script Praat®, réalisé à l'université de Montréal, à travers l'ensemble des sets afin d'extraire uniquement la vocalisation de l'enregistrement. Ensuite, un autre script Praat® nous a permis de recueillir les données fréquentielles pour l'ensemble des sets. Grâce à cela, nous avons pu récolter les données d'étendues fréquentielles émotionnelles pour chaque participant.

Je me suis focalisée ensuite uniquement sur les sets 1 et 2 afin d'étudier les variations de fréquences des participants en fonction des émotions et de leurs étendues fréquentielles.

VIII. Résultats

L'objectif principal de cette étude est d'investiguer l'impact de l'intensité de stimuli visuels émotionnels sur la fréquence fondamentale en regard de l'étendue fréquentielle de nos participants. D'un point de vue statistique, toutes les analyses ont été réalisées sur le logiciel SPSS. Nous commencerons par exposer des statistiques descriptives visant à décrire notre échantillon mais également les différentes données recueillies.

Nous avons effectué des statistiques inférentielles à l'aide d'analyses de variance avec schéma en mesures répétées (Anova à mesures répétées sur la moyenne). Ce test statistique nous a permis de tester plus de deux variables dépendantes en même temps. Nous avons contrôlé l'homoscédasticité des variances à l'aide du test de Levene qui permet d'accepter une hypothèse d'égalité de variance dans le cas où les probabilités de dépassement sont fortes par une non-significativité du résultat ($p > 0,05$). Nous avons réalisé un test de sphéricité de Mauchly afin de vérifier les conditions de sphéricité des variables. Si le résultat était non-significatif ($p > 0,05$), la sphéricité était satisfaite et nous pouvions employer les données de significativité du Lambda de Wilks sans ajustement du degré de liberté. Malgré les différentes valeurs rapportées par le logiciel (Lambda de Wilks, trace de Pillai, trace de Hotelling et la plus grande racine de Roy), le Lambda de Wilks est la plus couramment employée. En revanche, si le résultat du test de sphéricité de Mauchly était significatif ($p < 0,05$), nous devons appliquer la correction de Greenhouse-Geisser qui ajuste le degré de liberté. Ces valeurs correspondent aux résultats des effets intra-participants de l'Anova. Si elles étaient significatives, nous avons affiné les résultats à l'aide d'un test post-hoc avec une correction de Bonferroni qui permet de détailler la significativité des différents niveaux de la variable dépendante. Nous avons obtenu également les résultats de tests inter-participants qui nous renseignaient sur l'implication de certaines variables sur les résultats.

A. Description de la population

Nous remarquons (tableau 1) une différence de proportion d'hommes et de femmes entre les deux pays. En effet, en Belgique, dix participants de chaque genre ont été sélectionnés tandis qu'au Québec, il y a eu 16 femmes contre 4 hommes qui ont pris part à l'expérience. Les hommes sont, en moyenne, plus âgés que les femmes. En revanche, l'écart-type étant très important, les âges des hommes sont moins homogènes que ceux des femmes. Les moyennes d'âges entre les deux pays sont relativement similaires mais les écarts-types traduisent une hétérogénéité de la population, surtout au Québec.

Tableau 1 : description de l'effectif et de l'âge de la population de l'étude

	Effectif	Moyenne âge (ET)	Minimum âge	Maximum âge	
Femmes (N=26)	Québec	16	25,46 (8,33)	19	57
	Belgique	10			
Hommes (N=14)	Québec	4	32,64 (15,62)	22	68
	Belgique	10			
Québec	20	28,60 (13,56)	19	68	
Belgique	20	27,35 (9,92)	22	57	

B. Étendues fréquentielles des participants en demi-tons

1. Étendues fréquentielles et physiologiques

Lors de l'observation des fréquences conversationnelles relevées (annexe B.1), nous avons remarqué que les fréquences minimales sont très similaires pour l'ensemble des groupes. Cela nous a incité à employer les données du quantile 25 comme minimum pour nos analyses suivantes. Les fréquences minimales et maximales en Hertz ont également été relevées pour l'étendue physiologique (annexe B.2). Les fréquences en Hertz ont été converties en demi-tons selon la formule : $\frac{Hz2}{Hz1} (\text{Log}2) \times 12$.

Nous observons (tableau 2) une grande hétérogénéité des étendues lorsque nous employons les valeurs du quantile 25 et de fréquence maximale. En revanche, en utilisant comme limite le quantile 75, les étendues sont plus homogènes. De plus, nous relevons un ambitus de 23,28 demi-tons pour les hommes québécois dans le premier cas (quantile 25- F0 maximale), ce qui équivaut à presque 2 octaves, c'est une valeur aberrante pour une étendue fréquentielle conversationnelle. En effet, une étendue de 2 octaves est généralement récoltée lors d'une étendue fréquentielle absolue. Ensuite, concernant les étendues physiologiques, les participants belges, peu importe le genre, présentent de plus grandes étendues fréquentielles en

demi-tons. On relève également que l'écart-type est très important pour les hommes belges, ce qui indique une grande hétérogénéité des performances.

Lorsque l'on observe les étendues conversationnelles maximales et les étendues physiologiques (tableau 2), les différences ne sont pas celles attendues. En effet, nous nous attendions à de plus grands ambitus physiologiques pour tous les participants sauf que les hommes québécois présentent un ambitus conversationnel plus grand que le physiologique et surtout plus grand que tous les autres groupes (figure 7). En dehors de cela, nous pouvons voir que les belges ont effectivement des étendues fréquentielles physiologiques plus importantes que les québécois.

Tableau 2 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale, et entre les moyennes des quantiles 25 et 75 pour le conversationnel et entre la fréquence minimale et maximale pour le physiologique.

		Moyenne F0 Qt 25 – F0 max - conversationnel (ET)	Moyenne F0 Qt 25 – F0 Qt 75 - conversationnel (ET)	Moyenne des étendues physiologiques (ET)
F	Qc	13,38 (2,48)	2,83 (0,97)	19,32 (7,00)
	Be	15,48 (2,16)	3,43 (0,84)	24,06 (5,97)
H	Qc	23,28 (4,48)	2,35 (0,49)	20,34 (6,77)
	Be	14,13 (3,43)	3,48 (1,08)	24,23 (11,64)

Légende : F= femmes ; H= hommes ; Qc= Québec ; Be= Belgique ; ET= écart-type

Moyenne F0 Qt 25 – F0 max = Moyenne des étendues conversationnelles en demi-tons mesurées entre les moyennes des fréquences des quantiles 25 et de la fréquence maximale.

Moyenne F0 Qt 25 – F0 Qt 75 = Moyenne des étendues conversationnelles en demi-tons mesurées entre les moyennes des fréquences des quantiles 25 et des quantiles 75.

Moyenne des étendues physiologiques= Moyenne des étendues physiologiques en demi-tons mesurées entre les fréquences minimales et maximales relevées.

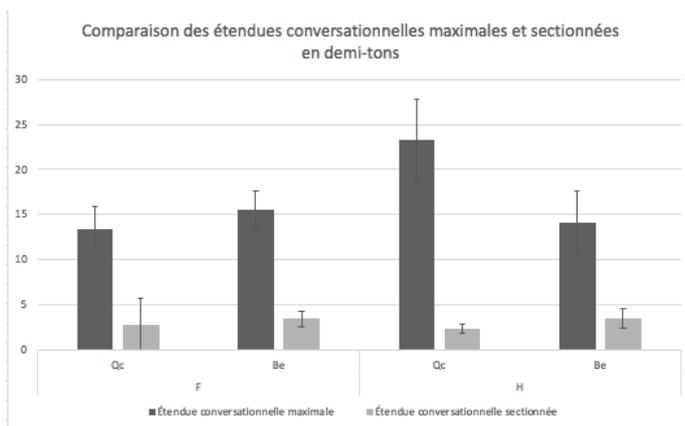
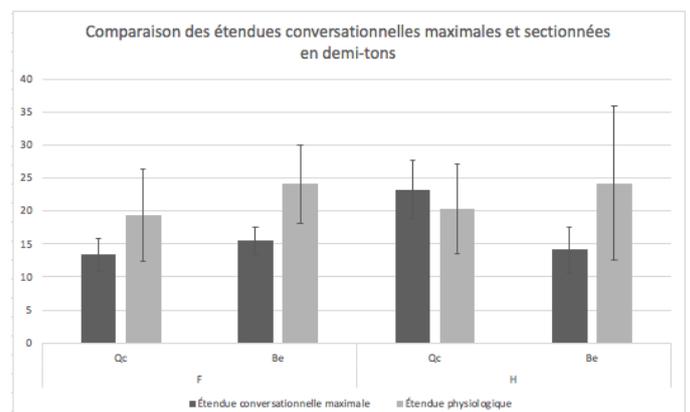


Figure 7. Comparaison des étendues fréquentielles conversationnelles et physiologiques

Figure 6. Comparaison étendues conversationnelles en demi-tons entre le quantile 25 et la F0 maximale et entre les quantiles 25 et 75



C. Vocalisations émotionnelles sur les sets 1 et set 2

1. Nombre de vocalisations pour les sets 1 et 2

Grâce à ces données (tableau 3), on peut observer que spontanément, les participants n'ont pas toujours vocalisé aux stimuli présentés. Nous pouvons également constater que nous avons recueilli plus de vocalisations en Belgique qu'au Québec. Ensuite, comme prévu, nous avons récolté plus de vocalisations pour le set 2 étant donné que celui-ci imposait de vocaliser. Mais les femmes ont plus répondu vocalement au set 1 qu'au set 2 tandis que les hommes ont réalisé plus de vocalisations au set 2 qu'au set 1.

Tableau 3 : Nombre de vocalisations émotionnelles pour les sets 1 et 2 par pays et par genre

		Set 1	Set 2	Total
Québec	F (N=16)	522	715	1237
	H (N=4)	107	167	274
Belgique	F (N=10)	405	403	808
	H (N=10)	390	440	830
Total		1424	1725	3149

2. Comparaison des étendues fréquentielles émotionnelles entre les sets 1 et 2

Lors de l'analyse des deux premiers sets de l'expérience, nous avons relevé les F0 moyennes, minimales, maximales, des quantiles 25 et 75 des vocalisations (annexe B.3).

De manière descriptive, nous avons l'impression que l'étendue fréquentielle est plus importante pour le set spontané (1) par rapport au set imposé (2) (tableau 4).

Tableau 4 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale, et entre les moyennes des quantiles 25 et 75 pour les vocalisations émotionnelles en demi-tons pour les sets 1 et 2

			Moyenne F0 Qt 25 – F0 max (ET)	Moyenne F0 Qt 25 – F0 Qt 75 (ET)
Set 1	F	Qc	8,34 (7,39)	3,84 (4,39)
		Be	6,94 (6,14)	3,67 (3,41)
	H	Qc	8,32 (6,88)	4,01 (3,92)
		Be	5,97 (4,78)	3,13 (2,87)
Set 2	F	Qc	7,74 (6,78)	3,63 (3,92)
		Be	7,13 (5,99)	3,77 (3,19)
	H	Qc	6,22 (5,24)	3,01 (2,72)
		Be	5,80 (4,80)	3,04 (2,88)

Légende : F= femmes ; H= hommes ; Qc= Québec ; Be= Belgique ; ET= écart-type ; set 1 = spontané ; set 2= imposé.
Moyenne F0 Qt 25 – F0 max = Moyenne des étendues émotionnelles en demi-tons mesurée entre les moyennes des fréquences des quantiles 25 et de la fréquence maximale des échantillons émotionnels.

Moyenne F0 Qt 25 – F0 Qt 75 = Moyenne des étendues émotionnelles en demi-tons mesurée entre les moyennes des fréquences des quantiles 25 et des quantiles 75 des échantillons émotionnels.

3. Comparaison des étendues fréquentielles conversationnelles et émotionnelles

Cette comparaison, basée sur les étendues fréquentielles conversationnelles (tableau 2) et émotionnelles (tableau 4) en demi-tons, nous montre que lorsque l'on emploie les fréquences maximales comme limite, aucune vocalisation émotionnelle ne dépasse l'étendue fréquentielle conversationnelle moyenne. En revanche, lorsque l'on emploie le quantile 75 comme limite, la plupart des vocalisations émotionnelles dépassent cet ambitus (figure 8).

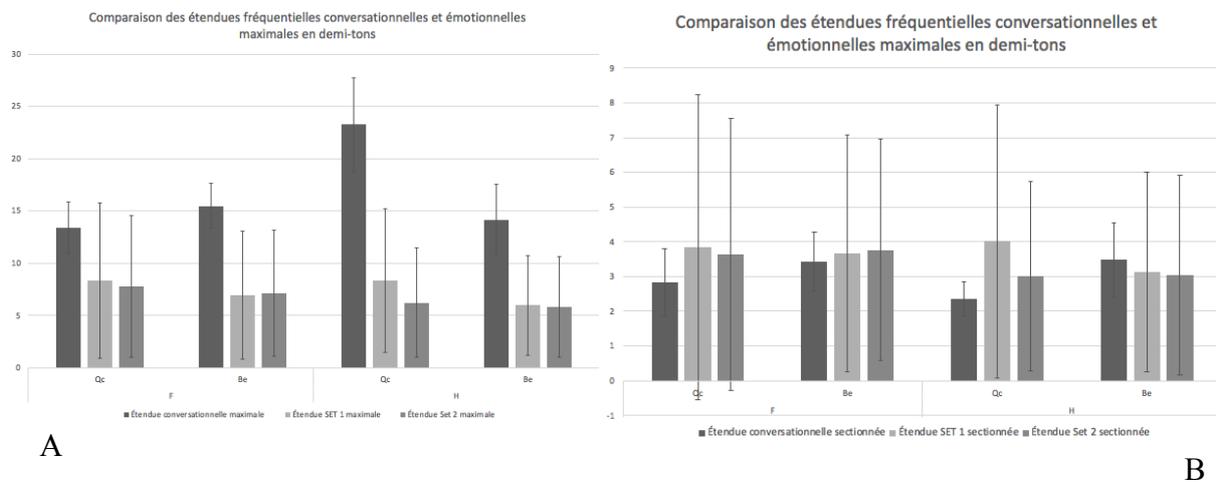


Figure 8. Comparaison des étendues fréquentielles conversationnelles et émotionnelles en demi-tons : A= moyennes entre les fréquences des quantiles 25 et des fréquences maximales ; B= moyennes entre les fréquences des quantiles 25 et 75

D. Vocalisations émotionnelles des sets 1 et 2 selon le niveau d'intensité émotionnelle estimé

1. Valeurs d'intensité et de valence émotionnelles relevées dans le set 2

Nous avons transformé les données d'intensité émotionnelle en rangs pour chacun des participants. À partir de ces rangs, nous avons réparti les données en 3 catégories d'intensité émotionnelle (faible, moyenne et élevée). Comme les données d'intensité émotionnelle individuelles sont manquantes pour le set 1, à cause d'une erreur de manipulation, cela n'a pu être fait que pour le set 2. Les moyennes des scores d'intensité et de valence ont été calculées pour chaque catégorie d'intensité (annexe B.4). Sur base de ces données, on peut observer (figure 9) que plus la catégorie d'intensité émotionnelle est élevée, plus le score moyen d'intensité émotionnelle cotée par les participants diminue. Nous voyons que les scores moyens de valence sont les plus élevés pour la catégorie d'intensité émotionnelle moyenne et les plus

bas pour la catégorie d'intensité émotionnelle élevée. Comme les valeurs faibles de valence correspondent à des stimuli négatifs et les valeurs élevées à des stimuli positifs, cela montre que le niveau d'intensité moyen contient des images en moyenne plus positives que les niveaux élevés et faibles et que le niveau d'intensité élevé contient des images en moyenne plus négatives que les niveaux faibles et moyens.

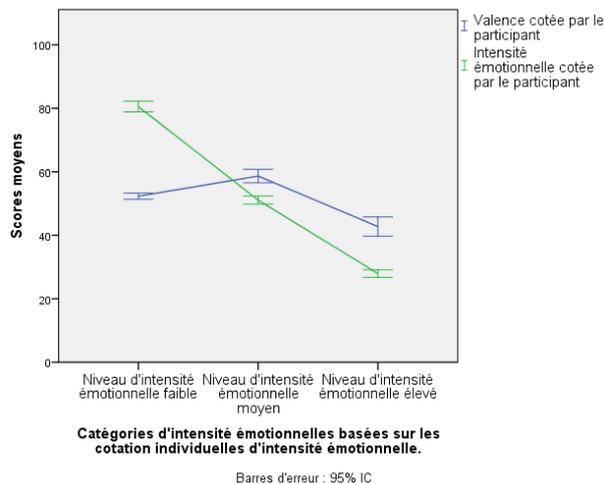


Figure 9. Catégories d'intensité émotionnelle basées sur les cotations individuelles d'intensité émotionnelle

2. Étendues fréquentielles des vocalisations émotionnelles du set 2 selon l'intensité émotionnelle faible, moyenne et élevée

Nous avons analysé la distribution de l'étendue fréquentielle des échantillons émotionnels en fonction des trois niveaux d'intensité émotionnelle. Comme pour l'étendue fréquentielle conversationnelle, nous avons calculé deux étendues fréquentielles pour chaque échantillon (tableau 5), soit une étendue allant de la F0 du quantile 25 à la F0 maximale (l'étendue maximale), et une étendue allant de la F0 du quantile 25 à la F0 du quantile 75 (l'étendue sectionnée).

Tableau 5 : Étendues fréquentielles des vocalisations émotionnelles en demi-tons mesurées en fonction de l'intensité émotionnelle cotée par les participants

	Moyenne F0 Qt 25 – F0 max (ET)	Moyenne F0 Qt 25 – F0 Qt 75 (ET)
1	6,46 (5,90)	2,99 (3,44)
2	7,23 (5,99)	3,76 (3,66)
3	7,19 (6,25)	3,61 (3,06)
Total	6,97 (6,05)	3,46 (3,41)

Légende : 1 = intensité faible ; 2 = intensité moyenne ; 3 = intensité élevée ; ET= écart-type

Moyenne F0 Qt 25 – F0 max = Moyenne des étendues émotionnelles en demi-tons mesurée entre les moyennes des fréquences des quantiles 25 et de la fréquence maximale des échantillons émotionnels.

Moyenne F0 Qt 25 – F0 Qt 75 = Moyenne des étendues émotionnelles en demi-tons mesurée entre les moyennes des fréquences des quantiles 25 et des quantiles 75 des échantillons émotionnels.

La distribution des données d'étendue fréquentielle nous montre une fréquence d'apparition élevée d'étendues en-dessous de 10 demi-tons pour l'étendue maximale et en-dessous de 5 demi-tons pour l'étendue sectionnée. Nous relevons également une distribution asymétrique positive sur toutes nos étendues (figure 10).

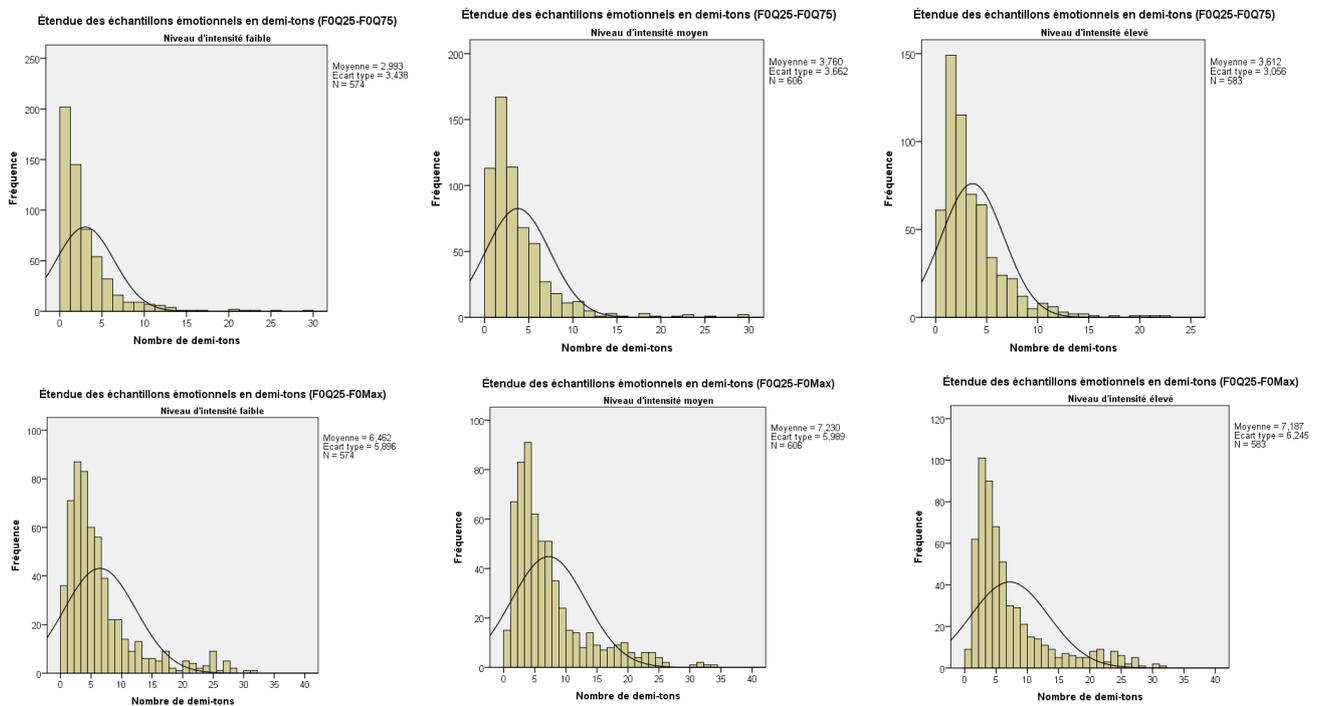


Figure 10. Histogrammes de distribution des données fréquentielles émotionnelles selon l'intensité émotionnelle, et courbe gaussienne : haut : étendue entre la fréquence du quantile 25 et le quantile 75 ; bas : étendue entre la fréquence du quantile 25 et la fréquence maximale

E. Résultats statistiques au regard de l'hypothèse 1 :

« L'étendue émotionnelle effectuera une ascension plus importante et dépassera plus l'étendue fréquentielle pour les émotions à forte intensité que pour les émotions à faible intensité, de manière régulière chez tous les participants et en ne tenant pas compte de la valence »

Nous avons utilisé les trois catégories d'intensité émotionnelle créées précédemment pour générer des scores moyens d'étendues fréquentielles en demi-tons des échantillons vocaux émotionnels de nos participants. Ces scores moyens seront utilisés par la suite dans les analyses de variance à mesures répétées (Anova).

1. Analyse de l'étendue fréquentielle des vocalisations émotionnelles selon les niveaux d'intensité cotée par les participants

Pour analyser la significativité des différences en fonction des intensités, nous avons réalisé les Anova sur l'étendue fréquentielle maximale (Qt25-F0max) et sur l'étendue fréquentielle sectionnée (Qt25-Qt75)

a) Analyse de l'étendue fréquentielle maximale en fonction du niveau d'intensité émotionnelle

De manière descriptive, nous pouvons observer une augmentation de l'étendue fréquentielle des vocalisations émotionnelles selon le niveau d'intensité émotionnelle (tableau 6, figure 11). En revanche, nous devons vérifier si ces résultats sont significatifs.

Pour cela, nous avons effectué une analyse de variance à mesures répétées (Anova). L'étendue fréquentielle maximale a été introduite comme variable dépendante et le niveau d'intensité émotionnelle comme variable indépendante (variable intra-participants). Le genre et le pays de l'étude ont été introduits comme facteurs inter-participants. Le test de sphéricité de Mauchly est significatif ($p=0,038$), cela nous amène à rejeter l'hypothèse de sphéricité et à ajuster les degrés de liberté selon la correction de Greenhouse-Geisser. Le test de Levene est non-significatif pour les trois niveaux d'intensité ($F_1(3,36) = 0,166$; $p = 0,919$, $F_2(3,36) = 2,557$; $p = 0,07$, $F_3(3,36) = 0,721$; $p = 0,546$), ce qui indique que les variables sont égales dans les différents groupes.

Tableau 6 : Étendues fréquentielles maximales des vocalisations émotionnelles en demi-tons mesurées en fonction de l'intensité émotionnelle cotée par le participant

		Moyenne F0 Qt 25 – F0 max Intensité 1 (ET)	Moyenne F0 Qt 25 – F0 max Intensité 2 (ET)	Moyenne F0 Qt 25 – F0 max Intensité 3 (ET)
F	Qc	7,48 (2,59)	7,80 (1,48)	8,01 (3,53)
	Be	6,91 (2,38)	7,18 (2,63)	6,92 (2,62)
H	Qc	4,96 (1,73)	7,04 (3,66)	6,78 (2,26)
	Be	5,19 (1,99)	5,88 (2,10)	6,46 (3,35)
Total F		7,26 (2,48)	7,56 (1,97)	7,59 (3,20)
Total H		5,13 (1,85)	6,21 (2,54)	6,55 (2,10)
Total		6,51 (2,48)	7,09 (2,25)	7,23 (3,13)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1= intensité faible, 2= intensité moyenne, 3 = intensité élevée, ET= écart-type

Moyenne F0 Qt 25 – F0 max = Moyenne des étendues émotionnelles en demi-tons mesurée entre les moyennes des fréquences des quantiles 25 et de la fréquence maximale des échantillons émotionnels.

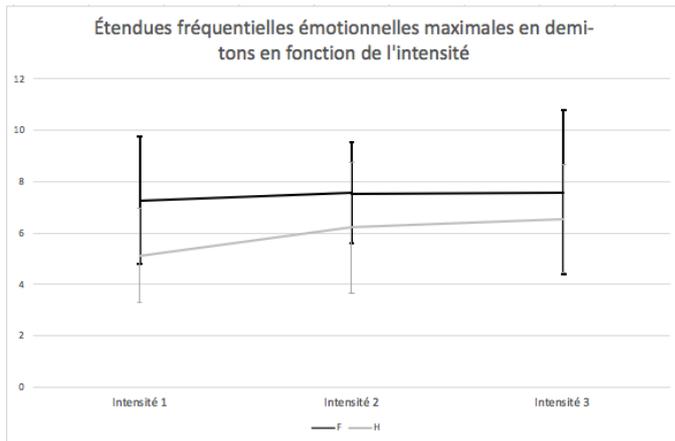


Figure 11. Étendues fréquentielles maximales des vocalisations émotionnelles en fonction de l'intensité

Les résultats sont non-significatifs. Il n'y a pas d'effet principal intra-participant du niveau d'intensité émotionnelle sur l'étendue fréquentielle maximale des échantillons émotionnels (tableau 7). Il n'y a pas non plus d'effet inter-participant du genre ou du pays sur l'étendue fréquentielle maximale des échantillons émotionnels (tableau 8).

Tableau 7 : Test des effets intra-participants pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en fonction de l'intensité émotionnelle

		F(ddl)	P	Eta-carré partiel
Greenhouse-Geisser	Intensité	F(1.71) = 1,66	0,201	0,044
	Intensité X Pays	F(1.71) = 0,23	0,760	0,006
	Intensité X genre	F(1.71) = 0,78	0,445	0,21
	Intensité X pays X genre	F(1.71) = 0,24	0,752	0,007

Tableau 8 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en fonction de l'intensité émotionnelle

	F(ddl)	P	Eta-carré partiel
Pays	F(1) = 0,74	0,395	0,020
Genre	F(1) = 3,81	0,059	0,096
Pays X Genre	F(1) = 0,064	0,802	0,002

b) Analyse de l'étendue fréquentielle sectionnée en fonction du niveau d'intensité émotionnel

De manière descriptive, nous pouvons observer une augmentation de l'étendue fréquentielle des vocalisations émotionnelles selon le niveau d'intensité émotionnelle pour les hommes. En revanche, l'étendue fréquentielle est plus élevée chez les femmes pour le niveau d'intensité moyen (tableau 9, figure 12).

Pour vérifier si ces résultats sont significatifs, nous avons effectué une analyse de variance à mesures répétées. L'étendue fréquentielle sectionnée a été introduite comme variable dépendante et le niveau d'intensité émotionnelle comme variable indépendante (variable intra-participants). Le genre et le pays de l'étude ont été introduits comme facteurs inter-participants. Le test de sphéricité de Mauchly étant non-significatif ($p=0,347$), les valeurs de F peuvent être présentées sans ajustement des degrés de liberté. Le test de Levene est non-significatif pour les

trois niveaux d'intensité ($F(3,36) = 1,947 ; p = 0,139$, $F(3,36) = 1,437 ; p = 0,248$, $F(3,36) = 0,587 ; p = 0,627$), ce qui indique que les variables sont égales dans les différents groupes.

Tableau 9 : Étendues fréquentielles sectionnées des vocalisations émotionnelles en demi-tons mesurées en fonction de l'intensité émotionnelle cotée par les participants

		Moyenne F0 Qt 25 – F0 Qt75 Intensité 1 (ET)	Moyenne F0 Qt 25 – F0 Qt75 Intensité 2 (ET)	Moyenne F0 Qt 25 – F0 Qt75 Intensité 3 (ET)
F	Qc	3,32 (1,54)	4,05 (1,24)	3,53 (1,45)
	Be	3,22 (1,64)	3,83 (1,45)	4,01 (1,40)
H	Qc	1,89 (0,19)	3,54 (2,46)	3,55 (1,87)
	Be	2,72 (0,99)	3,24 (1,20)	3,36 (1,11)
Total F		3,28 (1,54)	3,96 (1,30)	3,71 (1,42)
Total H		2,48 (0,92)	3,32 (1,55)	3,41 (1,29)
Total		2,99 (1,40)	3,74 (1,41)	3,61 (1,37)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1= intensité faible, 2= intensité moyenne, 3 = intensité élevée, ET= écart-type

Moyenne F0 Qt 25 – F0 Qt75 = Moyenne des étendues émotionnelles en demi-tons mesurée entre les moyennes des fréquences des quantiles 25 et des quantiles 75 des échantillons émotionnels.

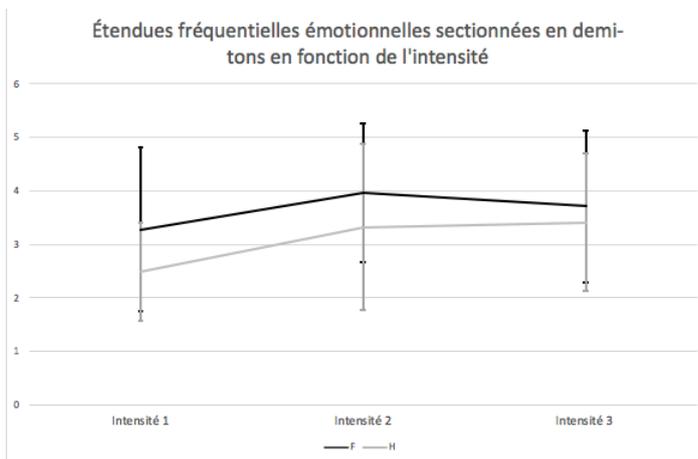


Figure 12. Étendues fréquentielles sectionnées des vocalisations émotionnelles en fonction de l'intensité

Nous observons (tableau 10) un effet intra-participants principal du niveau d'intensité sur l'étendue fréquentielle sectionnée ($F(2,35)=7,125$; $p=0,003$, éta-carré partiel= 0,289). Nous ne relevons aucun effet principal des facteurs inter-participants (tableau 11) de genre ou de pays, ni d'effet d'interaction. Les tests post-hoc avec correction de Bonferroni (annexe B.5) nous montrent que la différence d'étendue fréquentielle sectionnée est significative entre le niveau d'intensité faible par rapport aux niveaux moyens ($p=0,002$) et élevés ($p=0,025$) mais qu'il n'y a pas de différence significative entre les étendues fréquentielles sectionnées des niveaux d'intensité moyens et élevés.

Tableau 10 : Test multivarié pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en fonction de l'intensité émotionnelle cotée par les participants

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de Wilks	Intensité	$F(2,35)= 7,13$	0,003	0,289
	Intensité X Pays	$F(2,35)= 0,89$	0,421	0,048
	Intensité X genre	$F(2,35)= 0,65$	0,528	0,036
	Intensité X pays X genre	$F(2,35)= 0,98$	0,386	0,053

Tableau 11 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en fonction de l'intensité émotionnelle cotée par les participants

	F(ddl)	P	Eta-carré partiel
Pays	F(1)= 0,048	0,828	0,001
Genre	F(1)= 2,42	0,129	0,063
Pays X Genre	F(1)= 0,005	0,942	0,000

2. Analyse de distance entre l'étendue fréquentielle des échantillons vocaux émotionnels et l'étendue fréquentielle conversationnelle des participants

Nous voulions savoir si les vocalisations émotionnelles se situaient dans un registre particulier de l'étendue fréquentielle de nos participants. Nous faisons l'hypothèse que plus l'intensité émotionnelle est forte, plus les vocalisations se situeraient proches des limites aiguës de l'étendue fréquentielle du locuteur.

Nous avons analysé cela en définissant trois variables de distance en demi-tons. La première variable, DistMin, concerne la distance entre la fréquence du quantile 25 de la vocalisation émotionnelle et la fréquence du quantile 25 de l'étendue fréquentielle conversationnelle du locuteur. La deuxième variable, DistMax, concerne la distance entre la fréquence maximale de la vocalisation émotionnelle et la fréquence maximale de l'étendue fréquentielle conversationnelle du locuteur. La troisième variable, DistSect, concerne la distance entre la fréquence du quantile 75 de la vocalisation émotionnelle et la fréquence du quantile 75 de l'étendue fréquentielle conversationnelle du locuteur.

Pour démontrer notre hypothèse, nous avons fait une analyse de variance à mesures répétées pour analyser chacune des distances en fonction du niveau d'intensité émotionnel coté par le participant.

a) *Analyse de la distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels des participants (DistMin)*

Nous avons effectué une analyse de variance à mesures répétées. La variable DistMin a été définie comme variable dépendante et le niveau d'intensité émotionnelle coté par le participant a été défini comme variable indépendante (facteur intra-participants). Le genre et le pays de l'étude ont été introduits comme facteurs inter-participants. Le test de sphéricité de Mauchly étant non-significatif ($p=0,904$), les valeurs de F peuvent être présentées sans ajustement des degrés de liberté. Le test de Levene est non-significatif pour les trois niveaux d'intensité ($F_1(3,36) = 3,70 ; p = 0,200$, $F_2(3,36) = 0,93 ; p = 0,435$, $F_3(3,36) = 1,31 ; p = 0,287$), ce qui indique que les variables sont égales dans les différents groupes.

Tableau 12 : Distance entre les fréquences du quantile 25 des échantillons émotionnels et les fréquences du quantile 25 des étendues fréquentielles conversationnelles

		Moyenne DistMin Intensité 1 (ET)	Moyenne DistMin Intensité 2 (ET)	Moyenne DistMin Intensité 3 (ET)
F	Qc	1,69 (2,21)	2,29 (2,09)	2,46 (2,34)
	Be	0,40 (2,19)	1,21 (2,75)	2,40 (2,16)
H	Qc	1,08 (0,76)	2,01 (2,22)	1,49 (2,25)
	Be	0,96 (0,76)	1,41 (1,57)	1,67 (1,70)
Total F		1,19 (2,25)	1,87 (2,37)	2,44 (2,23)
Total H		0,99 (0,73)	1,58 (1,71)	1,62 (1,78)
Total		1,13 (1,85)	1,77 (2,15)	2,15 (2,09)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1= intensité faible, 2= intensité moyenne, 3 = intensité élevée, ET= écart-type

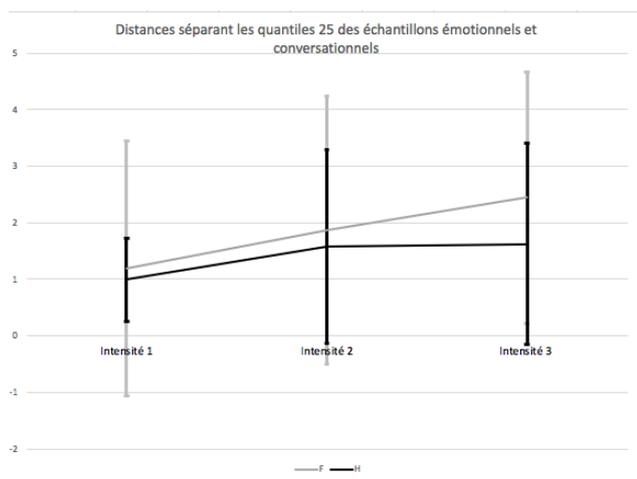


Figure 13. Distances séparant les quantiles 25 des échantillons émotionnels et de l'étendue conversationnelle

Nous observons un effet intra-participants principal du niveau d'intensité sur la variable DistMin ($F(2,35)=4,59$; $p=0,017$, éta-carré partiel=0,208) (tableau 13) mais aucun effet principal des facteurs inter-participants de genre ou de pays, ni d'effet d'interaction (tableau 14). Les tests post-hoc avec correction de Bonferroni (annexe B.6) nous montrent que la différence de DistMin est significative entre le niveau d'intensité émotionnelle faible par rapport au niveau d'intensité émotionnelle élevée ($p=0,014$). Les autres différences ne sont pas significatives.

Tableau 13 : Test multivarié pour la distance entre les fréquences du quantile 25 des échantillons émotionnels et les fréquences du quantile 25 des étendues fréquentielles conversationnelles

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de Wilks	Intensité	F(2,35)= 4,59	0,017	0,208
	Intensité X Pays	F(2,35)= 1,20	0,313	0,064
	Intensité X genre	F(2,35)= 1,11	0,341	0,060
	Intensité X pays X genre	F(2,35)= 0,276	0,761	0,016

Tableau 14 : Test des effets inter-participants pour la distance entre les fréquences du quantile 25 des échantillons émotionnels et les fréquences du quantile 25 des étendues fréquentielles conversationnelles

	F(ddl)	P	Eta-carré partiel
Pays	F(1)= 0,595	0,445	0,016
Genre	F(1)= 0,229	0,635	0,006
Pays X Genre	F(1)= 0,245	0,624	0,007

b) *Analyse de la distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels des participants (DistMax)*

Nous avons effectué une analyse de variance à mesures répétées. La variable DistMax a été définie comme variable dépendante et le niveau d'intensité émotionnelle coté par le participant a été défini comme variable indépendante (facteur intra-participants). Le genre et le pays de l'étude ont été introduits comme facteurs inter-participants. Le test de sphéricité de Mauchly étant non-significatif ($p=0,334$), les valeurs de F peuvent être présentées sans ajustement des degrés de liberté. Le test de Levene est non-significatif pour les trois niveaux d'intensité ($F_1(3,36) = 0,803 ; p = 0,500$, $F_2(3,36) = 0,314 ; p = 0,815$, $F_3(3,36) = 0,333 ; p = 0,801$), ce qui indique que les variables sont égales dans les différents groupes.

Tableau 15 : Distance entre les fréquences maximales des échantillons émotionnels et les fréquences maximales des étendues fréquentielles conversationnelles

		Moyenne DistMax Intensité 1 (ET)	Moyenne DistMax Intensité 2 (ET)	Moyenne DistMax Intensité 3 (ET)
F	Qc	-4,33 (3,74)	-3,42 (3,28)	-3,04 (4,27)
	Be	-7,75 (3,52)	-6,67 (3,98)	-6,54 (4,16)
H	Qc	-16,72 (3,62)	-13,71 (4,88)	-14,48 (3,93)
	Be	-8,22 (3,10)	-7,08 (3,38)	-5,92 (3,76)
Total F		-5,65 (3,96)	-4,67 (3,84)	-4,38 (4,50)
Total H		-10,65 (5,06)	-8,98 (4,80)	-8,37 (5,43)
Total		-7,40 (4,94)	-6,18 (4,63)	-5,78 (5,14)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1= intensité faible, 2= intensité moyenne, 3 = intensité élevée, ET= écart-type

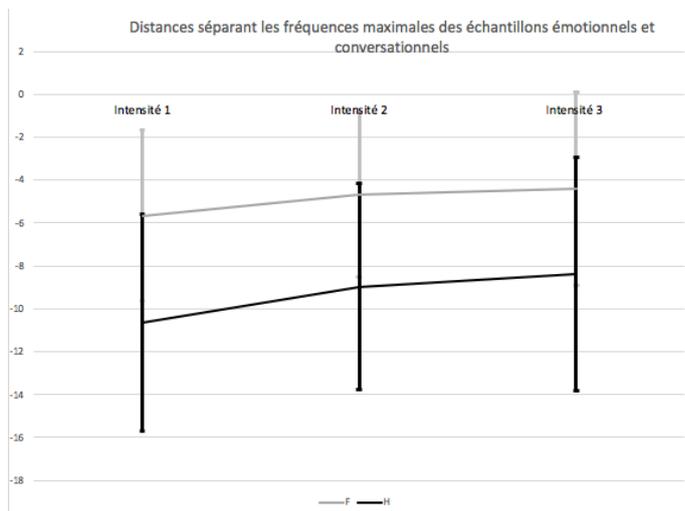


Figure 14. Distances séparant les fréquences maximales des échantillons émotionnels et de l'étendue conversationnelle

Nous observons un effet intra-participants principal du niveau d'intensité sur la variable DistMax ($F(2,35)=7,238; p=0,002$, éta-carré partiel=0,293) (tableau 16). Les tests post-hoc avec correction de Bonferroni (annexe B.7) nous montrent que la différence de DistMax est significative entre le niveau d'intensité émotionnelle faible par rapport aux niveaux

d'intensité émotionnelle moyens ($p=0,011$) et élevés ($p=0,009$). Les autres différences ne sont pas significatives.

Tableau 16 : Test multivarié pour la distance entre les fréquences maximales des échantillons émotionnels et les fréquences maximales des étendues fréquentielles conversationnelles

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de Wilks	Intensité	F(2,35)= 7,238	0,002	0,293
	Intensité X Pays	F(2,35)= 0,406	0,670	0,023
	Intensité X genre	F(2,35)= 0,748	0,481	0,041
	Intensité X pays X genre	F(2,35)= 0,606	0,551	0,033

Tableau 17 : Test des effets inter-participants pour la distance entre les fréquences maximales des échantillons émotionnels et les fréquences maximales des étendues fréquentielles conversationnelles

	F(ddl)	P	Eta-carré partiel
Pays	F(1)= 3,621	0,065	0,091
Genre	F(1)= 23,446	0,076	0,394
Pays X Genre	F(1)= 22,743	0,085	0,387

c) Analyse de distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels des participants (DistSect)

Nous avons effectué une analyse de variance à mesures répétées. La variable DistSect a été définie comme variable dépendante et le niveau d'intensité émotionnelle coté par le participant a été défini comme variable indépendante (facteur intra-participants). Le genre et le pays de l'étude ont été introduits comme facteurs inter-participants. Le test de sphéricité de Mauchly étant non-significatif ($p=0,85$), les valeurs de F peuvent être présentées sans ajustement des degrés de liberté. Le test de Levene est non-significatif pour les trois niveaux d'intensité ($F_1(3,36) = 1,844 ; p = 0,157$, $F_2(3,36) = 0,743 ; p = 0,534$, $F_3(3,36) = 0,115 ; p = 0,951$), ce qui indique que les variables sont égales dans les différents groupes.

Tableau 18: Distance entre les fréquences des quantiles 75 des échantillons émotionnels et les fréquences des quantiles 75 des étendues fréquentielles conversationnelles

		Moyenne DistSect Intensité 1 (ET)	Moyenne DistSect Intensité 2 (ET)	Moyenne DistSect Intensité 3 (ET)
F	Qc	2,15 (2,49)	3,47 (2,16)	3,12 (2,40)
	Be	0,45 (2,93)	1,88 (3,53)	2,99 (2,79)
H	Qc	0,59 (0,66)	3,17 (2,43)	2,67 (2,16)
	Be	0,16 (1,60)	1,12 (2,39)	1,56 (2,14)
Total F		1,50 (2,74)	2,86 (2,81)	3,07 (2,50)
Total H		0,28 (1,38)	1,71 (2,50)	1,88 (2,12)
Total		1,07 (2,41)	2,46 (2,73)	2,66 (2,42)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1= intensité faible, 2= intensité moyenne, 3 = intensité élevée, ET= écart-type

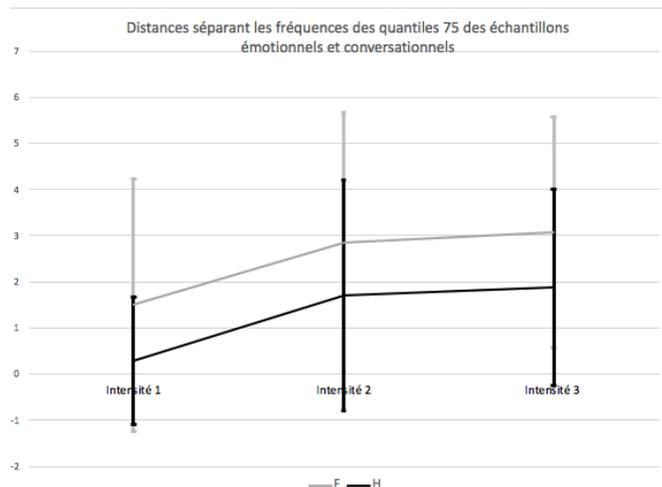


Figure 15. Distances séparant les fréquences des quantiles 75 des échantillons émotionnels et de l'étendue conversationnelle

Nous observons un effet intra-participants principal du niveau d'intensité sur la variable DistSect ($F(2,35) = 11,615$; $p < 0,001$, éta-carré partiel = 0,399) (tableau 19). Les tests post-hoc avec correction de Bonferroni (annexe B.8) nous montrent que la différence de DistSect est significative entre le niveau d'intensité émotionnelle faible par rapport aux niveaux d'intensité émotionnelle moyens ($p = 0,001$) et élevés ($p < 0,001$).

Tableau 19 : Test multivarié pour la distance entre les fréquences du quantile 75 des échantillons émotionnels et les fréquences du quantile 75 des étendues fréquentielles conversationnelles

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de Wilks	Intensité	$F(2,35) = 11,615$	<0,001	0,399
	Intensité X Pays	$F(2,35) = 1,065$	0,355	0,057
	Intensité X genre	$F(2,35) = 0,168$	0,846	0,010
	Intensité X pays X genre	$F(2,35) = 1,053$	0,360	0,057

Tableau 20 : Test des effets inter-participants pour la distance entre les fréquences du quantile 75 des échantillons émotionnels et les fréquences du quantile 75 des étendues fréquentielles conversationnelles

	F(ddl)	P	Eta-carré partiel
Pays	$F(1) = 2,426$	0,128	0,063
Genre	$F(1) = 1,144$	0,292	0,031
Pays X Genre	$F(1) = 0,001$	0,970	0,000

3. Réponse à l'hypothèse 1

Nous observons une augmentation de l'amplitude des étendues fréquentielles des échantillons émotionnels lorsque le niveau d'intensité coté par le participant augmente. Cette différence est statistiquement significative pour l'étendue sectionnée entre les niveaux d'intensité faibles et moyens et entre les niveaux d'intensité faibles et élevés (annexe B.8).

De plus, nous observons également que plus le niveau d'intensité émotionnelle augmente, plus les échantillons vocaux émotionnels se déplacent vers le registre aigu du locuteur. On le remarque grâce au fait que plus le niveau d'intensité est élevé, plus la fréquence du quantile 25 des échantillons émotionnels s'éloigne de la fréquence du quantile 25 conversationnelle

(tableaux 12 et 13) et plus la fréquence maximale des échantillons émotionnels se rapproche de la fréquence maximale conversationnelle (tableaux 15, 16).

Ensuite, lorsque l'on observe la fréquence du quantile 75 des échantillons émotionnels (figure 16), on peut remarquer que celle-ci est toujours plus élevée que la fréquence du quantile 75 conversationnelle. La distance les séparant augmente avec les niveaux d'intensité émotionnelle.

En conclusion, nos résultats soutiennent notre première hypothèse, les étendues fréquentielles des participants augmentent en fonction de l'intensité émotionnelle ressentie par le participant.

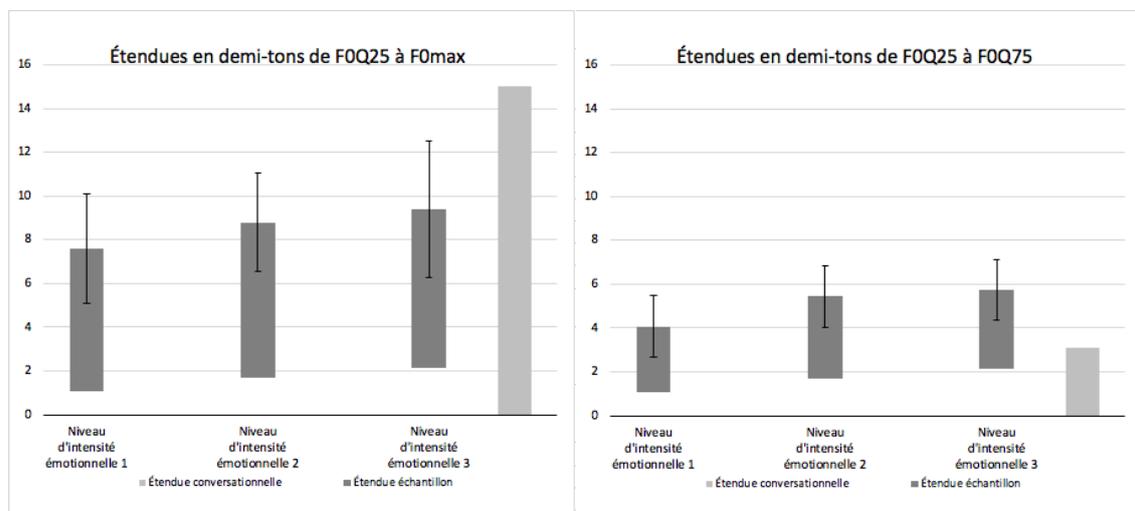


Figure 16. Comparaison des excursions des étendues fréquentielles maximales et sectionnées en fonction de l'intensité

F. Résultats statistiques au regard de l'hypothèse 2 :

« Les étendues fréquentielles des vocalisations spontanées seront plus petites que celles de vocalisations imposées »

L'idéal pour répondre à cette hypothèse aurait été d'employer également le niveau d'intensité émotionnelle côté par le participant. Malheureusement, les données n'ont pas été enregistrées pour le set spontané (1). Pour pallier cela, nous allons utiliser les scores de valence pour tenter de trouver une corrélation entre les scores d'intensité et les scores de valence. Nous n'emploierons pas les données d'intensité de l'IAPS car celles-ci ne se sont pas montrées justes pour notre première hypothèse.

1. Liens entre les intensités et valences estimées du participant

Nous avons cherché à trouver quelle variable nous permettrait d'estimer au mieux le niveau d'intensité des participants du set 1. Pour cela, nous avons réalisé des régressions linéaires et des régressions logistiques (annexe B.9)

Ces analyses nous ont indiqué qu'il était préférable de nous baser sur les scores de valence émotionnelle cotés par le participant, plutôt que sur les scores de l'IAPS, afin de réaliser des catégories de niveaux d'intensité émotionnelle pour le set 1. En effet, 42 % de la variance des scores de l'intensité émotionnelle s'expliquait par les scores de valence ($r^2=0,424$, $p<0,001$). En revanche, les scores d'intensité de l'IAPS ne pouvaient pas nous aider à estimer ces valeurs ($r^2<0,240$, $p<0,001$). Ensuite, nous avons réalisé une dernière analyse entre les scores de valence émotionnelle de l'IAPS et ceux des sets 1 et 2, il en ressort que les participants se sont comportés de la même manière et qu'une relation forte existe entre leurs scores de valence et d'intensité personnels.

2. Analyses des différences d'étendues fréquentielles émotionnelles selon le set

Nous sommes partis du principe que peu importe que la valence soit positive ou négative, plus le score est élevé ou bas, plus l'intensité émotionnelle perçue est élevée. Nous avons transformé les scores de valence de manière à ce que toutes les données de 50-100 soient considérées sur la même échelle que celles de 0-50. Pour faire cela, nous avons soustrait 50 à tous les scores de valence et ensuite transformé les scores en valeurs positives. Ainsi, les scores sont maintenant tous ordonnés de 0-50 où 0 correspond aux intensités les plus faibles et 50 aux

intensités les plus élevées. Ce sont donc des scores d'intensité émotionnelle estimés à partir des scores de valence cotés par les participants, ceci étant la meilleure manière de remplacer les données perdues lors des enregistrements.

Suite à cela, pour tous les participants, les scores ont été transformés en rangs pour les deux sets combinés. En effet, d'autres analyses ont mis en évidence une différence de taux de réponses en fonction du set (annexe B.10), il était donc préférable de les combiner. En partant des rangs calculés, les scores ont été transformés en trois catégories d'intensité estimées (faible, moyenne et élevée). Nous allons comparer les excursions fréquentielles émotionnelles par rapport aux sets. Nous allons donc regarder comment se déplacent les étendues émotionnelles au regard de l'étendue conversationnelle que nous considérons comme fixe.

Les analyses présentées comparent d'abord les différences entre les étendues fréquentielles maximales, entre les étendues fréquentielles sectionnées puis les différentes distances entre les quantiles 25 (DistMin), quantiles 75 (DistSect) et fréquences maximales (DistMax).

a) Analyses des différences d'étendues fréquentielles émotionnelles maximales en fonction du niveau d'intensité émotionnelle estimé et du set

Nous avons effectué une Anova à mesures répétées. L'étendue fréquentielle a été définie comme variable dépendante et les sets et les niveaux d'intensité estimés ont été définis comme variables indépendantes (facteurs intra-participants) et le genre et le pays comme facteurs inter-participants.

Le test de sphéricité de Mauchly est non-significatif ($p=0,835$), nous pouvons donc utiliser les données sans ajustement du degré de liberté. Le test de Levene est significatif pour les étendues du set 1 et niveau 3 et pour le set 1 et niveau 2¹ ($F_{s1n1}(2,35)= 0,891 ; p=0,456$, $F_{s2n1}(2,35)=0,976 ; p=0,415$, $F_{s1n2}(3,35)= 6,418 ; p=0,001$, $F_{s2n2}(3,35)=1,615 ; p=0,203$, $F_{s1n3}(3,35) = 3,807 ; p= 0,018$, $F_{s2n3}(3,35)= 0,157 ; p=0,925$), ce qui indique que les variables ne sont pas égales dans les différents groupes et que les résultats seront à prendre avec précaution.

¹ F_{s*n*} = Valeur F du set * pour le niveau *

Tableau 21 : Étendues fréquentielles maximales des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 1

		Moyenne des étendues intensité faible (ET)	Moyenne des étendues intensité moyenne (ET)	Moyenne des étendues intensité élevée (ET)
F	Qc	7,52 (2,91)	8,71 (3,07)	8,51 (1,87)
	Be	6,02 (2,99)	6,74 (2,76)	7,50 (3,56)
H	Qc	7,12 (1,35)	11,31 (9,52)	9,27 (1,35)
	Be	6,17 (2,36)	5,82 (1,52)	5,85 (2,28)
Total F		6,92 (2,98)	7,92 (3,05)	8,10 (2,66)
Total H		6,44 (2,12)	7,39 (5,39)	6,82 (2,56)
Total		6,75 (2,68)	7,73 (3,99)	7,64 (2,66)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Tableau 22 : Étendues fréquentielles maximales des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 2

		Moyenne des étendues intensité faible (ET)	Moyenne des étendues intensité moyenne (ET)	Moyenne des étendues intensité élevée (ET)
F	Qc	6,91 (2,48)	8,10 (2,59)	8,03 (2,93)
	Be	7,20 (2,73)	7,35 (3,39)	7,22 (1,96)
H	Qc	4,76 (1,30)	7,18 (1,98)	6,54 (3,40)
	Be	5,22 (2,32)	5,92 (1,77)	6,48 (2,84)
Total F		7,03 (2,53)	7,80 (2,89)	7,71 (2,57)
Total H		5,09 (2,04)	6,28 (1,85)	6,50 (2,87)
Total		6,33 (2,52)	7,25 (2,65)	7,27 (2,71)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Nous observons un effet intra-participants principal du niveau d'intensité estimé (F (2,34) =4,648, p=0,016, éta-carré partiel= 0,215) et du set (F (2,34) =5,344, p=0,027, éta-carré partiel= 0,132) sur l'étendue des vocalisations (tableau 23). Plus le niveau d'intensité émotionnelle est élevé, plus l'étendue fréquentielle augmente chez les participants. Nous relevons également un effet intra-participants d'interaction entre le set et le pays (F (2,34) =8,591, p=0,006, éta-carré partiel= 0,197) et entre le set et le genre (F (2,34) =4,949, p=0,006, éta-carré partiel= 0,124). Nous observons aussi un effet principal du facteur inter-participants (tableau 24) pays (F (1,35) =4,106, p=0,050, éta-carré partiel=0,927), les québécois ont des étendues plus importantes que les belges au set 1. La différence entre le set 1 et le set 2 est significative, l'étendue fréquentielle maximale est plus grande que pour le set 2. Les tests post-hoc avec correction de Bonferroni (annexe B.11) nous montrent que la différence d'étendue fréquentielle selon le set est significative entre le niveau d'intensité faible et le niveau élevé (p=0,047) et entre les niveaux faibles et moyens (p=0,025) mais qu'il n'y a pas de différence significative entre les étendues fréquentielles des niveaux d'intensité moyens et élevés.

Tableau 23 : Test multivarié pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimée

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de	Intensité estimée	F(2,34)= 4,648	0,016	0,215
Wilks	Intensité estimée X Pays	F(2,34)=2,206	0,126	0,115
	Intensité estimée X genre	F(2,34)=0,514	0,603	0,029
	Intensité estimée X pays X genre	F(2,34)=0,826	0,446	0,046
	Set	F(1,35)= 5,344	0,027	0,132
	Set X Pays	F(1,35)=8,591	0,006	0,197
	Set X Genre	F(1,35)=4,949	0,033	0,124
	Set X pays X genre	F(1,35)=1,931	0,173	0,052
	Intensité estimée X set	F(2,34)=0,087	0,917	0,005
	Intensité estimée X set X pays	F(2,34)=0,233	0,793	0,014
	Intensité estimée X set X genre	F(2,34)=0,699	0,504	0,040
	Intensité estimée X set X pays X genre	F(2,34)=0,928	0,405	0,052

Tableau 24 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle maximale des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimée

	F(ddl)	P	Eta-carré partiel
Pays	F(1,35)=4,106	0,050	0,927
Genre	F(1,35)=1,016	0,320	0,105
Pays X Genre	F(1,35)=0,76	0,544	0,028

b) Analyses des différences d'étendues fréquentielles émotionnelles sectionnées en fonction du niveau d'intensité émotionnelle estimé et du set

Nous avons effectué une Anova à mesures répétées. L'étendue fréquentielle a été définie comme variable dépendante et les sets et niveaux d'intensité estimés ont été définis comme variables indépendantes (facteurs intra-participants) et le genre et le pays comme facteurs inter-participants. Le test de sphéricité de Mauchly est non-significatif ($p=0,580$; $p=0,275$), nous pouvons donc utiliser les données sans ajustement du degré de liberté. Le test de Levene est significatif pour les étendues du set 1 et niveau 2 ($F_{s1n1}(2,35)= 2,351$; $p=0,089$, $F_{s2n1}(2,35)=0,873$; $p=0,464$, $F_{s1n2}(3,35)= 8,815$; $p<0,001$, $F_{s2n2}(3,35)=1,363$; $p=0,270$, $F_{s1n3}(3,35) = 0,768$; $p= 0,520$, $F_{s2n3}(3,35)= 1,906$; $p=0,147$), ce qui indique que les variables ne sont pas égales dans les différents groupes et que les résultats seront à prendre avec précaution pour cette interaction.

Tableau 25 : Étendues fréquentielles sectionnées des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 1

		Moyenne des étendues intensité faible (ET)	Moyenne des étendues intensité moyenne (ET)	Moyenne des étendues intensité élevée (ET)
F	Qc	3,22 (1,56)	4,12 (1,92)	4,15 (1,68)
	Be	3,32 (1,41)	3,49 (1,20)	4,14 (2,03)
H	Qc	2,97 (1,62)	5,94 (5,12)	4,63 (2,04)
	Be	2,52 (0,85)	3,28 (0,66)	3,45 (1,30)
Total F		3,26 (1,48)	3,86 (1,67)	4,14 (1,79)
Total H		2,65 (1,08)	4,04 (2,81)	3,79 (1,56)
Total		3,04 (1,36)	3,93 (2,12)	4,02 (1,70)

Tableau 26 : Étendues fréquentielles sectionnées des vocalisations émotionnelles en demi-tons mesurées en fonction des scores d'intensité estimés pour le set 2

		Moyenne des étendues intensité faible (ET)	Moyenne des étendues intensité moyenne (ET)	Moyenne des étendues intensité élevée (ET)
F	Qc	3,05 (1,27)	3,85 (1,22)	3,60 (1,33)
	Be	3,50 (1,35)	4,13 (1,80)	4,12 (1,14)
H	Qc	2,17 (0,49)	3,22 (1,63)	3,46 (2,58)
	Be	2,53 (1,03)	3,35 (1,09)	3,36 (1,15)
Total F		3,23 (1,30)	3,96 (1,45)	3,81 (1,26)
Total H		2,43 (0,90)	3,31 (1,20)	3,39 (1,57)
Total		2,94 (1,22)	3,73 (1,38)	3,66 (1,37)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Nous observons un effet intra-participant principal du niveau d'intensité estimé ($F(2,34)=10,724$, $p<0,001$, éta-carré partiel= 0,387) et du set ($F(1,35)=5,147$, $p=0,030$, éta-carré partiel= 0,128) sur l'étendue des vocalisations (tableau 27). Plus le niveau d'intensité émotionnelle est élevé, plus l'étendue fréquentielle augmente chez les participants. Nous relevons également un effet intra-participants du set ($F(1,35)=5,147$, $p=0,030$, éta-carré partiel= 0,128), le set 1 présente une plus grande étendue fréquentielle que le set 2. Nous relevons aussi un effet intra-participants d'interaction entre le set et le pays ($F(2,34)=9,031$, $p=0,005$, éta-carré partiel= 0,205) et entre le set et le genre ($F(2,34)=4,458$, $p=0,042$, éta-carré partiel= 0,113). En effet, les québécois présentent une plus grande étendue fréquentielle pour le set 1 que les belges. Cette différence est vraie aussi pour les hommes par rapport aux femmes. Nous n'observons aucun effet des facteurs inter-participants. Les tests post-hoc avec correction de Bonferroni (annexe B.12) nous montrent que la différence d'étendue fréquentielle sectionnée est significative entre le niveau d'intensité faible et le niveau élevé ($p=0,003$) ainsi qu'entre les niveaux faibles et moyens ($p<0,001$) mais qu'il n'y a pas de différence significative entre les étendues fréquentielles des niveaux d'intensité faibles et élevés.

Tableau 27 : Test multivarié pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimés

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de Wilks	Intensité estimée	F(2,34)= 10,724	<0,001	0,387
	Intensité estimée X Pays	F(2,34)=1,662	0,205	0,089
	Intensité estimée X genre	F(2,34)=1,398	0,261	0,076
	Intensité estimée X pays X genre	F(2,34)=0,352	0,706	0,020
	Set	F(1,35)= 5,147	0,030	0,128
	Set X Pays	F(2,34)=9,031	0,005	0,205
	Set X Genre	F(2,34)=4,458	0,042	0,113
	Set X pays X genre	F(2,34)=1,802	0,188	0,049
	Intensité estimée X set	F(2,34)=0,445	0,645	0,026
	Intensité estimée X set X pays	F(2,34)=0,985	0,384	0,055
	Intensité estimée X set X genre	F(2,34)=0,896	0,418	0,050
	Intensité estimée X set X pays X genre	F(2,34)=0,339	0,715	0,020

Tableau 28 : Test des effets inter-participants pour l'analyse de l'étendue fréquentielle sectionnée des vocalisations émotionnelles en demi-tons mesurée en fonction des scores d'intensité estimés

	F(ddl)	P	Eta-carré partiel
Pays	F(1,35)=0,547	0,464	0,015
Genre	F(1,35)=0,766	0,387	0,021
Pays X Genre	F(1,35)=1,133	0,295	0,031

c) Analyses de la distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé et du set (DistMin)

Nous avons effectué une Anova à mesures répétées. Les sets et les niveaux d'intensité estimés ont été définis comme facteurs intra-participants et le genre et le pays comme facteurs inter-participants. Les distances entre les quantiles 25 ont été définies comme variables dépendantes. Le test de sphéricité de Mauchly est non-significatif ($p=0,103$; $p=0,264$), nous pouvons donc utiliser les données sans ajustement du degré de liberté. Le test de Levene est non-significatif pour toutes les données (Fs1n1 (2,35) = 2,313 ; $p=0,093$, Fs2n1 (2,35)=3,236 ; $p=0,340$, Fs1n2 (3,35)= 0,660; $p=0,582$, Fs2n2 (3,35)=0,752 ; $p=0,529$, Fs1n3 (3,35) = 0,471; $p= 0,705$, Fs2n3 (3,35)= 0,669 ; $p=0,577$), ce qui indique que les variables sont égales dans les différents groupes.

Tableau 29 : Distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 1

		DistMin intensité faible (ET)	DistMin intensité moyenne (ET)	DistMin intensité élevée (ET)
F	Qc	2,19 (2,07)	1,20 (2,36)	2,36 (1,68)
	Be	1,35 (2,07)	1,12 (1,84)	1,78 (2,29)
H	Qc	1,77 (2,28)	0,86 (1,71)	2,95 (2,56)
	Be	0,87 (1,27)	1,44 (1,96)	1,29 (1,83)
Total F		1,85 (2,07)	1,64 (2,17)	2,13 (2,25)
Total H		1,13 (1,58)	1,27 (1,84)	1,77 (2,11)
Total		1,59 (1,92)	1,51 (2,04)	1,99 (2,18)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Tableau 30 : Distance entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 2

		DistMin intensité faible (ET)	DistMin intensité moyenne (ET)	DistMin intensité élevée (ET)
F	Qc	1,88 (2,02)	2,11 (2,23)	2,58 (2,70)
	Be	0,54 (1,66)	1,63 (3,03)	1,88 (1,80)
H	Qc	0,92 (1,08)	1,87 (1,86)	1,73 (2,54)
	Be	1,21 (0,99)	1,14 (1,24)	1,64 (1,85)
Total F		1,34 (1,97)	1,92 (2,53)	2,30 (2,36)
Total H		1,13 (0,99)	1,35 (1,41)	1,66 (1,96)
Total		1,26 (1,67)	1,71 (2,19)	2,07 (2,22)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Nous observons un effet intra-participants d'interaction entre le set, le pays et le genre ($F(2,34) = 4,890$, $p = 0,034$, éta-carré partiel = 0,123) ainsi qu'entre l'intensité estimée et le set ($F(2,34) = 3,449$, $p = 0,043$, éta-carré partiel = 0,169) (tableau 31). Pour le second set, plus l'intensité augmente, plus la distance entre les quantiles 25 augmente. De plus, les femmes québécoises ont une étendue fréquentielle plus grande que les femmes belges dans les deux sets, surtout pour le premier set. Nous n'observons aucun effet des facteurs inter-participants (tableau 32). Les tests post-hoc avec correction de Bonferroni (annexe B.13) ne nous montrent aucune différence significative entre les étendues fréquentielles des niveaux d'intensité faibles, moyens et élevés.

Tableau 31 : Test multivarié pour l'analyse des distances entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé

		F(ddl, erreur ddl)	P	Eta-carré partiel
Lambda de Wilks	Intensité estimée	F(2,34)= 2,154	0,132	0,112
	Intensité estimée X Pays	F(2,34)=0,022	0,978	0,001
	Intensité estimée X genre	F(2,34)=0,382	0,686	0,022
	Intensité estimée X pays X genre	F(2,34)=0,247	0,783	0,014
	Set	F(1,35)= 2,397	0,131	0,064
	Set X Pays	F(2,34)=1,794	0,189	0,049
	Set X Genre	F(2,34)=0,077	0,783	0,002
	Set X pays X genre	F(2,34)=4,890	0,034	0,123
	Intensité estimée X set	F(2,34)=3,449	0,043	0,169
	Intensité estimée X set X pays	F(2,34)=0,552	0,581	0,031
	Intensité estimée X set X genre	F(2,34)=0,329	0,722	0,019
	Intensité estimée X set X pays X genre	F(2,34)=0,455	0,638	0,026

Tableau 32 : Test des effets inter-participants pour l'analyse des distances entre la F0 du quantile 25 des échantillons vocaux émotionnels et la F0 du quantile 25 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé

	F(ddl)	P	Eta-carré partiel
Pays	F(1,35)=0,944	0,338	0,026
Genre	F(1,35)=0,240	0,627	0,007
Pays X Genre	F(1,35)=0,093	0,763	0,003

d) Analyses de la distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé et du set (DistSect)

Nous avons effectué une Anova à mesures répétées. Les sets et les niveaux d'intensité estimés ont été définis comme facteurs intra-participants et le genre et le pays comme facteurs inter-participants. Les distances entre les quantiles 25 ont été définies comme variables dépendantes. Le test de sphéricité de Mauchly est significatif ($p = 0,006$; $p = 0,010$), nous devons ajuster le degré de liberté par la correction de Greenhouse-Geisser. Le test de Levene est non-significatif pour toutes les données ($F_{s1n1}(3,35) = 1,065$; $p = 0,377$, $F_{s1n2}(3,35) = 0,761$;

$p=0,524$, $F_{s1n3}(3,35)=0,282$; $p=0,838$, $F_{s2n1}(2,35)=1,308$; $p=0,287$, $F_{s2n2}(2,35)=1,354$; $p=0,273$, $F_{s2n3}(3,35)=0,335$; $p=0,800$, ce qui indique que les variables sont égales dans les différents groupes.

Tableau 33 : Distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons en fonction du niveau d'intensité émotionnelle estimé pour le set 1

		DistSect intensité faible (ET)	DistSect intensité moyenne (ET)	DistSect intensité élevée (ET)
F	Qc	2,62 (2,43)	3,31 (2,42)	3,72 (2,67)
	Be	1,25 (2,77)	1,19 (2,43)	2,50 (2,78)
H	Qc	2,37 (1,63)	4,44 (4,08)	5,22 (2,61)
	Be	-0,08 (1,72)	1,25 (2,27)	1,2 (2,32)
Total F		2,08 (2,61)	2,46 (2,60)	3,23 (2,72)
Total H		0,62 (1,99)	2,16 (3,10)	2,40 (2,11)
Total		1,55 (2,48)	2,35 (2,76)	2,93 (2,80)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Tableau 34 : Distance entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 2

		DistSect intensité faible (ET)	DistSect intensité moyenne (ET)	DistSect intensité élevée (ET)
F	Qc	2,14 (2,27)	3,18 (2,27)	3,39 (2,48)
	Be	0,63 (2,58)	2,34 (3,27)	2,58 (2,30)
H	Qc	0,72 (1,32)	2,71 (1,63)	2,82 (2,86)
	Be	0,27 (1,46)	1,01 (2,17)	1,53 (2,31)
Total F		1,53 (2,46)	2,84 (2,68)	3,06 (2,40)
Total H		0,40 (1,39)	1,50 (2,12)	1,90 (2,44)
Total		1,13 (2,19)	2,36 (2,55)	2,54 (2,44)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Nous observons (tableau 35) un effet principal de l'intensité estimée sur les distances entre les quantiles 75 ($F(1,58,55,44)=9,286$, $p=0,001$, éta-carré partiel = 0,21). On observe également un effet intra-participants d'interaction entre l'intensité estimée, le pays et le genre ($F(1,58,55,44)=3,589$, $p=0,044$, éta-carré partiel= 0,093) ainsi qu'entre l'intensité estimée et le set ($F(1,62,56,59)=8,951$, $p<0,001$, éta-carré partiel= 0,204). Les québécois présentent une distance plus importante entre les quantiles 75 que les belges, leur excursion est plus importante. Nous observons un effet inter-participants pour le pays ($F(1,35)=6,025$, $p=0,019$, éta-carré-partiel= 0,147) (tableau 36). Les tests post-hoc avec correction de Bonferroni (annexe B.14) nous montrent une différence significative entre les étendues fréquentielles entre les niveaux d'intensité faibles et élevés ($p=0,002$) et entre les niveaux d'intensité moyens et élevés ($p=0,044$). Par contre, nous ne relevons aucune différence significative entre les niveaux d'intensité faibles et moyens.

Tableau 35 : Test des effets intra-participants pour l'analyse des distances entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé

		F(ddl, erreur ddl)	P	Eta-carré partiel
Greenhouse-Geisser	Intensité estimée	F(1,58,55,44)= 9,286	0,001	0,21
	Intensité estimée X Pays	F(1,58,55,44)=3,589	0,044	0,093
	Intensité estimée X genre	F(1,58,55,44)=2,900	0,075	0,077
	Intensité estimée X pays X genre	F(1,58,55,44)=1,486	0,235	0,041
	Set	F(1,35)= 3,349	0,076	0,087
	Set X Pays	F(1,35)=1,055	0,311	0,029
	Set X Genre	F(1,35)=0,182	0,673	0,101
	Set X pays X genre	F(1,35)=3,922	0,056	0,101
	Intensité estimée X set	F(1.62,56.59)=8,951	<0,001	0,204
	Intensité estimée X set X pays	F(1.62,56.59)=0,330	0,674	0,009
	Intensité estimée X set X genre	F(1.62,56.59)=0,107	0,858	0,003
Intensité estimée X set X pays X genre	F(1.62,56.59)=0,182	0,672	0,005	

Tableau 36 : Test des effets inter-participants pour l'analyse des distances entre la F0 du quantile 75 des échantillons vocaux émotionnels et la F0 du quantile 75 des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé

	F(ddl)	P	Eta-carré partiel
Pays	F(1,35)=6,025	0,019	0,147
Genre	F(1,35)=0,391	0,536	0,011
Pays X Genre	F(1,35)=0,365	0,550	0,010

e) *Analyses de la distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé et du set (DistMax)*

Nous avons effectué une Anova à mesures répétées. Les sets et les niveaux d'intensité estimés ont été définis comme facteurs intra-participants et le genre et le pays comme facteurs inter-participants. Les distances entre les fréquences maximales ont été définies comme variable dépendante. Le test de sphéricité de Mauchly est significatif pour l'intensité émotionnelle estimée ($p=0,003$; $p=0,287$), nous devons ajuster le degré de liberté par la correction de Greenhouse-Geisser. Le test de Levene est significatif pour l'intensité émotionnelle moyenne du set 1 ($F_{s1n1}(3,35)=0,950$; $p=0,427$, $F_{s1n2}(3,35)=3,275$; $p=0,032$, $F_{s1n3}(3,35)=1,118$; $p=0,355$, $F_{s2n1}(2,35)=0,263$; $p=0,851$, $F_{s2n2}(2,35)=0,327$; $p=0,806$, $F_{s2n3}(3,35)=0,275$; $p=0,843$), ce qui indique que nous devons utiliser ces données prudemment.

Tableau 37 : Distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 1

		DistMax intensité faible (ET)	DistMax intensité moyenne (ET)	DistMax intensité élevée (ET)
F	Qc	-3,74 (3,48)	-2,76 (3,47)	-2,58 (3,30)
	Be	-8,49 (4,66)	-8,00 (4,76)	-6,58 (4,93)
H	Qc	-13,87 (3,23)	-10,58 (9,82)	-10,53 (5,34)
	Be	-7,00 (3,94)	-6,79 (3,50)	-6,91 (3,79)
Total F		-5,64 (4,56)	-4,86 (4,73)	-4,18 (4,41)
Total H		-8,97 (4,85)	-7,88 (5,82)	-7,95 (4,41)
Total		-6,83 (4,88)	-5,94 (5,28)	-5,53 (4,72)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Tableau 38 : Distance entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé pour le set 2

		DistMax intensité faible (ET)	DistMax intensité moyenne (ET)	DistMax intensité élevée (ET)
F	Qc	-4,67 (3,99)	-3,24 (3,51)	-2,84 (4,07)
	Be	-8,12 (3,65)	-6,88 (3,93)	-6,76 (3,39)
H	Qc	-17,07 (3,37)	-13,71 (3,26)	-14,48 (4,90)
	Be	-7,62 (3,08)	-7,00 (3,26)	-5,93 (3,60)
Total F		-6,05 (4,16)	-4,70 (4,04)	-4,41 (4,22)
Total H		-10,32 (5,37)	-8,92 (4,44)	-8,38 (5,53)
Total		-7,58 (5,01)	-6,21 (4,61)	-5,83 (5,05)

F= femmes, H= hommes, Qc= Québec, Be= Belgique, 1 = intensité faible, 2 = intensité moyenne, 3 = intensité élevée, ET= écart-type

Nous observons (tableau 39) un effet principal de l'intensité estimée sur les distances entre les fréquences maximales ($F(1,55,54.38) = 7,331, p = 0,003$, éta-carré partiel = 0,173). On observe également un effet intra-participants d'interaction entre l'intensité estimée et le pays ($F(1,55,54.38) = 5,164, p = 0,014$, éta-carré partiel = 0,129) ainsi qu'entre l'intensité estimée et le set ($F(1,87,65.37) = 4,806, p = 0,013$, éta-carré partiel = 0,121). En effet, les hommes gardent une plus grande distance entre leur fréquence maximale émotionnelle et leur fréquence maximale conversationnelle. Les hommes québécois sont ceux présentant une plus petite excursion fréquentielle émotionnelle. Nous observons (tableau 40) un effet inter-participants pour le genre ($F(1,35) = 14,901, p < 0,001$, éta-carré-partiel = 0,299) et pour l'interaction entre le pays et le genre ($F(1,35) = 18,978, p < 0,001$, éta-carré-partiel = 0,350). Les tests post-hoc avec correction de Bonferroni (annexe B.15) nous montrent une différence significative entre les étendues fréquentielles des niveaux d'intensité faibles et élevés ($p = 0,006$).

Tableau 39 : Test des effets intra-participants pour l'analyse des distances entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé

		F(ddl, erreur ddl)	P	Eta-carré partiel
Greenhouse-Geisser	Intensité estimée	F(1.55,54.38)= 7,331	0,003	0,173
	Intensité estimée X Pays	F(1.55,54.38)=5,164	0,014	0,129
	Intensité estimée X genre	F(1.55,54.38)=2,994	0,122	0,062
	Intensité estimée X pays X genre	F(1.55,54.38)=2,602	0,096	0,069
	Set	F(1,35)= 2,802	0,103	0,074
	Set X Pays	F(1,35)=2,078	0,158	0,056
	Set X Genre	F(1,35)=0,360	0,553	0,010
	Set X pays X genre	F(1,35)=2,537	0,120	0,068
	Intensité estimée X set	F(1.87,65.37)=4,806	0,013	0,121
	Intensité estimée X set X pays	F(1.87,65.37)=0,984	0,374	0,027
	Intensité estimée X set X genre	F(1.87,65.37)=0,825	0,436	0,023
	Intensité estimée X set X pays X genre	F(1.87,65.37)=0,150	0,848	0,004

Tableau 40 : Test des effets inter-participants pour l'analyse des distances entre la F0 maximale des échantillons vocaux émotionnels et la F0 maximale des échantillons conversationnels en fonction du niveau d'intensité émotionnelle estimé

	F(ddl)	P	Eta-carré partiel
Pays	F(1,35)=0,900	0,349	0,025
Genre	F(1,35)=14,901	<0,001	0,299
Pays X Genre	F(1,35)=18,878	<0,001	0,350

3. Réponse à l'hypothèse 2

Nous voulions vérifier que les étendues fréquentielles des vocalisations spontanées étaient plus petites que celles de vocalisations imposées mais nos résultats ne permettent pas de confirmer cette hypothèse. En effet, nous relevons des étendues fréquentielles maximales et sectionnées, statistiquement plus grandes pour le set spontané que pour le set imposé (tableaux 23 et 27). Cette différence est significative pour les étendues sectionnées entre les niveaux d'intensité faibles et moyens et les niveaux d'intensité faibles et élevés. De plus, pour le set imposé, nous remarquons (figure 17) que plus l'intensité augmente, plus les étendues fréquentielles se déplacent vers un registre plus aigu de l'étendue conversationnelle. En effet, plus l'intensité augmente, plus les quantiles 25 des vocalisations émotionnelles s'éloignent des quantiles 25 des échantillons conversationnels et plus les fréquences maximales émotionnelles se rapprochent des fréquences maximales conversationnelles.

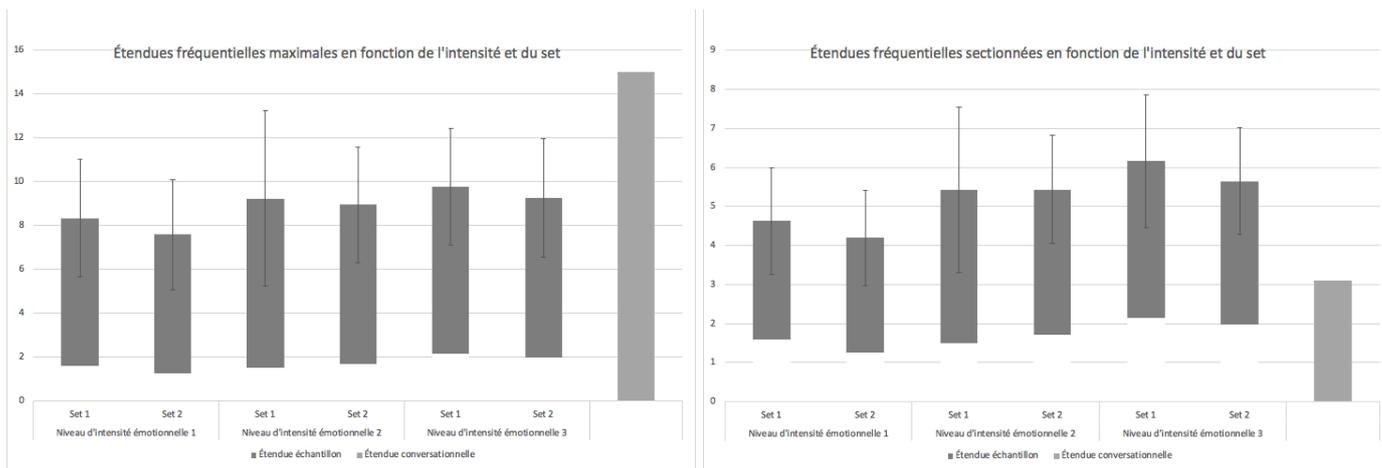


Figure 17. Comparaison des excursions fréquentielles maximales et sectionnées en fonction de l'intensité et du set.

G. Résultats statistiques au regard de l'hypothèse 3 :

« Les étendues fréquentielles des vocalisations émotionnelles se situeront au-delà de l'étendue fréquentielle conversationnelle des locuteurs »

1. Comparaison des étendues fréquentielles émotionnelles et des étendues fréquentielles conversationnelles

Pour répondre à cette dernière hypothèse, nous avons voulu comparer les étendues fréquentielles émotionnelles, sans distinction de niveaux d'intensité, aux étendues fréquentielles conversationnelles. Nous avons abordé cette comparaison de manière descriptive dans la figure 8, nous pouvons constater que lorsque nous utilisons les fréquences maximales, aucune vocalisation émotionnelle ne dépassait l'étendue fréquentielle conversationnelle. En revanche, en utilisant l'étendue sectionnée au quantile 75, la plupart des vocalisations des deux sets dépassaient l'étendue conversationnelle. Mais ces résultats sont-ils statistiquement significatifs ?

a) *Comparaison des étendues fréquentielles émotionnelles et des étendues fréquentielles conversationnelles maximales*

Nous avons effectué une Anova à mesures répétées. Les différentes conditions (conversationnelles et émotionnelles) ont été définies comme facteurs intra-participants et le genre et le pays comme facteurs inter-participants. Les étendues fréquentielles maximales ont

été définies comme variables dépendantes. Le test de sphéricité de Mauchly est non-significatif ($p=0,125$), nous pouvons employer les valeurs sans ajustement du degré de liberté.

Tableau 41 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale pour les vocalisations émotionnelles et conversationnelles

		Moyenne F0 Qt25-F0 max Conversationnelle (ET)	Moyenne F0 Qt25-F0 max Émotionnelle (ET)
F	Qc	13,51 (2,56)	7,76 (2,62)
	Be	15,50 (2,30)	7,15 (2,48)
H	Qc	22,75 (4,65)	6,26 (2,61)
	Be	14,05 (3,66)	5,80 (2,52)
Total F		14,27 (2,61)	7,53 (2,57)
Total H		16,54 (5,56)	5,93 (2,52)
Total		15,06 (3,98)	6,97 (2,65)

Nous observons (tableau 42) un effet intra-participants principal de la condition d'expérience ($F(1,116) = 707,493$, $p < 0,001$, éta-carré partiel = 0,859), de l'interaction entre la condition et le pays ($F(1,116) = 14,918$, $p < 0,001$, éta-carré partiel = 0,114), ainsi qu'entre la condition et le genre ($F(1,116) = 53,314$, $p < 0,001$, éta-carré partiel = 0,315). L'interaction entre la condition, le genre et le pays est également significative ($F(1,116) = 55,139$, $p < 0,001$, éta-carré partiel = 0,322). Nous observons (tableau 43) également un effet du facteur inter-participants du pays ($F(1,116) = 20,015$, $p < 0,001$, éta-carré partiel = 0,147), du facteur genre ($F(1,116) = 8,097$, $p = 0,005$, éta-carré partiel = 0,065), et de l'effet d'interaction entre le pays et le genre ($F(1,116) = 36,692$, $p < 0,001$, éta-carré partiel = 0,240). Lorsque l'on tient compte de l'étendue maximale, aucune vocalisation ne dépasse l'étendue conversationnelle. La population belge présente des étendues fréquentielles significativement plus petites que la population québécoise. Dans le cadre de l'expérience, les femmes présentent des étendues fréquentielles plus importantes que les hommes mais pas dans le cas de la tâche conversationnelle.

Tableau 42 : Test des effets intra-participants pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale pour les vocalisations émotionnelles et conversationnelles

		F(ddl)	P	Eta-carré partiel
Lambda de Wilks	Condition	$F(1,116) = 707,493$	<0,001	0,859
	Condition X Pays	$F(1,116) = 14,918$	<0,001	0,114
	Condition X genre	$F(1,116) = 53,314$	<0,001	0,315
	Condition X pays X genre	$F(1,116) = 55,139$	<0,001	0,322

Condition = tâche conversationnelle ou set 1 ou set 2

Tableau 43 : Test des effets inter-participants pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de fréquence maximale pour les vocalisations émotionnelles et conversationnelles

	F(ddl)	P	Eta-carré partiel
Pays	$F(1,116) = 20,015$	<0,001	0,147
Genre	$F(1,116) = 8,097$	0,005	0,065
Pays X Genre	$F(1,116) = 36,692$	<0,001	0,240

b) **Comparaison des étendues fréquentielles émotionnelles et des étendues fréquentielles conversationnelles sectionnées**

Nous avons effectué une Anova à mesures répétées. Les différentes conditions (conversationnelles et émotionnelles) ont été définies comme facteurs intra-participants et le genre et le pays comme facteurs inter-participants. Les étendues fréquentielles maximales ont été définies comme variables dépendantes. Le test de sphéricité de Mauchly est non-significatif ($p=0,248$), nous pouvons employer les valeurs sans ajustement du degré de liberté.

Tableau 44 : Étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et de quantile 75 pour les vocalisations émotionnelles et conversationnelles

		Moyenne F0 Qt25-F0 Qt75 Conversationnelle (ET)	Moyenne F0 Qt25-F0 Qt75 Émotionnelle (ET)
F	Qc	2,83 (0,97)	3,63 (1,42)
	Be	3,43 (0,84)	3,77 (1,44)
H	Qc	2,35 (0,49)	2,99 (1,81)
	Be	3,48 (1,08)	3,04 (1,12)
Total F		3,08 (0,96)	3,68 (1,42)
Total H		3,16 (1,08)	3,02 (2,52)
Total		3,10 (0,99)	3,45 (1,42)

Nous observons (tableau 45) un effet intra-participants principal de la condition d'expérience ($F(1,116) = 4,087$, $p=0,046$, éta-carré partiel= 0,034) ainsi qu'entre la condition et le pays ($F(1,116) = 5,192$, $p=0,020$, éta-carré partiel= 0,043). Nous pouvons donc observer une différence d'étendue fréquentielle émotionnelle par rapport à l'étendue conversationnelle. Seuls les hommes belges ne présentent pas une plus grande étendue fréquentielle dans la deuxième condition. Nous n'observons aucun effet inter-participants.

Tableau 45 : Test multivarié pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et quantile 75 pour les vocalisations émotionnelles et conversationnelles

		F(ddl)	P	Eta-carré partiel
Lambda de Wilks	Condition	F(1,116)= 4,087	0,46	0,034
	Condition X Pays	F(1,116)= 5,192	0,025	0,043
	Condition X genre	F(1,116)= 2,117	0,148	0,018
	Condition X pays X genre	F(1,116)= 0,974	0,326	0,008

Condition = tâche conversationnelle ou set 1 ou set 2

Tableau 46 : Test des effets inter-participants pour l'analyse des étendues fréquentielles en demi-tons entre les moyennes de quantile 25 et quantile 75 pour les vocalisations émotionnelles et conversationnelles

	F(ddl)	P	Eta-carré partiel
Pays	F(1,116)= 6,059	0,240	0,050
Genre	F(1,116)= 5,860	0,478	0,048
Pays X Genre	F(1,116)= 0,376	0,541	0,003

2. Réponse à l'hypothèse 3

Nous souhaitons vérifier que les étendues fréquentielles des vocalisations émotionnelles se situaient au-delà des étendues fréquentielles conversationnelles. Lorsque l'on tient compte de l'étendue maximale, aucune vocalisation ne dépasse l'étendue conversationnelle. Mais, comme expliqué précédemment, nous considérons que l'étendue maximale n'est pas très fiable étant donné la présence de fréquences aberrantes dans les enregistrements. De ce fait, lorsque nous nous penchons sur l'étendue fréquentielle sectionnée au quantile 75, nous observons effectivement une différence significative entre les étendues émotionnelles et conversationnelles. Cette différence est visible à la figure 18, les étendues ne sont pas beaucoup plus grandes en termes de demi-tons mais l'excursion vers les fréquences aiguës est plus importante.

En conclusion, nos résultats soutiennent notre hypothèse, les étendues fréquentielles émotionnelles dépassent les étendues conversationnelles, lorsque l'on exclut les valeurs extrêmes des fréquences afin de garder les cinquante pourcents de fréquences les plus souvent présentes.

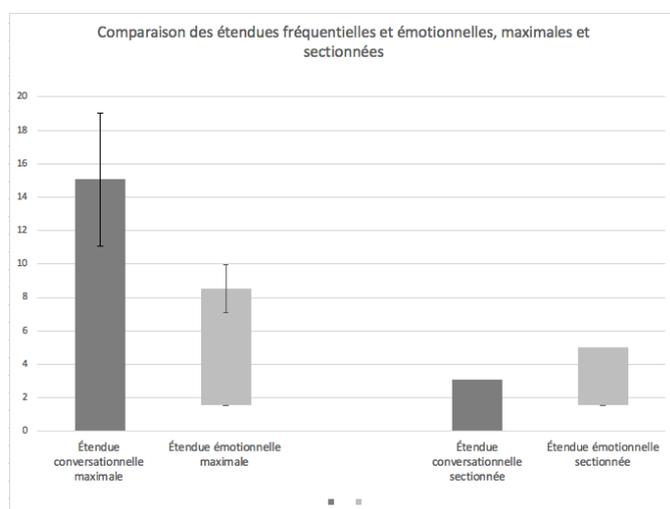


Figure 18. Comparaison des étendues fréquentielles conversationnelles et émotionnelles sans distinction d'intensité

IX. Discussion

Le but de cette recherche est d'explorer les variations fréquentielles des participants en fonction de l'intensité émotionnelle du stimulus visuel présenté. Dans cette section, nous discuterons des résultats obtenus qui nous permettent de soutenir ou non nos hypothèses.

A. Analyses des statistiques descriptives et choix méthodologiques

1. Description de la population

a) *Origine de la population*

Dans un premier temps, il semble opportun d'aborder les difficultés liées à notre échantillon. Nous n'avons pas contrôlé les pays d'origine ou de résidence de nos participants. En ce qui concerne les participants belges, nous savons qu'ils sont tous francophones à l'exception d'une participante dont la langue maternelle est l'espagnol mais qui est parfaitement bilingue. En revanche, au Québec, les participants peuvent être originaires de différents pays et parler différentes langues. Nous partons malgré tout du principe qu'une comparaison interculturelle est possible étant donné que nous avons récolté les données de 20 participants belges et de 20 participants québécois.

b) *Effectif de la population*

Si en Belgique le nombre de femmes et d'hommes a été contrôlé et respecté, c'est-à-dire 10 participants de chaque genre, au Québec les participants n'ont pas été sélectionnés sur ce critère-là. Ceci aurait pu nous poser des difficultés quant à la comparaison inter-genres des résultats mais avons pu observer que les données concordaient entre les hommes belges et québécois malgré la taille de notre effectif.

2. Étendues fréquentielles des participants

a) *Le quantile 25*

Il a été décidé d'employer le quantile 25 de l'étendue fréquentielle comme valeur minimale car, chez tous les participants, la fréquence minimale avoisinait les 70 Hz. Cette valeur s'explique par l'utilisation du mécanisme Fry par les participants mais nous avons décidé de ne pas considérer ces valeurs comme des fréquences minimales. Il a été démontré que les valeurs fréquentielles minimales des femmes se situaient généralement en-dessous de 100 Hz

à cause de l'utilisation de ce mécanisme. C'est une particularité vocale qui peut se développer, notamment chez les jeunes femmes de différentes origines, par le biais de l'influence sociale des personnes connues qui emploient ce mécanisme dans leur discours (Oliveira, G., Davidson, A., Holczer, R., Kaplan, S. et Paretzky, A., 2016). De plus, si l'on tenait compte de ces valeurs avoisinant les 70 Hz, on se situerait au niveau de la fréquence minimale physiologique des hommes et bien en-deçà de celle des femmes (Sanchez, K., Oates, J., Dacalis, G. et Holmberg, E.B., 2014). Il nous a donc semblé plus judicieux d'employer le quantile 25 de ces données, qui reflète des valeurs fréquentielles plus proches de la réalité.

b) Le quantile 75

Nous hésitions entre l'utilisation du quantile 75 ou de la fréquence maximale de la tâche conversationnelle. Malgré la grande similitude entre la fréquence fondamentale moyenne et le quantile 75, nous avons observé qu'un grand nombre de fréquences maximales provenait en fait de sons, tels qu'un rire ou une inspiration bruyante, ne faisant pas réellement partie de la voix du participant. De plus, dans la plupart des langues, l'étendue fréquentielle conversationnelle employée par les locuteurs est généralement située dans la partie inférieure de leur étendue fréquentielle car l'utilisation des fréquences les plus élevées demande un effort plus important comme lors d'un cri (Traunmüller, H. et Eriksson, A., 1995).

c) Étendue fréquentielle conversationnelle

Le fait que les fréquences les plus élevées soient peu utilisées (Traunmüller, H. et Eriksson, A., 1995) nous a également poussé à ne pas employer l'étendue fréquentielle physiologique pour la comparaison avec les étendues fréquentielles émotionnelles. Nous pensons que l'étendue fréquentielle physiologique n'est pas une tâche à recommander lors d'études de ce genre, elle ne nous apporte pas d'informations valides étant donné les difficultés rencontrées par certains participants lors de la passation. En effet, on peut observer que l'étendue conversationnelle est parfois plus large que l'étendue absolue. Si l'on compare ces résultats à ceux de la littérature, les femmes devraient monter, en moyenne, à une fréquence maximale avoisinant les 880 Hz (Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., 2002), ce qui n'est pas le cas de notre population. Nous expliquons cela par l'effet du stress imposé par l'expérience. En effet, lors d'une consultation médicale ou logopédique, les patients viennent à la recherche d'une solution ou de réponses à leurs questions et font donc confiance à la personne qui se trouve en face d'eux. Dans ce cas-ci, les participants n'ont pour

la plupart jamais exploré leurs capacités vocales et se retrouvent à être le centre de l'attention d'une seule personne qui ne revêt pas cette image médicale. La population belge présente une plus grande étendue physiologique qui est expliquée par les différences de promiscuité entre les participants et l'expérimentateur. En effet, les participants belges me connaissaient pour la plupart d'entre eux, ce qui a pu favoriser un climat de confiance. Pour la suite de nos analyses, nous avons donc décidé de n'employer que les étendues vocales conversationnelles de nos participants en raison de l'invalidité des étendues fréquentielles physiologiques. Notons tout de même que la tâche conversationnelle qui consiste à raconter une recette de cuisine, n'est pas si neutre qu'elle n'y paraît. En effet, cette tâche peut provoquer des émotions chez certains de nos participants. Par exemple, l'évocation d'une recette de grand-mère rendant le participant nostalgique, la recette préférée d'un proche décédé ou une recette qu'il aime particulièrement. Ces émotions pourraient biaiser nos données d'étendues fréquentielles mais cet élément n'a pas été pris en compte par notre étude.

3. Conversion des fréquences hertziennes en demi-tons

Afin de pouvoir comparer de manières équivalentes les différentes étendues fréquentielles de nos participants, nous avons transformé les différentes valeurs fréquentielles relevées en demi-tons. Il semble en effet plus approprié d'employer une transformation logarithmique de la fréquence, nommée intervalle de hauteur et qui se mesure en cents pour comparer des étendues fréquentielles (Matteson, S.E., Streit Olness, G., Caplow, N.J., 2013). Étant donné qu'un cent correspond à un centième de demi-ton, nous pouvons effectuer cette transformation par l'utilisation de la formule $\frac{Hz_2}{Hz_1} (\text{Log}_2) \times 12$. Grâce à celle-ci, nous multiplions par 12 le logarithme en base 2 de l'équation de la fréquence maximale divisée par la fréquence minimale, ce qui nous donne une étendue en demi-tons, plus facilement visualisable qu'une étendue en cents. Avec les étendues fréquentielles en demi-tons, toutes les valeurs fréquentielles du quantile 25, de chaque participant, seront équivalentes à « 0 ». De cette manière, tous les participants ont le même point de départ et toutes les données seront équivalentes. Grâce à cette conversion, et pour appuyer notre décision d'employer le quantile 75, nous avons pu relever un ambitus de presque 2 octaves pour les hommes en conversationnel lorsque nous avons utilisé la fréquence maximale comme limite. Cette valeur nous semblait démesurée et ce n'était pas la seule valeur aberrante relevée lors des étendues conversationnelles et émotionnelles. Ces valeurs correspondent généralement à des sources

extérieures comme une inspiration bruyante ou un bruit réalisé avec le matériel. En conséquence, l'utilisation de l'étendue sectionnée semble toujours la plus adaptée.

4. Nombre de vocalisations émotionnelles

Au niveau du recueil des vocalisations émotionnelles, nous avons remarqué (tableau 6) que nous avons récolté plus de vocalisations émotionnelles pour le premier set chez les femmes belges. Ceci est interpellant étant donné que lors du second set, les participants étaient obligés de vocaliser. Dans les faits, lors de l'expérience, des participants belges féminins ont déclaré avoir passé certaines images de manière involontaire, en appuyant sur la barre espace trop rapidement. Malheureusement, aucune manipulation ne nous permettait de revenir en arrière. Il aurait fallu recommencer le set mais cela aurait complètement biaisé nos résultats.

5. Intensité émotionnelle cotée par le participant

Lors de l'analyse des données concernant les vocalisations émotionnelles, nous avons été étonnées de voir que les stimuli de forte intensité ne provoquaient pas toujours de réactions vocales. En allant observer les évaluations d'intensité remplies par les participants lors de l'expérience, nous nous sommes rendues compte que tous les stimuli n'étaient pas cotés de la même manière. En effet, tous les participants ne réagissaient pas de la même manière face au même stimulus. Par exemple, certains participants étaient très surpris et apeurés face à l'image représentant une personne blessée, tandis que d'autres, faisant partie du corps médical ou de la police, n'étaient pas choqués. Les valeurs d'intensité émotionnelle sont donc très personnelles et nous avons décidé de rassembler et d'employer les valeurs d'intensité cotées par le participant afin de rechercher une différence significative entre les différentes valeurs d'intensité. Ces valeurs, qui allaient de 10 à 110 sur l'échelle analogique, seront divisées en trois groupes (faible, moyenne et élevée) et nous ne nous baserons plus que sur ces valeurs. Malheureusement, ces analyses ne pourront porter que sur le set 2. En effet, nous nous sommes rendues compte que les données d'intensité pour le set 1 ne s'étaient pas enregistrées lors de l'expérience à cause d'une erreur de script.

B. Discussion des résultats et des hypothèses

1. Discussion au regard des résultats de l'hypothèse 1

Notre première hypothèse énonçait un dépassement plus important de l'étendue fréquentielle pour les émotions à forte intensité en comparaison des émotions à faible intensité.

a) Analyses sur le set imposé

Au départ, il était prévu de réaliser nos analyses sur les deux premiers sets de l'expérience mais il y a eu tellement de non-réponses pour le set spontané (1) que nous avons décidé de n'employer que le set imposé (2). Ces résultats sont donc à prendre avec vigilance étant donné que nous nous trouvons dans une situation où l'on force une personne à vocaliser sur des images de valence et d'intensité différentes. La variation des cotations d'intensité nous avait poussé à ne plus utiliser les valeurs de cotation de l'IAPS. En effet, nos analyses auraient été biaisées puisque chaque participant porte un regard et une interprétation émotionnelle différente pour chaque stimulus. Par exemple, une personne du corps médical ne sera pas dérangée par la vue du sang, contrairement à un participant extérieur à ce milieu.

b) Résultats : validation de l'hypothèse

Grâce à ces analyses, nous avons pu vérifier et confirmer notre hypothèse. En effet, si l'on s'en tient à l'emploi des étendues fréquentielles sectionnées, ce qui nous semble plus logique, on peut observer un dépassement significatif de l'étendue conversationnelle par les trois niveaux d'intensité. Et ce dépassement augmente effectivement avec l'augmentation de l'intensité émotionnelle même si cela n'est significatif qu'entre les niveaux faibles et moyens, et faibles et élevés. Plus l'intensité émotionnelle est élevée, plus les fréquences des participants se déplacent vers le registre fréquentiel aigu. Ce lien peut s'expliquer par la relation existant entre les déclarations émotionnelles de valence positive et les émotions d'intensité forte, qui entraîne une gamme fréquentielle plus élevée que d'habitude (Cruttenden, A., 1986, cité par Reckling, F. et Kügler, F., 2011). En effet, il a été démontré que la fréquence fondamentale maximum des émotions positives est en moyenne plus élevée que celle des émotions négatives. Par contre, la fréquence minimale des énoncés de valence négative est en moyenne plus élevée que celle de valence positive. Les gammes de fréquences de valence positive et négative sont donc différentes en terme de hauteur tonale mais également d'étendue fréquentielle. Les énoncés de valence négative sont définis par une étendue plus petite que ceux de valence

positive (Reckling, F. et Kügler, F., 2011). Nous pouvons rapprocher nos résultats à ceux-ci étant donné les corrélations qui ont été établies entre les scores de valence et d'intensité pour notre seconde hypothèse. Il aurait été intéressant d'investiguer les variations au niveau de la valence émotionnelle pour pouvoir établir des liens solides entre ces deux versants de l'émotion mais ceci n'a pu être réalisé par manque de temps.

Nos résultats vont également dans le sens d'autres recherches qui soutiennent que la gamme fréquentielle est ce qui permet le mieux de détecter les émotions de son interlocuteur. En effet, une gamme fréquentielle étroite représente une émotion d'intensité faible tandis qu'une gamme fréquentielle large correspond à une émotion d'intensité élevée, pouvant être de valence négative ou positive (Scherer, K.R., 1995). De ce fait, la colère et la joie sont toutes les deux représentées par une hausse de la fréquence fondamentale (Russel, J.A., Bachorowski, J.A. et Fernandez-Dols, J.M., 2003) et par une étendue fréquentielle plus large. Pourtant, ces deux traits émotionnels sont opposés au niveau de la valence, c'est une preuve supplémentaire des liens significatifs existants entre la valence et l'intensité, et un soutien supplémentaire à nos résultats qui stipulent que l'étendue fréquentielle augmente avec l'intensité.

c) Limites

Jusqu'ici, l'étude de la fréquence fondamentale nous semble donc être un bon outil pour identifier l'intensité émotionnelle. En revanche, cela ne semble pas suffisant pour distinguer les émotions particulières les unes des autres. Les recherches précédentes ont investigué les différents facteurs qui permettraient de distinguer les émotions telles que la joie ou la surprise (Scherer, K.R., 2003). Cependant, rien ne nous permet encore actuellement une distinction précise. Les résultats de ce mémoire permettent de distinguer des catégories d'intensité émotionnelle contribuant à l'élaboration d'un profil acoustique émotionnel mais des analyses supplémentaires, joignant la valence à l'intensité émotionnelle, permettraient de distinguer l'ensemble des catégories émotionnelles.

d) Différences d'étendues fréquentielles entre les niveaux d'intensité cotée par le participant

En ce qui concerne la non-significativité des différences d'étendues fréquentielles entre les niveaux d'intensité moyenne et élevée, nous trouvons cela étonnant d'observer une hausse si importante d'étendue pour l'intensité moyenne alors que ce n'est pas si marquant lorsque l'on passe au niveau d'intensité élevé. On retrouve même des étendues plus importantes pour

le niveau d'intensité moyenne par rapport au niveau élevé (tableau 9). On peut expliquer cela par une sorte de seuil fréquentiel que les participants atteignent rapidement lorsque l'intensité augmente. Une autre explication serait que la manière de coter l'intensité des stimuli n'est pas la plus adaptée. En effet, l'échelle analogique offre une grande liberté aux participants mais celui-ci aurait peut-être toujours tendance à coter vers la médiane de l'échelle, en prévision d'un stimulus beaucoup plus intense que celui qui vient d'être visualisé. Les différentes intensités relevées lors des expériences se trouvent en annexes (annexe B.16).

e) ***Différences inter-genres***

Nous avons également pu constater que les hommes et les femmes ne se comportaient pas de la même manière quant à la distance qui séparait la fréquence du quantile 25 des vocalisations émotionnelles de la fréquence du quantile 25 de l'étendue conversationnelle. En effet, plus l'intensité émotionnelle augmente, plus la distance entre ces fréquences augmente chez les femmes. En revanche, chez les hommes, il y a une augmentation pour l'intensité moyenne mais peu de différences entre cette distance et celle séparant l'intensité faible de l'intensité élevée. Pourtant, ils se comportent de manière similaire en ce qui concerne la distance séparant les fréquences maximales émotionnelles et conversationnelles. On peut expliquer cela par le fait que les hommes ont tendance à plus intérioriser leurs émotions, contrairement aux femmes qui sont plus expressives (Buck, R., 1977). Les hommes effectuent une excursion moins importante vers le registre aigu que les femmes.

2. Discussion au regard des résultats de l'hypothèse 2

En ce qui concerne la seconde hypothèse, qui stipulait que les étendues fréquentielles des vocalisations spontanées seraient plus petites que celles des vocalisations imposées, nous avons rencontré d'importantes difficultés méthodologiques.

a) ***Estimation des intensités émotionnelles à l'aide de la valence cotée par le participant***

Nous voulions comparer les deux sets à l'aide des différents niveaux d'intensité étant donné que nous avons pu vérifier dans notre première hypothèse que ce facteur était concluant. Malheureusement, les données d'intensité cotées par le participant n'ont pas été enregistrées pour le set spontané pour les deux populations. Les analyses réalisées afin de trouver une valeur nous permettant d'estimer l'intensité des stimuli n'auraient jamais été réalisées dans le cadre

d'un article scientifique. Cependant, nous avons essayé de pallier la perte des données pour avoir une idée du résultat. Pour cela, il était préférable de nous baser sur les scores de valence émotionnelle cotés par le participant. De ce fait, il est important de souligner que les niveaux d'intensité estimés pour l'hypothèse 2 sont différents des niveaux d'intensité de l'hypothèse 1 qui reprenaient les valeurs réelles des participants. Les scores utilisés sont en fait des scores d'intensité extraits des scores de valence cotés par le participant.

Un autre élément important à prendre en compte est l'importante quantité de non-réponses au set 1. En effet, étant donné que les participants n'étaient pas obligés de vocaliser sur le set spontané, énormément d'images sont restées sans réaction. Nous avons remarqué que plus l'intensité cotée par le participant est élevée, plus le taux de réponses est important tandis que les images de faible intensité restaient sans réponse. Par exemple, un participant avait répondu à 3 items du set 1 et avait coté leurs intensités à 100, 90 et 80, tandis qu'il avait répondu aux 45 items du set 2 mais les avait évalués à 90, 60, 40, 30 ou 10. Cela n'avait donc pas de sens de calculer des niveaux d'intensité à partir d'items auxquels les participants n'avaient pas répondu. Si nous avons décidé d'employer les niveaux d'intensité cotés par le participant pour le set 2, presque toutes les réponses vocales du set 1 se seraient retrouvées dans le niveau d'intensité élevée du set 2 car les 3 niveaux d'intensité qui se seraient établis pour le set 1 auraient tous contenu des images à forte intensité. Il nous a donc semblé plus judicieux de transformer les niveaux d'intensité afin de pouvoir regarder les sets 1 et 2 de manière combinée. Cela nous a permis de catégoriser les stimuli dans les bons niveaux d'intensité avec de nouvelles valeurs estimées pour le set spontané et également pour le set imposé. Par rapport à cette catégorisation, nous avons dû exclure un participant québécois car il n'avait répondu à aucun stimulus du set 1 et nous ne pouvions pas nous baser uniquement sur ses réponses au set 2 au risque de biaiser nos résultats.

b) Résultats : hypothèse non-vérifiée

Grâce aux estimations d'intensité émotionnelle, notre hypothèse n'est pas vérifiée. Il se trouve que les étendues fréquentielles des vocalisations spontanées sont généralement plus larges que les étendues fréquentielles des vocalisations imposées. Lorsque nous avons émis cette hypothèse, nous pensions que les participants auraient tendance à exagérer les vocalisations imposées. Cette idée était émise sur le principe des acteurs à qui il avait été demandé, dans des études antérieures, de produire des énoncés de contenus neutres mais avec une intonation émotionnelle particulière (Scherer, K.R., 1995). En effet, dans le cas où l'on

impose ce type de tâches, les émotions produites par des comédiens sont généralement exprimées de manière extravertie. Cependant, notre étude repose sur l'expression de personnes non-comédiennes, elles n'ont donc pas l'habitude de s'exprimer pour un public et d'exagérer leurs émotions. Une autre explication pourrait être le fait que lors du set spontané, les participants ne réagissaient qu'aux images de vives intensités, leurs réactions étaient donc plus importantes en moyenne pour ce set, contrairement au set imposé pour lequel les participants ont réagi à des images d'intensités émotionnelles diverses. En effet, la communication vocale et faciale jouant un rôle central dans les interactions sociales (Campanella, S. et Belin, P., 2007), le but des participants était de faire passer rapidement un message émotionnel face à une image qui les a réellement interpellée. La différence de taux de réponses étant très importante entre les deux sets, les participants ont effectivement réagi aux images qui les faisaient vraiment réagir, ce sont souvent des réactions de surprise, de dégoût ou de bonheur si l'image est très positive. La consigne donnée était de réagir naturellement et spontanément sans se forcer, les images sans réponse lors du set spontané ont été passées car elles ne provoquaient pas d'émotions aux participants. Même si nous avons équilibré les niveaux d'intensité, les images de très faible intensité du set 2 entraînent une diminution des étendues fréquentielles moyennes.

De plus, un autre élément pouvant expliquer cette différence d'étendue fréquentielle est un phénomène de lassitude et d'habituation. En effet, le set imposé comprend 45 images auxquelles le participant est obligé de réagir vocalement. Cette situation peut vite ennuyer les participants qui réagissent peut-être de manière moins importante au fur et à mesure de l'expérience, contrairement au set spontané où les participants réagissent quand ils le souhaitent. On peut également ajouter que le set spontané est le premier de l'expérience, les participants sont encore dans la découverte alors que pour le set imposé, l'expérience est présentée sous la même forme exceptée en ce qui concerne la consigne. Malheureusement, ce point n'a pas été investigué.

c) *Distances inter-quantiles et étendues fréquentielles des vocalisations*

Lors de l'analyse de la distance entre les fréquences des quantiles 25 des vocalisations émotionnelles et les quantiles 25 des vocalisations conversationnelles, nous avons pu remarquer que celle-ci diminue entre l'intensité faible et moyenne pour le set 1 (tableau 29). C'est étonnant car cela ne va pas dans le sens de nos résultats, nous ne nous attendions pas à obtenir une baisse de cette distance alors que l'intensité augmente. Pourtant, lorsque l'on observe les résultats d'étendues fréquentielles sectionnées (tableau 25), on peut effectivement observer des étendues

plus larges pour le set 1. Nous pensons donc que les étendues sont plus larges malgré une excursion moins importante et que les participants se dirigent tout de même vers leur registre aigu.

Cette excursion vers le registre aigu de l'étendue fréquentielle des participants se remarque également par la diminution des distances entre les fréquences maximales des étendues fréquentielles émotionnelles et conversationnelles ainsi qu'entre les fréquences des quantiles 75 de ces deux étendues. C'est encore une preuve qu'enlever les 25 % de fréquences les plus hautes ne nous empêche pas de situer l'intensité émotionnelle par rapport à l'étendue conversationnelle. En effet lorsque l'on tient compte de l'étendue maximale, les fréquences maximales des vocalisations émotionnelles se rapprochent de la fréquence maximale conversationnelle. Dans le cas de l'étendue sectionnée, on peut voir que les fréquences les plus hautes dépassent de plus en plus la fréquence la plus haute de l'étendue conversationnelle. En plus, comme nous éliminons les fréquences aberrantes, les représentations des étendues sectionnées sont plus cohérentes avec les étendues fréquentielles conversationnelles naturelles.

Finalement, au niveau du set imposé, nous pouvons dire que, comme nous leur avons imposé de vocaliser, ils ont atteint plus facilement des registres aigus pour tous les niveaux mais ces étendues sont moins amples que celles spontanées car elles sont moins naturelles.

3. Discussion au regard des résultats de l'hypothèse 3

Notre dernière hypothèse visait à vérifier que les étendues fréquentielles des vocalisations émotionnelles se situaient au-delà de l'étendue fréquentielle conversationnelle des participants.

a) Résultats : validation de l'hypothèse

Grâce à nos analyses, nous avons pu valider cette hypothèse. Si l'on tient compte des étendues sectionnées de nos échantillons, on peut observer que les étendues fréquentielles émotionnelles se situent au-delà des étendues fréquentielles conversationnelles. Les étendues fréquentielles ne sont pas plus larges en termes de demi-tons mais on observe une nette excursion vers les fréquences aiguës des locuteurs. Comme pour la première hypothèse, si l'on s'en tient aux étendues fréquentielles maximales, les étendues émotionnelles ne dépassent pas les étendues conversationnelles, mais cela n'est pas une représentation réaliste du comportement vocal de la population. Ces résultats peuvent être expliqués par le fait qu'exprimer ses émotions de manière authentique et naturelle est utile au maintien de bonnes

relations sociales. Au contraire, si l'on adopte une attitude de restriction émotionnelle, en essayant de les intérioriser au maximum, cela peut conduire à un amoindrissement des relations sociales (Gross, J.J., 2002 ; Butler, E.A., Egloff, B., Wilhelm, F.H. et Gross, J.J., 2003). De ce fait, intérioriser des émotions telles que la colère entraîne un moins bon fonctionnement social (Chervonsky, E. et Hunt, C., 2017). Cela dit, le fait d'exprimer souvent de la colère peut également entraîner des interactions sociales plus négatives parce que cela envoie une image menaçante et nuisible à l'interlocuteur. Il est donc nécessaire de réguler ses émotions, qu'elles soient intériorisées ou exprimées, afin d'obtenir un équilibre social sain (Arslan, C. 2009). Pour cela, les locuteurs doivent apprendre à exprimer de manière adaptée leurs différentes émotions, en exprimant librement des émotions positives ou en réprimant des émotions entraînant une situation agressive et menaçante pour l'interlocuteur (Chervonsky, E. et Hunt, C., 2017).

Nos résultats portent sur les émotions de manière générale mais s'inscrivent dans les connaissances de ce domaine. Les étendues fréquentielles ont été étudiées dans le cadre d'acteurs produisant des émotions types et il se trouve que pour l'ensemble de ces émotions, on peut observer une augmentation de l'étendue fréquentielle. Pour la colère, le bonheur et la peur, les étendues fréquentielles sont beaucoup plus larges. En revanche, pour le dégoût et la tristesse, les étendues sont légèrement plus étroites (Murray, I.R. et Arnott, J.L., 1992). Même si la valence émotionnelle n'a pas été investiguée dans notre étude, les corrélations établies entre la valence et l'intensité nous permettent d'affirmer que nos résultats vont dans le même sens mais nous ne pouvons pas vérifier qu'il y a une diminution de l'étendue fréquentielle pour le dégoût et la tristesse. Ce que l'on pourrait proposer, c'est que ces types d'émotions ne provoquent pas une forte intensité émotionnelle, ce qui expliquerait que l'excursion fréquentielle est amoindrie.

b) Différences inter-genres

Lors de nos analyses, nous avons pu observer que les femmes présentaient des étendues fréquentielles plus larges que les hommes lors des tâches expérimentales mais pas dans le cadre de la tâche d'étendue conversationnelle. Ceci peut être dû au fait que les femmes ont tendance à être plus expressives que les hommes (Brody L.R., Hall, J.A., 1993, cité par Chaplin, T.M., 2015). Les femmes ont tendance à plus exprimer leurs émotions (Buck, R., 1977), probablement à cause des normes et attentes sociales de la société environnante. D'ailleurs, il est plus accepté socialement qu'un homme intériorise ses émotions qu'une femme, cela correspond plus aux normes sociales (Chervonsky, E. et Hunt, C., 2017). Mais cette intériorisation émotionnelle pourrait jouer sur leurs relations sociales.

c) *Différences interculturelles*

Lorsque l'on observe les étendues fréquentielles émotionnelles maximales et sectionnées, la population québécoise présente des étendues fréquentielles plus larges que la population belge. Mais lorsque l'on se penche sur les étendues fréquentielles sectionnées, c'est la population belge qui présente de plus larges étendues fréquentielles. Nous pouvons expliquer cela par des excursions plus fréquentes vers le registre aigu pour les québécois.

C. Forces et faiblesses de cette recherche

1. Limites

Notre étude comporte quelques limites qui ont été évoquées dans les résultats et la discussion de ce travail.

a) *Population*

Nous pouvons commencer par les participants composant notre effectif. Il aurait été plus intéressant et significatif de travailler sur une population équilibrée au niveau du genre, avec un nombre équivalent d'hommes et de femmes dans chaque population.

Il aurait aussi été intéressant d'équilibrer la répartition de l'âge dans chaque groupe. En effet, la seule restriction était d'être majeur, sauf que nous savons que l'âge est un facteur qui modifie la voix. L'idéal aurait donc été de recruter des personnes d'âge divers de manière imposée par la méthodologie. Cela impliquerait de travailler sur un échantillon de taille beaucoup plus importante afin de vérifier de manière statistique les données que l'on trouverait.

Nous savons que les émotions subissent des variations interculturelles malgré une universalité présente pour les émotions de base. Il aurait été intéressant de contrôler les origines des participants afin de travailler uniquement sur une population francophone belge et une population francophone québécoise, en s'assurant que ces participants étaient suffisamment imprégnés de la culture du pays pour le représenter.

b) *Méthodologie*

La plus grande difficulté rencontrée lors de la réalisation de cette étude provient d'une erreur de script qui nous a privé de précieuses données d'intensité émotionnelle. Les analyses

réalisées ne paraissent donc pas scientifiquement correctes mais cela nous a permis de compenser cette perte. L'idéal serait donc de recommencer cette étude avec l'ensemble des données correctement enregistrées.

c) Mesures physiologiques

Il est dommage que nous n'ayons pas pu récolter des mesures physiologiques lors de ces expériences, étant donné que nous savons que les changements physiologiques affectent la voix. Par exemple, l'anxiété peut entraîner une augmentation des tensions musculaires du larynx et des plis vocaux, entraînant une augmentation de la hauteur tonale. Ce champ a déjà été investigué dans des études précédentes, mais elles travaillaient essentiellement sur des comédien.ne.s. Il existe donc peu de travaux réalisés sur un discours émotionnel plus authentique et naturel (Laukka, P., Linman, C., Ahs, F., Pissioti, A., Frans, Ö., Faria, V., Michelgard, A., Appel, L., Fredrikson, M. et Furmark, T., 2008).

2. Point positifs

Notre étude présente également des points positifs et novateurs.

a) Population non professionnelle de la voix

Le point le plus positif de cette recherche est qu'elle a été menée sur des participants qui n'étaient pas des professionnels de la voix. La plupart des études précédentes ont travaillé sur base de vocalisations produites par des acteurs. Généralement, les acteurs étaient invités à représenter des émotions dans le but qu'elles soient reconnues par des auditeurs. Il est évident que ces recherches nous ont apporté énormément de connaissances sur les expressions vocales mais elles comportent quelques limites et biais par l'aptitude des comédien.ne.s à gérer l'expression de leurs émotions, contrairement à des participants tout-venants qui n'ont pas appris à les maîtriser (Laukka, P., Linman, C., Ahs, F., Pissioti, A., Frans, Ö., Faria, V., Michelgard, A., Appel, L., Fredrikson, M. et Furmark, T., 2008). De plus, il est probable que certains aspects de la voix véhiculant des émotions ne soient pas produits de manière volontaire (Bachorowski, J.A. et Owren, M.J., 1995). Notre population était composée de participants aux personnalités variées, ce qui nous offre un point de vue assez proche de la population générale. En effet, nous avons pu entendre des participants plus ou moins expressifs ou sensibles face aux stimuli. Cette diversité de personnalité nous offre un champ d'investigation plus proche de la réalité.

b) *Étude internationale*

Notre étude avait également l'avantage et l'intérêt d'être menée sur deux populations très éloignées géographiquement. L'universalité des émotions ayant été très étudiée dans le domaine psychologique, c'est une manière de prolonger ces recherches en confirmant qu'il y a effectivement un caractère international à l'expression émotionnelle mais également des variations interculturelles. De ce fait, peu importe l'origine des participants, on observe une étendue fréquentielle émotionnelle qui augmente avec l'intensité émotionnelle. En revanche, on remarque plus de variations vocales dans la population québécoise qui présente une étendue émotionnelle plus large que la population belge alors que cette dernière présentait des étendues fréquentielles conversationnelles et physiologiques plus importantes.

X. Conclusions et perspectives

Dans quelles mesures les émotions affectent-elles notre voix ? Cette question est au cœur de la recherche depuis de nombreuses années. Les études précédentes nous apportent un flux d'informations très important quant aux paramètres qui varient avec l'état émotionnel d'une personne. Toutefois, peu d'études se sont basées sur une population de participants toutvenants. De plus, les émotions étaient généralement étudiées de manière globale et arbitraire car les acteurs étaient tenus de représenter une émotion de manière stéréotypée.

Notre étude inverse cette tendance en s'attardant sur l'intensité émotionnelle réellement ressentie par le participant et les modifications vocales qui en découlent. Nous ne traitons pas les différents stimuli comme des représentants d'émotions particulières, comme le bonheur ou la colère, mais plutôt comme des stimuli dont la réaction dépendra du vécu et de la personnalité des participants. Cette recherche est centrée sur les modifications fréquentielles émotionnelles en comparaison de l'étendue fréquentielle conversationnelle des participants. L'aspect psychologique de ces données sera exploité ultérieurement par d'autres études.

A ces fins, un dispositif expérimental plaçant les participants dans différentes conditions communicationnelles a été mis au point. Nous avons d'abord étudié les étendues fréquentielles conversationnelles et physiologiques de chaque population et il s'est révélé que nous devions nous baser sur l'étendue conversationnelle. Suite à cela, les participants ont été confrontés à 4 corpus d'images de valence et d'intensité émotionnelles différentes. Notre recherche était basée sur les corpus spontané et imposé afin d'investiguer les probables différences de paramètres vocaux entre ces conditions. Les participants devaient également coter chaque stimulus en termes de valence et d'intensité afin de réaliser un classement équilibré entre tous les stimuli.

En passant outre les difficultés méthodologiques, notre étude a permis d'investiguer ce domaine et de répondre à nos interrogations. Nous avons pu démontrer que les étendues fréquentielles émotionnelles effectuaient une ascension dépassant les étendues fréquentielles conversationnelles. Nous avons également trouvé que cette excursion vers la gamme de fréquences aiguës dépendait de l'intensité émotionnelle ressentie par le participant. En revanche, nos résultats concernant les différences d'étendues fréquentielles entre les conditions imposées et spontanées ne nous ont pas permis de valider notre hypothèse. En effet, il se trouve

que c'est lors de la condition spontanée que les étendues fréquentielles émotionnelles effectuent une plus grande excursion au travers la gamme de fréquences.

Nos résultats nous poussent à établir des perspectives d'avenir et d'amélioration pour ce champs de recherche.

Il serait intéressant, au même titre que pour l'intensité émotionnelle cotée par le participant, d'investiguer l'impact de la valence émotionnelle sur l'étendue fréquentielle de nos participants et d'établir des liens statistiques entre ces deux facteurs.

Il serait également enrichissant d'étudier l'évolution des étendues fréquentielles émotionnelles des participants au cours de l'expérience. Nous ferions peut-être face à un déclin de celles-ci suite à l'apparition d'une certaine lassitude des participants, comme énoncé précédemment.

Concernant la tâche d'étendue fréquentielle conversationnelle, il serait judicieux de mettre au point une tâche totalement neutre chez tous les participants, tout en restant dans une condition naturelle pour ne pas biaiser nos données.

En conclusion, cette étude ouvre un champ d'investigation intéressant dans le cadre de notre profession de logopèdes car nous savons que l'état psychologique peut affecter le fonctionnement vocal. En effet, un impact émotionnel répétitif sur l'appareil vocal peut entraîner des pathologies vocales couramment rencontrées. Il est de ce fait indispensable de tenir compte du poids de l'implication émotionnelle et psychologique lors de nos prises en charge.

XI. Bibliographie

- Allen, J.G., Haccoun, D.M. (1976). Sex differences in emotionality: A multidimensional approach. *Human Relations*, 29, 711–722. doi: 10.1177/001872677602900801
- Altrov, R. (2013). Aspects of cultural communication in recognizing emotions. *TRAMES*, 17 (67/62), 2, 159-174. doi: 10.3176/tr.2013.2.04
- Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology*, 8, 291–322. doi: 10.1037/1089-2680.8.4.291
- Arslan, C. (2009). Anger, self-esteem, and perceived social support in adolescence. *Social Behavior and Personality*, 37, 555-564. doi: 10.2224/sbp.2009.37.4.555
- Averill, J.R., Opton Jr, E.M., Lazarus, R.S. (1969). Cross-Cultural studies of psychophysiological responses during stress and emotion. *International Journal of Psychology*, 4 (2), 83-102. doi: 10.1080/00207596908247258
- Bandura, A. *Principles of behavior modification*. Holt, Rinehart & Winston; New York: 1969.
- Banse, R., Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70, 614–636. doi: 0022-3514/96/53.00
- Bänziger, T., Patel, S. et Scherer, K.R. (2014). The role of perceived voice and speech characteristics in vocal emotion communication. *Journal of Nonverbal behavior*, 38(1), 31-52. doi: 10.1007/s10919-013-0165-x
- Barsties, B., & Maryn, Y. (2015). The improvement of the internal consistency of the Acoustic Voice Quality Index. *American Journal of Otolaryngology - Head and Neck Medicine and Surgery*, 36(5), 647-656. doi: 10.1016/j.amjoto.2015.04.012
- Belyk, M., Brown, S. (2014). The acoustic correlates of valence depend on emotion family. *Journal of Voice*, 28 (4), 523. doi: 10.1121/1.5000482
- Blair, R.J.R. (2003). Facial expressions, their communicatory functions and neuro– cognitive substrates. *Philosophical Transactions of the Royal Society of London. Series B*, 358, 561–572. doi: 10.1098/rstb.2002.1220
- Brody, LR.; Hall, JA. *Gender and emotion*. In: Lewis, M.; Haviland, JM., editors. Handbook of emotions. Guilford Press; New York, NY: 1993. p. 447-460.
- Buck, R. (1977). Nonverbal communication of affect in preschool children: Relationships with personality and skin conductance. *Journal of Personality and Social Psychology*, 4, 225-236.

Butler, E.A., Egloff, B., Wilhelm, F.H. et Gross, J.J (2003) The Social Consequences of Expressive Suppression, *Emotion*, 3(1), 48-67. doi: 10.1037/1528-3542.3.1.48

Campanella, S. et Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, 11(12), 535 – 543. doi: 10.1016/j.tics.2007.10.001

Chaplin, T.M. (2015). Gender and Emotion Expression: A Developmental Contextual Perspective. *Emotion review: journal of the International Society for Research on Emotion*, 7(1), 14-21. doi: 10.1177/1754073914544408

Chaplin, T.M., Hong, K.A., Bergquist, K., Sinha, R., (2008). Gender differences in response to emotional stress: An assessment across subjective, behavioral, and physiological domains and relations to alcohol craving. *Alcoholism: Clinical and Experimental Research*, 32, 1242–1250. doi: 10.1111/j.1530-0277.2008.00679.x.

Chervonsky, E., et Hunt, C. (2017). Suppression and Expression of Emotion in Social and Interpersonal Outcomes: A Meta-Analysis. *Emotion*. Advance online publication. doi: 10.1037/emo0000270

Cohen, S.M., Kim, J., Roy, N., Asche, C. et Courey, M. (2012). Prevalence and causes of dysphonia in a large treatment-seeking population. *Laryngoscope*, 122(2), 343-348. doi : 10.1002/lary.22426.

Côté, G., Gosselin, P., Dagenais, I. (2013). Evaluation multidimensionnelle de la régulation des émotions : propriétés psychométriques d'une version francophone du Difficulties in Emotion Regulation Scale. *Journal de Thérapie Comportementale et Cognitive*, 23 (2), 63-72. doi : 10.1016/j.jtcc.2013.01.005

Coyle, S.M., Weinrich, B.D. et Stemple J.C. (2001). Shifts in Relative Prevalence of Laryngeal Pathology in a Treatment-Seeking Population. *Journal of Voice*, 15(3), 424-440. doi : 10.1016/S0892-1997(01)00043-1

Cuttenden, A. (1986) *Intonation*. Cambridge : CUP

Damasio, A.R. (2010). *L'erreur de Descartes: la raison des émotions*. Paris: Odile Jacob.

Dan-Glauser, E.S. et Scherer, K.R. (2013). The Difficulties in Emotion Regulation Scale (DERS) Factor structure and consistency of a french translation. *Swiss Journal of Psychology*, 72 (1), 5-11. doi : 10.1024/1421-0185/a000093

Devillers, L., Tahon, M., Sehili, M.A. & Delaborde, A. (2014). Détection des états affectifs lors d'interactions parlées- robustesse des indices non verbaux, *Traitement Automatique des Langues*, 55 (2), 123-149

- Dietrich, M., Verdolini Abbott, K., Gartner-Schmidt, J. et Rosen, C.A. (2008). The frequency of perceived stress, anxiety, and depression in patients with common pathologies affecting voice. *Journal of Voice*, 22(4), 472-488. doi : 10.1016/j.jvoice.2006.08.007
- Dietrich, M., Verdolini Abbott, K. (2012). Vocal function in introverts and extraverts during a psychological stress reactivity protocol. *Journal of Speech, Language and Hearing Research*; 55 (3), 973-987. doi : 10.1044/1092-4388(2011/10-0344).
- Dobres, R., Lee, L., Stemple, J.C., Kummer, A.W. et Kretschmer, L.W. (1990). Description of Laryngeal Pathologies in Children Evaluated by Otolaryngologists. *Journal of Speech and Hearing Disorders*, 55(3), 526-532. doi : 10.104/jshd.5503.526.
- Drevets, W.C. (2000). Neuroimaging studies of mood disorders. *Biological Psychiatry* 48, 813-829. doi : 10.1016/S0006-3223(00)01020-9.
- Ekman, P. (1992) Are there basic emotions? *Psychological Review*, 99 (3), 550-553. doi : 0033-295X/92/53.00.
- Frijda, N.H. (1988) The laws of emotion. *American Psychologist*. 43, 349–358. doi : 10.1037/0003-066X.43.5.349.
- Frühholz, S., Trost, W., & Grandjean, D. (2014). The role of the medial temporal limbic system in processing emotions in voice and music. *Progress in neurobiology*, 123, 1-17. doi : 10.1016/j.pneurobio.2014.09.003.
- Genter, A. (2014). *Definition and representation of user experience intentions in the early phase of the industrial design process: a focus on the kansei process* (Thèse) Product Design and Innovation Laboratory (LCPI), Arts et Metiers ParisTech.
- Ghio A. (2007). *L'évaluation acoustique*. Dans Auzou P.; Rolland V.; Pinto S. ; Ozsancak C. *Les dysarthries*, Solal, 2007.
- Giovanni, A., Robert, D., Estublier, N., Teston, B., Zanaret, M., Cannoni, M. (1996). Objective evaluation of dysphonia: preliminary results of a device allowing simultaneous acoustic and aerodynamic measurements. *Folia Phoniatica Logopaedia*, 48, 175- 185.
- Gramming, P., Sundberg, J., Terström, S., Leanderson, R., et Perkins, W.H. (1988). Relationship between changes in Voice Pitch and Loudness. *Journal of Voice*, 2 (2), 118-126. doi: 10.1016/j.jvoice.2015.03.006
- Gratz, K.L. et Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment*, 26, 41-54. doi: 10.1007/s10862-008-9102-4.

- Gross, J.J. (2002). Emotion regulation: Affective, cognitive, and social consequences, *Psychophysiology*, 39, 281-291. doi: 10.1017/S0048577201393198
- Guimarães, I. (2007). *A Ciência e a Arte da Voz Humana*. Escola Superior de Saúde de Alcoitão.
- Heuillet-Martin, G., Garson-Bavard, H. et Legré, A. (1995). *La voix pour tous*. Marseille : Solal.
- Heylen, L., Wuyts, F.L., De Bodt, M. and Van de Heyning, P.H., (2002). Normative Voice Range Profiles of Male and Female Professional Voice Users. *Journal of Voice*, 16 (1), 1-7. doi: 10.1016/S0892-1997(02)00065-6
- Hubbard, D.J., Assmann, P.F. et AA (2013). Perceptual adaptation to gender and expressive properties in speech: the role of fundamental frequency. *The journal of the acoustical society of America*, 133(4), 2367-2376
- Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review*, 91, 153–184. doi: 10.1037/0033-295X.91.2.153
- James, W., 1884. What is an Emotion? *Mind* 9, 188–205. doi: 10.1093/mind/os-IX.34.188
- Jang, D. and Elfenbein H. A. (2015). Emotion, Perception and Expression of. *International Encyclopedia of the Social & Behavioral Sciences*, 2nd edition, 7, 483-489. doi: 10.1016/B978-0-08-097086-8.25052-6
- John Laver: *The Phonetic Description of Voice Quality*, pages 93–140. *Cambridge University Press*, 1980.
- Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*, 1(4), 381- 412. doi: 10.1037/1528-3542.1.4.381
- Kim, J., Bouchard, C., Ryu, H., Omhover, J.-F., & Aoussat, A. (2012). Emotion finds a way to users from designers: assessing product images to convey designer's emotion. *Journal of Design Research*, Vol. 10, No. 4, 307-323.
- LaFrance, M., Hecht, M.A., Levy Paluck, E.L. (2003) The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin*. 129, 305–334. doi:10.1037/0033-2909.129.2.305.
- Lang, P.J., (1979). A bio-Informational Theory of Emotional Imagery. *Psychophysiology*, 16 (6), 495-512. doi: 10.1111/j.1469-8986.1979.tb01511.x

- Lardi, C., Billieux, J., d'Acremont, M., et Van der Linden, M., (2008). Adaptation of a short version of the Sensitivity to Punishment and Sentivity to Reward Questionnaire (SPSRQ). *Personality and Individual Differences*, 45, 722-725. doi:10.1016/j.paid.2008.07.019.
- Laukka, P., Linman, C., Ahs, F., Pissioti, A., Frans, Ö., Faria, V., Michelgard, A., Appel, L., Fredrikson, M. et Furmark, T. (2008). In a Nervous Voice: Acoustic Analysis and Perception of Anxiety in Social Phobics' Speech. *Journal of Nonverbal Behavior*, 32, 195-214. doi: 10.1007/s10919-008-0055-9.
- Laukka, P. et Elfenbein, H. A. (2012). Emotion appraisal dimensions can be inferred from vocal expressions. *Social Psychological and Personality Science*, 3(5), 529-536. doi : 10.1177/1948550611428011
- Le Bagousse, S. (2008). Etude Perceptive et acoustique du timbre de la voix chantée dans le contexte des répertoires de tradition orale. Mémoire de Master ATIAM, Université Pierre et Marie Curie Paris 6, Paris.
- Le Huche, F & Allali, A. (2010). *La voix. Anatomie et physiologie des organes de la voix et de la parole*. Tome 1. Collection: phoniatrie, Paris, Elsevier Masson.
- Lhote, E. (1982). *La parole et la voix: analyse et synthèse de faits de langue au niveau du larynx*. Hamburger phonetische Beiträge 37. Hamburg: Buske. 560 p.
- Liben, L.S., Bigler, R.S., (2002), The developmental course of gender differentiation : Conceptualizing, measuring and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, 6, serial No 269.
- Liebenthal, E., Silbersweig, D.A. & Stern, E. (2016). The language, tone and prosody of emotions: neural substrates and dynamics of spoken-word emotion perception. *Frontiers in neuroscience*, 10 (506) doi: 10.3389/fnins.2016.00506
- Livingstone, S. R., Choi, D. H. et Russo, F. A. (2014). The influence of vocal training and acting experience on measures of voice quality and emotional genuineness. *Frontiers in psychology*, 5. doi: 10.3389/fpsyg.2014.00156
- Luminet, O. (2007). *Psychologie des émotions. Confrontation et évitement*. Bruxelles, De Boeck Université, 3ème édition.
- Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J. et Frith, C.D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97, 4398-4403. doi; 10.1073/pnas.070039597
- Martin, C.L., Halverson, C.F.Jr., (1981) A schematic processing model of sex typing and stereotyping in Children. *Child Development*, 52, 1119-1134.

- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324. doi: 10.3758/s13428-011-0168-7
- Matteson, S.E., Streit Olness, G., Caplow, N.J., (2013). Toward a quantitative account of pitch distribution in spontaneous narrative: Method and validation, *Journal Acoustic Society of America*, 133 (5), 2953-2971. doi: 10.1121/1.4796111
- Méchain, C. (2014) Utilisation du triangle vocalique dans l'apprentissage de la voix oesophagienne. *Sciences Cognitives* <dumas-01081035>
- Menin-Sicard, A., Sicard, E. (2016). *Évaluation et réhabilitation de la voix : approche clinique et objective*. Louvain-la-Neuve, De Boeck-Supérieur, « Orthophonie ».
- Meunier, C. (2007). Phonétique acoustique. Dans Auzou, P., Rolland-Monnoury, V., Pinto, S. et Ozsancak, C. *Les dysarthries* (p.164-173). Paris, De Boeck-Solal, <hal-00250272>
- Morel, M., Bänziger, T., (2002). Le rôle de l'intonation dans la communication vocale des émotions : test par la synthèse. *Cahiers de l'Institut de Linguistique de Louvain, Peeters*, 30, 207-232 <hal-00100347>
- Morsomme, D. (2001). *Contribution à la détermination des paramètres subjectifs et objectifs pour l'étude de la voix*. Thèse de doctorat, Louvain-la-Neuve, Université Catholique de Louvain.
- Murray, I.R. et Arnott, J.L. (1992). Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion, *Journal Acoustic Society of America*, 93 (2), 1097-1108. doi: 10.1007/978-3-540-78293-3_5
- Oliveira, G., Davidson, A., Holczer, R., Kaplan, S. et Paretzky (2016). A comparison of the Use of Glottal Fry in the Spontaneous Speech of Young and Middle-Aged American Women. *Journal of Voice*, 30 (6), 684-687. doi: 10.1016/j.jvoice.2015.08.015
- Patel, S., Scherer, K. R., Sundberg, J. et Björkner, E. (2010). Acoustic markers of emotions based on voice physiology. In *Proceedings of the speech prosody* (100865), 1-4.
- Reckling, F. et Kügler, F. (2011). Pitch range in positive and negative connoted statements of german. *The 17th International Congress of Phonetic Sciences, Hong Kong 2011*
- Roy, N., Bless, D.M., (2000). Personality traits and psychological factors in voice pathology: a foundation for a future research. *Journal of Speech, Language and Hearing Research*, 43, 737-748. doi:10.1044/jslhr.4303.737

- Roy, N., Bless, D.M., Heisey, D. (2000). Personality and voice disorders: A multitrait-multidisorder analysis. *Journal of Voice*, 14(4), 521-548. doi: 10.1016/S0892-1997(00)80009-0
- Russel, J.A., Bachorowski, J.A. et Fernandez-Dols, J.M. (2003). Facial and Vocal Expressions of Emotion. *Annual Review of Psychology*, 54, 329-349. doi: 10.1146/annurev.psych.54.101601.145102
- Sanchez, K., Oates, J., Dacalis, G. et Holmberg, E.B. (2014). Speech and Voice range profiles of adults with untrained normal voices: methodological implications. *Logopedics Phoniatrics Vocology*, 39, 62-71. doi: 10.3109/14015439.2013.777109
- Sataloff R.T. (1995). G. Paul Moore Lecture: Rational thought: the impact of voice science upon voice care. *Journal of Voice*, 9, 215-234. doi: /10.1016/S0892-1997(05)80230-9
- Scherer, K.R. (1994). Evidence for Universality and Cultural Variation of Differential Emotion Response Patterning. *Journal of Personality and Social Psychology*, 66, 2, 310-328. doi: 10.1037/0022-3514.66.2.310
- Scherer, K.R. (1995). Expression of Emotion in Voice and Music. *Journal of Voice*, 9 (3), 235-248. doi: 10.1016/S0892-1997(05)80231-0
- Scherer, K.R. (2003). Vocal communication of emotion: a review of research paradigms, *Speech Communication*, 40, 227-256. doi: 10.1016/S0167-6393(02)00084-5
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44, 695-729. doi: 10.1177/0539018405058216
- Scherer, K.R. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Computer Speech & Language*, 27(1), 40-58. doi: 10.1016/j.csl.2011.11.003
- Schmidt J., Janse, E. & Scharenborg, O. (2016). Perception of Emotion in Conversational Speech by Younger and Older Listeners. *Frontiers in Psychology*, 7, 781. doi: 10.3389/fpsyg.2016.00781
- Smith, C.A., Ellsworth, P.C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, 48, 813–838. doi: 10.1037/0022-3514.48.4.813
- Soderstrom, M., Reimchen, M., Sauter, D. & Morgan, J.M. (2017). Do infants discriminate non-linguistic vocal expressions of positive emotions. *Cognition and emotion*, 31 (2), 298-311. doi: 10.1080/02699931.2015.1108904
- Sundberg, J. (1972). A perceptual function of the 'singing formant'. *STL-QPSR*, 13, 61-63.

- Sundberg, J., (1994). Perceptual Aspects of Singing. *Journal of Voice*, 8 (2), 106-122. doi: 10.1016/S0892-1997(05)80303-0
- Tavris, C., Wade, C. (2007). *Introduction à la Psychologie : les grandes perspectives*. Québec : Editions du renouveau pédagogique inc.
- Teixeira, J.P., Oliveira, C., Lopes, C., (2013). Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters. *Procedia Technology*, 9, 1112 – 1122. doi: 10.1016/j.protcy.2013.12.124
- Titze, I.R. (1992). Acoustic Interpretation of The Voice Range Profile (Phonetogram). *Journal of Speech and Hearing Research*, 35, 21-34. doi:10.1044/jshr.3501.21
- Traunmüller, H. et Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults.
- Van Houtte, E., Van Lierde, K., D’Haeseleer, E. et Claeys, S. (2010). The prevalence of laryngeal pathology in a treatment-seeking population with dysphonia. *Laryngoscope*, 120(2), 306-312. doi: 10.1002/lary.20696
- Van Mersbergen, M. (2011). Voice Disorders and Personality Understanding Their Interactions. *Perspectives on Voice and Voice Disorders*, 21(1), 31-38. doi: 10.1044/vvd21.1.31
- Van Mersbergen, M. et Delany, M. (2014). Vocal responses to emotional picture viewing. *Logopedics Phoniatrics Vocology*, 39(3), 99-107. doi: 10.3109/14015439.2013.777108
- Waaramaa, T. & Kankare, E. (2013). Acoustic and EEG analyses of emotional utterances. *Logopedics phoniatrics vocology*, 38 (1), 11-18
- Yao, S.N., Note ; I., Fanget, F., Albuissou, E., Bouvard, M., Jalenques, I., Cottraux, J., (1999). L’anxiété sociale chez les phobiques sociaux : validation de l’échelle d’anxiété sociale de Lieboitz (version française). *L’encéphale*, 25, 429-435.
- Zeman, J., Garber, J. (1996). Display rules for anger, sadness, and pain: It depends on who is watching. *Child Development*, 67, 957-973. doi : 10.2307/1131873